Review

Journal of the Brazilian Chemical Society

https://dx.doi.org/10.21577/0103-5053.20240159

J. Braz. Chem. Soc. **2024**, *35*, 10, e-20240159, 1-18
©2024 Sociedade Brasileira de Química

Drug Repositioning and Artificial Intelligence: Is It a Promising Approach to be Used for Neglected Diseases?

Fernanda Bongiovani,[a] Marcos A. Ferreira-Junior,[a,b] Soraya S. Santos,[c] Jeanine G. Vargas[a] and Elizabeth I. A. Ferreira[,a]*

[a]*Laboratório de Planejamento e Síntese de Químioterápicos Potencialmente Ativos em Doenças Negligenciadas (LAPEN), Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, 05508-000 São Paulo-SP, Brazil*

[b]*Aché Laboratórios Farmacêuticos S.A., 07034-904 Guarulhos-SP, Brazil*

[c]*Instituto de Ciências da Saúde, Universidade Paulista, 11013-551 Santos-SP, Brazil*

Drug repositioning involves the use of a determined drug in a different indication and has been widely used for some therapeutic classes. Artificial intelligence, in turn, has been a trend in the modern world of innovation, including in drug design. Those approaches are apparently paradoxical (the former is not considered as properly an innovative method to introduce new molecules/drugs into the market, although the latter is) can be used as complementary. In this review, we present some concepts of both methods, their advantages, possible disadvantages, and some applications in drug design in general to improve different aspects of drug development. Examples of the use of both methods together have been given for many therapeutic classes. Notwithstanding, their application in the search for drugs for neglected diseases, although somehow stimulated, deserves more discussion, mainly in light of the social aspects of these infections.

Keywords: drug design, drug discovery, medicinal chemistry, artificial intelligence, drug repositioning, neglected diseases

1. Introduction

According to International Union of Pure and Applied Chemistry (IUPAC), "Medicinal chemistry is a chemistry-based discipline, also involving aspects of biological, medical and pharmaceutical sciences. It is concerned with the invention, discovery, design, identification and preparation of biologically active compounds, the study of their metabolism, the interpretation of their mode of action at the molecular level and the construction of structure-activity relationships".[1] This highly multidisciplinary area, it depends on Chemistry, Biology, Physics, Pharmacology, Toxicology, and their subareas, has substantially contributed to the introduction of new drugs in therapeutics.[2] The development of the correlated areas, mainly in the twentieth century, reflects the evolution of Medicinal Chemistry. The twenty-first century assisted a rapid advance in the area with the introduction of many computational tools (Figure 1).

The development of the correlated fields, mainly in the twentieth century, contributed to the evolution of Medicinal Chemistry. The twenty-first century assisted a fast advance in the area with the introduction of many computational approaches, the most challenging being Artificial Intelligence (AI)[2] (Figure 1).

AI encloses advanced technologies, which in conjunction, allows computers to perform several functions, using adequate algorithms. This technology has been applied to several areas, including drug design, reducing time and financial resources substantially to introduce innovation in the field.[3]

Considering the basic phases comprehended in the introduction of new drugs in the therapeutics (Figure 2), Medicinal Chemistry involves the essential paths toward the preclinical studies of molecules, whose main goals are to identify drug candidates for further clinical development.[4]

1.1. Evolution of medicinal chemistry-brief historical aspects[5]

Alfred Burger, in 1970, when talking about the evolution of Medicinal Chemistry over the years, stated: "The great

*e-mail: elizabeth.igne@gmail.com
Editor handled this article: Hector Henrique F. Koolen (Associate)

This review was written in honor of Prof Eliezer Jesus Barreiro, a great pharmacist who dedicated his life to the development of Medicinal Chemistry in Brazil, strongly contributing to the search for "green and yellow" drugs.

This is an open-access article distributed under the terms of the Creative Commons Attribution License.

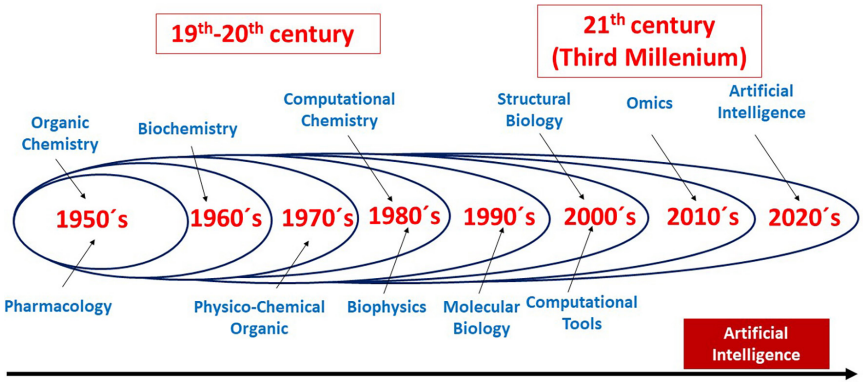


Figure 1. Development of related fields of Medicinal Chemistry from 19th to 20th centuries up to 21th century (third Millennium).[2]

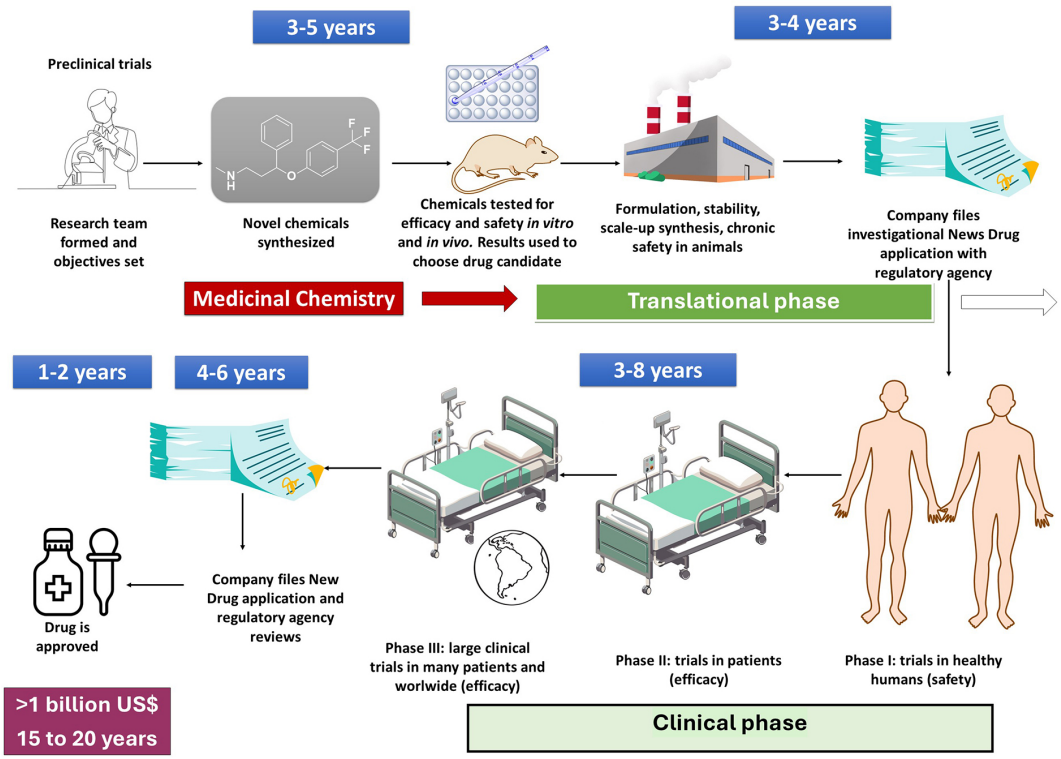


Figure 2. Phases related to the introduction of new drugs in the therapeutics.[4]

advance of medicinal chemistry has been achieved by two types of investigators: those with the genius of prophetic logic, who have opened a new field by interpreting correctly a few well-placed experiments, whether they pertained to the design or the mechanism of action of drugs; and those who have varied patiently the chemical structures of physiologically active compounds until a useful drug could be evolved as a tool evolved as a tool in medicine."[5] Based on this statement, it is important to analyze the ideas and concepts from the point of view of their evolution that led to the knowledge available in the present.[5]

Starting with Antiquity, we must look back to the ancient people, such as the Greeks, Chinese, Hindus, and the Mayans, from Central America, for example, and the use of natural products. This was extended to the Middle Ages, with emphasis on Paracelsus, around the 15th and 16th, whose history is related to the cure power of antimony and its salt, thus, emphasizing the potentiality of chemicals to heal diseases. The 19th century was considered the Age of Innovation and Chemistry. Several active principles were then extracted, with cocaine, morphine, caffeine, and physostigmine as the most important examples. The 20th century witnessed the birth of the Pharmaceutical Industry, triggered by the advances in correlated areas as shown in Figure 2. The introduction of many chemotherapeutic agents at the beginning of the century, such as sulfonamides and penicillin, to cite the most important ones, was a great impulse for the discovery of many potent anti-infective

agents, changing the paradigm of medical practice. Drugs from several classes were discovered and had their chemical structures modified to improve their activity and other properties. Agents for central nervous system diseases, endocrine therapy, steroids, and drugs affecting renal and cardiovascular functions were further introduced. The lessons from World Wars I and II led to the discovery of some anticancer agents and the molecular studies about their actions stimulated the search for antineoplastic agents with advanced mechanisms of action. Meanwhile, it is also important to point out that the limits of drug discovery and development in terms of costs and time-consuming, give rise to the use of an approach, that has the objective of discovering new indications of drugs already in the market, named drug repositioning. Processes of medicinal chemistry can be used in conjunction with repositioning, decreasing the time and money spent by Pharmaceutical Industries.[5,6]

Genomics constituted the next wave of drug discovery, in combination with other approaches such as high-throughput screening and medicinal chemistry. The pharmacogenetics that arose from the latter wave changed the scenario and opened doors for discovering particular medicines for specific patients. It is worth mentioning the high relevance of the discovery of molecular targets that can explain how drugs work at the cellular level, tissues, and organs, helping the researchers of the area to find more potent and advanced molecules for many diseases still waiting for specific drugs.[5]

The development of highly sophisticated technology, AI being the most relevant and challenging, could aid the design of new molecules, and the 21st century will certainly change the pattern of drug design. AI encloses advanced techniques, which in conjunction, allow computers to perform several functions, using adequate algorithms. This technology has been applicable to several areas, including drug design, reducing time and financial resources substantially to introduce innovation in the field.[7]

It is worth highlighting that, since the very beginning, the insight, determination, knowledge, and ability of the researchers in managing the molecules were decisive in giving them higher potency to cure most of the illnesses already discovered and others that are still waiting for proper therapy. Although AI could enhance productivity in terms of drug design, it might not be overestimated compared to the contribution of the scientists.

In Brazil, groups are working on Medicinal Chemistry, some of them with the aid of AI. Notwithstanding, Barreiro's group (LASSBio-Laboratory of Evaluation and Synthesis of Bioactive Substances, created in 1994 and coordinated up till now by Prof Eliezer Jesus Barreiro)[8] has been using processes of Classical Medicinal Chemistry with elevated success. Agents for cardiovascular, for metabolic for cancer disease, and anti-inflammatories in addition to agents for neglected diseases such as leishmaniasis and Chagas, have been designed.[9] They composed a library of more than 2,000 new molecules,[10] which is being studied in partnership with Eurofarma, a Brazilian Pharmaceutical Industry with a profile of innovation.

It must be emphasized that chemical intuition has been strongly considered by Pedreira *et al.*,[11] in their extremely interesting review, as a relevant part of Medicinal Chemistry, which includes many drug discoveries. To illustrate, a LASSBio example (LASSBio-579) can be presented (Figure 3). This compound is a new neuroactive drug candidate acting as a D2-receptor agonist. It was designed from clozapine, an atypic antipsychotic, through the process of Medicinal Chemistry known as Molecular Simplification.[12] The strategy was to use it to produce *N*-phenylpyrazone and *N*-phenyl[1,2,3]triazole derivatives through chemistry intuition.

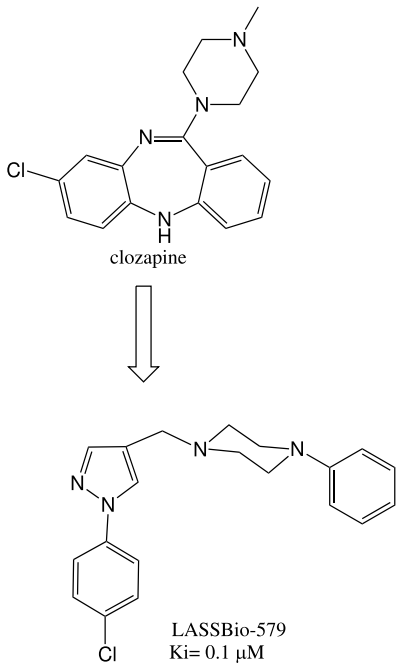


Figure 3. Design of LASSBio-579 from clozapine.

2. Drug Repositioning: Concept, Applications, Advantages and Disadvantages

Ashburn and Thor[13] raised concerns about the need to grow the productivity of biopharmaceutical industries, as the discovery of new compounds was proving to be an increasingly difficult and expensive process. Thus, they brought drug repositioning as a promising concept to modify this challenging scenario.[14] As defined, "drug

repositioning" or "drug repurposing" comprises the use of an active pharmaceutical ingredient, which has a defined biological activity (an "old drug"), for a new therapeutic indication.[13-15]

According to Jourdan *et al*.,[15] to apply the concept of drug repositioning, there is no structural modification of the bioactive agent. Thereby, there is a new therapeutic indication based on the molecular bases of the disease, through a deep knowledge of the human genome and, consequently, regarding biological targets. Possibly, the dose, formulation, and route of administration can be modified, however, a preliminary understanding of pharmacokinetic and pharmacology properties greatly assists the drug development process.[15]

There are different strategies for drug repositioning, such as observation of unexpected side effects during clinical trials, investigation involving phenotypes of specific diseases, computational approaches (signature matching, molecular docking, dynamics, genome-wide association studies, pathway mapping, artificial intelligence, among others), which can also result in the discovery of a new drug target.[16,17]

This interesting strategy allows the designing of novel compounds with low cost, shorter timeline, and high efficiency since the previous knowledge of candidates can be considered in the development,[16-18] Additionally, facilities such as bioinformatics databases for drug candidates (Entrez-Gene, DrugBank/Drug Central/PubChem) have greatly assisted the practice, since it is possible to elucidate promising targets for ligand interactions, without major experiments.[15] Traditional steps required for approval by the Food and Drug Administration (FDA)[19] may not be necessary. Furthermore, phase I of clinical trials on humans may be discarded, resulting in significant time and financial benefits.[20]

It is important to highlight the estimated cost of launching a new drug on the market: 314 million to 2.8 billion dollars, in total of 12-17 years, considering the classic steps of development, for example, drug design, and pre-clinical and clinical studies.[18,20,21] However, with the repositioning strategy, the time can be reduced to 3 to 12 years, with the great advantage of a considerable reduction in cost, making it a very attractive approach for the pharmaceutical industries.[18] Considering the above-mentioned, 30% of new FDA-approved drugs and vaccines in recent years come from drug repositioning method.[22] There are many examples of successful drugs repositioned in therapy, such as zidovudine, originally utilized for cancer treatment and now is applied in human immunodeficiency virus (HIV) infection; ketoconazole is an antifungal agent repositioned to treat Cushing syndrome; aspirin was developed for analgesia and has a new indication for colorectal cancer; sildenafil was designed for angina and the new indication is erectile dysfunction.[14]

It is also essential to emphasize some disadvantages of the drug repurposing methodology. The patent application can be a great concern since the inventor can protect the discovery of any new process until the final deadline of the document. Therefore, if the selected repositioned compound is under patent application, there is a real problem to advance in the drug development phases.[17] Although, it is possible to encourage a partnership with the companies that own the invention, which may benefit from this new use.[23] Moreover, there is a lack of professionals who deeply understand the legal issues related to the repositioning of pharmaceutical ingredients, which can influence negatively the evolution of the new application.[24]

Oprea *et al*.[24] highlighted the difficulties regarding dosing and safety, as well as the lack of integration with pharmaceutical sciences and toxicology. The former concern is mainly associated with safety aspects, since there is a new target and the dose required for the biological effect may be different, consequently, the exposure to the active component may be greater. The latter is concerning new formulations and delivery mechanisms to avoid unexpected toxic effects *in vivo*.[24]

Considering project funding, the difficulty is evident especially when the drug candidate shows a problem in their original indication, which generates concerns regarding future applications. In this scenario, some funding agencies do not sponsor the proposal.[17]

In academia, there are several applications of drug repositioning, and a few examples will be briefly described below. Bayraktar *et al*.[25] carried out a careful search on Alzheimer's disease, analyzing the main genes expressed in this disorder, and aiming to find a promising therapeutic target. As a result, they identified glutaminase, since its overexpression may be related to several conditions, causing, for instance, neurodegeneration. Through repositioning, they studied eight compounds that could be interesting for interaction with glutaminase, highlighting bortezomib and parbendazole (Figure 4), which demonstrated interesting results, including *in silico* parameters evaluated by SwissADME.[26]

Ovarian cancer represents a global health problem and the treatment includes particularly toxic drugs. An interesting activity-based drug repositioning strategy, which considered 54 FDA-approved molecules capable of inhibiting the viability of human epithelial ovarian cancer (SKOV-3 cells), selected disulfiram (Figure 4) as the most promising. The original biological activity of this compound is related to the control of alcohol abuse.

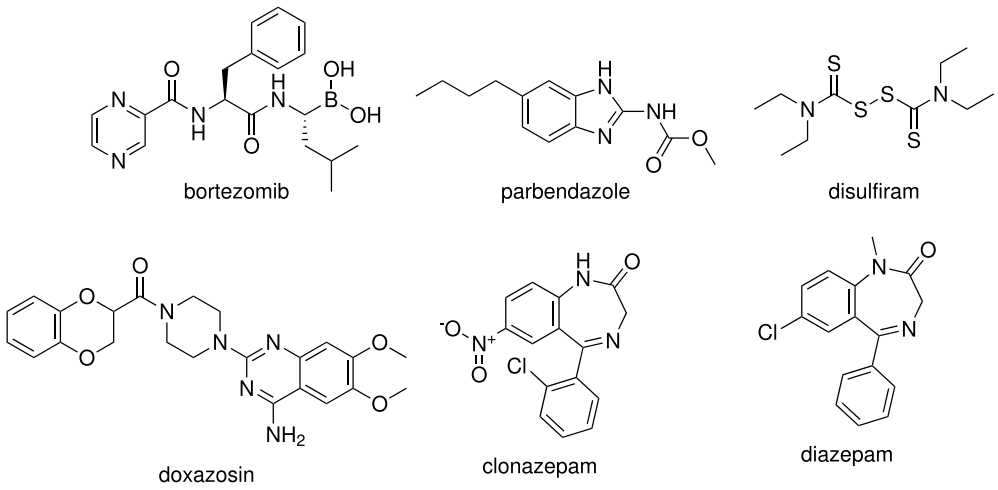


Figure 4. Chemical structures of some repurposed drugs in academia.

Interesting findings showed, among other results, the ability of this drug to decrease the expression of anti-apoptosis B-cell lymphoma/leukemia-2 (Bcl-2). Moreover, the treatment with disulfiram provided an increased level of proinflammatory cytokines in mice models. The complex with copper was also evaluated, revealing, for example, a better inducer of reactive oxigen species (ROS) production in cancer cells. *In vivo* data for disulfiram and copper gluconate were even more positive, indicating decreased tumor size and survival rate in a murine model. Therefore, the repositioned action of disulfiram with copper may represent a hopeful future for the treatment of ovarian cancer.[27]

The action of α-adrenergic blocker, doxazosin (Figure 4), on *Proteus mirabilis* and *Pseudomonas aeruginosa* was studied by Elfaky *et al*.[28] The authors developed experiments to evaluate, especially, the action against virulence, using, for example, the assessment of anti-biofilm activity and the effect on bacterial motility, quantification of the expression of virulence-controlling genes, analysis of anti-proteolytic activity, among other investigations. Important *in vitro* and *in vivo* findings showed the capacity of the replaced drug to act as an anti-virulence compound, through the reduction of virulence enzymes, pigments, biofilm formation, and motility. Therefore, initial studies demonstrated the potential of doxazosin in this area of great importance for human health.[28]

Considering infectious diseases caused by the ESKAPE group, named *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *and Enterobacter* spp., there is an important investigation regarding the benzodiazepines repurposing. Thus, clonazepam and diazepam (Figure 4), in combination with ciprofloxacin, were evaluated for their antibacterial activity. Some examples of experimental assays performed were: the determination of minimum inhibitory concentration (MIC), determination of minimum bactericidal concentration (MBC), checkerboard synergy assay, and determination of fractional inhibitory concentration index (FICI), among others. The MIC results were 64 to 1024 μg mL^{-1} for ciprofloxacin; 128 to 256 μg mL^{-1} for clonazepam and 1024 to 2048 μg mL^{-1} for diazepam. The combined drug assay showed interesting findings: 128 to 32 μg mL^{-1} and 64 to 32 μg mL^{-1} (ciprofloxacin-clonazepam and ciprofloxacin-diazepam, respectively). Moreover, considering almost all tested isolates, synergistic effects were observed in the checkerboard and FICI assays. Therefore, these preliminary repurposing results may be useful for future drug design for infectious diseases caused by ESKAPE.[29]

3. Artificial Intelligence: Concept, Application, Challenges and Recommendations

As mentioned before, the search and development of a new drug can take up to 12-15 years, and several resources, from labor hours to reagents, to money spent on this investment (Figure 2). The high cost and consumption of human resources to develop these new treatments have led researchers to try and figure out new ways to reduce these expenses and to save time, allowing these resources can be rearranged to other areas of expertise.[19]

Several tools have been created to facilitate this process, e.g., molecular docking, virtual screening, high-throughput screening, chemical synthesis, and many others, although they still take a lot of time to create and analyze these data. This is where artificial intelligence (AI) comes into play. The idea that computers can be employed to simulate intelligent behavior is something that researchers have been exploring since its creation. In 1950, Alan Turing in his seminal paper *Computers and Intelligence*[30] thought

of a test that would see if computers were able to mimic human intelligence, thus creating the "Turing test". In the beginning, AI was resumed as a rule-based system, with time, more intricate algorithms were developed which allowed a subdivision in the field, such as machine learning, applied to identify and analyze patterns whilst the machine uses the data provided to improve itself; deep learning, a multi-layered neural network that allows machines to learn and to make decisions, ensuring that the model presents the ability to mold and change itself; natural language processing, employed to make decisions based on the information drawn from human language; and computer vision, where the computers analyze and acquire information from videos and images.[31] In this context, the development of machine learning and deep learning tools may be useful the drug discovery, often used to help in identifying the pharmacological properties of different compounds, which may act in several pathologies and molecular targets.[32] These tools have shown great potential to accelerate and improve many aspects of drug discovery, especially in the synthetic planning of small molecules and the predictive chemistry field.[33]

Deep learning tools, especially convolutional neural networks, as stated before, belong to a subgroup of machine learning. The neural network structure is organized in multiple layers and uses images, formulas, and patterns as a way to "learn" and develop itself, allowing for a faster resolution of problems and faster pattern testing. This approach is currently being used in the studies of medical images, as well several deep learning algorithms are being developed, which may allow these tools to become part of the routine in clinical analysis in the near future.[34]

Thereby, with machine learning technologies and the knowledge of molecular target properties, it is possible to create and develop artificial intelligence tools that may help in the prediction of compatibility of a certain molecule with its target, correlating its physicochemical properties with the target characteristics, allowing researchers to see if an already available and commercialized drug may be compatible with targets of interest from different diseases. AI can be employed to integrate a series of heterogeneous data to discover patterns to understand at the molecular level the mechanisms by which a pathology or molecule works, helping in the identification of a target. It can also be used in the development and generation of lead molecules and their optimization through score function and quantitative structure-activity relationships (QSAR). Therefore, the use of AI tools accelerates research development through the extraction of new and important information, based on the large amount of data generated by the drug discovery process.[35]

One of the greatest advantages of using AI instead of more traditional approaches, such as virtual screening and molecular docking, is that it can be programmed to be more a self-sufficient tool, which could eliminate possible challenges faced by conventional tools. Deep learning and Machine Learning algorithms are already being used in several steps of the research of new drugs, from structure and ligand-based virtual screenings, synthesis of peptides and small molecule design, to toxicity and drug dosage prediction, drug repositioning, and much more.[33]

These computational techniques may even assist in the search for new drugs for the treatment of neglected diseases by facilitating the identification of proteomic, genomic, and transcriptomic targets, and the interaction they may have with potential drug candidates.[36]

Another example is the development of different methods for drug repositioning using AI such as the one developed by Lei *et al.*[37] named VGAEDR, which proposes the use of a heterogeneous network of multiple attributes associated with drugs and an autoencoder graph (VGAE) to predict associations between molecules and pathologies. The tool built a drug-disease heterogeneous network based on various correlations, considering the drug properties, the disease features, and the association between drug-disease. The VGAEDR module is divided into two parts, the VGAE one receives a heterogeneous network as an input, then it learns and extracts its redimensioned representation, and in a multilayered convolution module, which perfects the learning on top of the extraction obtained by the VGAE. The association of drug-disease is then predicted by the data provided to this neural network.[37]

Although AI can decrease the time spent on drug development and can reduce the cost of this process, the use of machine learning and deep learning tools still have a lot of challenges. In order to begin creating such a tool, it is necessary to understand and have sufficient information about the target molecule and/or the bioactive molecule of interest, there needs to be accurate and well-defined data to feed the AI tool. Without the correct and validated information, the results may be flawed and misleading, which could lead to wasted time and resources. The labeling of information needs to be complex and as complete as possible to reflect the nature of drugs in biological systems, and therefore, it is necessary to be an investment in the understanding of the drug mechanisms.[38]

There is also a need for well-trained human resources to create such tools, and to understand what kind of data needs to be employed in the construction of the algorithm. The cost of developing such software is still on the higher side, demanding a great amount of processing power and

competent people to operate it. Furthermore, the available data is not always sufficient to create a reliable tool. Often, there is not enough information regarding the nature of the target molecule, the disease pathophysiology being studied, and how the target molecules interact with the lead molecule in the study. Another problem is the type of available data, which usually is created in various formats that are not compatible with one another, which makes the correlation between them virtually impossible. In this case, human interference is required to refine the data used to feed to the algorithm and to proofread the findings to check their consistency and validity. This takes up a lot of time and may also lead to human error or bias.[35,39] Finally, AI tools bring many benefits to the search for new drugs, such as time and resource savings, though only when made with good and validated data through a well-developed algorithm.

4. Advantages and Challenges of Employing Artificial Intelligence for Drug Repurposing

AI is playing a growing role in identifying potential drug repurposing opportunities. In this section, the major advantages, and challenges of using AI for drug repurposing are discussed.

4.1. Opportunities

AI enables a systematic approach for generating drug repurposing hypotheses through the development of predictive computational models that can be employed for screening large compound libraries of potential candidates. Different successful drug repurposing cases described in the literature, such as the landmark repurposing of sildenafil, were based on serendipitous findings from clinical trial data or experimental testing using *in vitro* or *in vivo* models.[7] Although successful, those approaches are more difficult to reproduce on a large scale and they present intrinsic challenges, including the need to have physical compound libraries, the required work for high throughput assay development, and the labor and materials for running the screening campaigns, which can require significant investments and time.[40,41] Therefore, the use of AI is a promising alternative to reduce the time and cost of generating hypotheses for experimental validation, which are advantages aligned with the major goal of the drug repurposing field.[7]

AI provides the possibility of analyzing very large databases containing chemical and biological data to identify similarities and interactions between compounds, biomacromolecules, or disease conditions, which can be employed to generate the repurposing hypotheses. This capacity is relevant considering the large and increasing amounts of biomedical data publicly available with information on the disease and drug profiles.[42] This scenario is the result of novel high-throughput technologies that allow the collection of vast amounts of chemical and biological results, such as multi-omics (genomics, transcriptomics, proteomics, metabolomics), compound chemical characterization, drug-biomacromolecule interaction, phenotypic endpoints from traditional and high-content screening methodologies, clinical trial data (therapeutic and side-effects) and real-world clinical findings (e.g., electronic health records).[40-43] In this context, Wang *et al.*[44] developed a deep learning framework to integrate large-scale heterogeneous multi-omics data (genomics, transcriptomics, and epigenomics), chemical properties, and drug-target interaction information to predict the response of cancer cells to drugs. Liu *et al.*[45] described a deep learning framework for drug repurposing using real-world patient data from electronic health records and insurance claims to predict the effect of the drugs on the disease outcome. The model was used to find new therapies and combinations for the treatment of coronary artery disease using a dataset with 107.5 million patient records. Exploration of very large databases is now possible as a result of the advances in computational capacity and AI algorithms.[42]

AI brings the possibility of integrating different types of data from diverse sources, which further expands the possibility of generating systematic repurposing hypotheses.[43] Each type of data provides a different aspect of the disease, biomacromolecule, or drug and, therefore, adds another layer to their underlying connections.[46] Data integration enables the development of methods that evaluate the molecule effects not only at a single target level but also its interplay with other biomacromolecules, which share a common biological pathway.[42] Therefore, integrating different types of data can help expand the applicability domain of the models and unravel interplays between diseases, targets, and bioactive compounds.[46]

Amiri *et al.*[47] developed convolutional neural networks to predict new indications for existing drugs following data integration on compound properties (chemical structure, side-effects, and therapeutic target), disease patterns (human phenotype and target protein), and drug-disease associations of 593 FDA-approved drugs.[47] A neural network model was developed to predict new targets and drug repurposing opportunities by embedding chemical, genomic, phenotypic, and cellular network data to generate a heterogeneous drug-gene-disease framework.[48] Furthermore, a deep learning model was developed for

drug repurposing containing chemical structure and drug-target interaction to capture off-target profiles for model development. Although there are several reports in the literature on the AI application for drug repurposing, there are important challenges that should be considered for any endeavor in the field, which depend either on the biomedical data or on the model development and validation.[49]

4.2. Challenges and recommendations

There are large quantities of experimental findings, such as data generated by high-throughput deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) sequencing, mass spectrometry, metabolomics, transcriptomic data, phenotyping, and also clinical data that are increasingly becoming available from electronic health records. These data sets are so large or complex that traditional data processing methods are inadequate. In this context, AI is an important tool for improving our understanding of disease and developing strategies for drug design. It is crucial to highlight that there is a gap between our ability to generate big biomedical data and our ability to integrate, analyze, and interpret the data. However, there are many results, such that much-generated data are unstructured and heterogeneous, such as imaging and structural findings, which increase the complexity integration. There is an urgent need for technology solutions that can combine heterogeneous data sets and integrate, analyze, and interpret them. Another challenge is access to different types of databases. Some are publicly available databases for transcriptomic data, while others are limited and rare, such as clinical trial data and structural, *in vitro*, or imaging data. Therefore, it is important to have publicly accessible and standardized databases. Other challenges that can emerge are related to patent and regulatory considerations besides organizational hurdles.[14,50-53]

The performance of AI models depends directly on the availability of high-quality and comprehensive experimental data for modeling. Despite the huge expansion in public biomedical information, it is well known that the data can be incomplete and even contain noise, bias, and inaccurate results. Noise can be a product intrinsic to the methodology, from human error or biological variability. Bias can arise from the selection of a particular subset of compounds for testing or from uncontrolled experimental conditions that were not acknowledged during the data report. Limited negative results can also interfere with the development of robust AI models.[43,54,55] For this reason, careful consideration of the data quality must be taken to guarantee that reliable and relevant models are generated. In this context, there are initiatives from the scientific community to standardize the reporting of biomedical data, such as the FAIR Principles, which aim at improving data reusability for *in silico* modeling.[56]

Besides data quality, AI applicability for drug repurposing is limited to the information that can be extracted from the existing data. The applicability domain of AI models is determined by the structural and pharmacological diversity of the training set.[57] Therefore, AI-generated drug repurposing hypotheses are subject to bias, since there is more published information and databases for a selected set of drugs, while limited information is available for other compounds.[43,58] In this sense, more diverse datasets and data integration are promising strategies to expand the applicability domain by enabling novel connections between drugs, biomacromolecules, and diseases, thus providing the identification of a wider variety of drug repurposing candidates.[46]

Data integration from different sources can also be challenging due to the high heterogeneity of the chemical (structure, physicochemical properties), biological (multi-omics, phenotypic profile), and clinical results.[46] In addition, integration can be difficult owing to inconsistencies, sparse data, variations in terms of format and annotation, a variety of experimental conditions, and the level of agreement of equivalent endpoints from different sources.[53,54] This scenario demonstrates the need for extensive manual data curation, which prevents more widespread use of AI for drug repurposing[14] Similarly, there are initiatives focusing on harmonizing data from different databases to facilitate its use for computational models.[59]

Even with robust and comprehensive biomedical data available, the incomplete understanding of the biological systems and the complex and multifactorial biology of diseases makes it challenging to employ AI to predict the therapeutic potentials of existing drugs.[60] AI algorithms cannot completely model the complex interplay between drugs, biomacromolecules, and disease conditions. Moreover, most biomedical results are generated using *in vitro* assays that evaluate proxy endpoints of human clinical efficacy, safety, and pharmacokinetics. These results only capture partial aspects of the disease biology and thus present limited translation to the effect that will be observed in humans.[61] In addition, there are aspects of drug pharmacokinetics and pharmacodynamics that are not usually considered during hypothesis generation, which may have a significant impact on the safety and efficacy profile of the compounds. Those features include, for example, different pharmacokinetic properties and toxicity for a novel target population (e.g., a drug approved for adults may have a completely different safety profile in a special population, such as the elderly) and for different disease conditions.[62]

In addition to the quality of the data set, the performance of AI models depends on the AI framework and proper model development and validation. Several neural network structures exhibit superior performance for specific data formats, such as convolutional neural networks for image-like data and recurrent neural networks for data available as sequences.[63] Therefore, the development of AI models for drug repurposing can still be challenging and require extensive work for validation. AI models tend to result in overfitting, in which predictive performance is high during training and internal validation, though significantly decreases for unseen data, thus leading to erroneous predictions and limited applicability in real-world settings.[43] This may be a result of the limited amount of data and the complex model architecture.[63] Some strategies can be employed to minimize overfitting, which may involve, for example, removing layers from the network or terminating the training phase earlier.[64] Additionally, the limited availability of reliable benchmarking databases for model validation makes it difficult to compare the performance reported for different frameworks in the literature.[65] Moreover, deep learning neural networks still lack interpretability, making it challenging to evaluate the variables that impact the prediction scoring and, thus, resulting in a "black box" computational model.[66] Although the above-mentioned strategies for further validating the AI model outputs are promising, experimental validation using preclinical experiments before human testing is paramount in any drug repurposing initiative.[14,53]

AI algorithms coupled with large databases of biomedical data can be a promising approach to generate hypotheses on drug repurposing opportunities. The selection of the most adequate algorithm for each case should consider the type, quality, and amount of available data. Although AI has limitations, it can provide strong drug repurposing hypotheses to support experimental testing.[44]

Pushpakom *et al.*,[14] in 2019, describe some recommendations to help realize the drug repurposing, such as the need for better integrative platforms for data analysis, as well as the access and integration, remain a bottleneck, particularly for findings from industry-sponsored phase II-IV clinical trials, especially for discontinued drugs, which may open the opportunity of repurposing through external searches. Another recommendation is to carry out novel safety assays to repurposed drugs, due to the new interactions between the drug and the disease for which it is repurposed, use in new populations, or differences in the dosing schedule. More incentives are needed to finance projects involving the repositioning of drugs, which can also happen through new financial sources, such as crowdsourcing and parent entrepreneurs, especially for neglected diseases, which do not arouse the interest of the pharmaceutical industry in researching new drugs. Another recommendation is to share libraries of bioactive compounds. Lastly, actions are required regarding the patent and regulatory barriers, which could include better data exclusivity time for repurposed indications, royalty arrangements between companies, or other legislative changes to ensure the retrieval of investment from drug repurposing programs.

5. Drug Discovery and Development for Neglected Diseases: Use of Drug Repositioning and Artificial Intelligence

According to World Health Organization (WHO),[65] "Neglected Tropical Diseases (NTDs) are a diverse group of conditions caused by a variety of pathogens and associated with devastating health, social and economic consequences". The tendency nowadays is to consider NTDs as diseases of neglected populations, as they are prevalent in poor communities in tropical areas, where the conditions are normally dreadful. In general, the interest of governments in improving the health situations of those areas is missing, as the people involved do not have political voices to change the scenario. NTDs affect more than 1 billion people, and comprise more than 20 diseases, such as: Buruli ulcer, Chagas disease, dengue and chikungunya, dracunculiasis, echinococcosis, foodborne trematodiases, human African trypanosomiasis, leishmaniasis, leprosy, lymphatic filariasis, mycetoma, chromoblastomycosis and other deep mycoses, noma, onchocerciasis, rabies, scabies and other ectoparasitosis, schistosomiasis, soil-transmitted helminthiases, snakebite envenoming, taeniasis/cysticercosis, trachoma, and yaws. Malaria and tuberculosis have not been included as NTDs given the investments that have been made in these diseases. Nevertheless, they have been considered diseases of neglected people. Table 1 shows some of the diseases of neglected people and their distributions around the world. Nevertheless, if one considers the populations needing preventive and curative interventions, this number reaches 1.6 billion, which corresponds to about 5% of the world population.

NTDs normally have a very complex epidemiology, being closely related to the environment.[65] Many are vector-borne diseases and have complex life cycles and animal reservoirs. These characteristics make them difficult to control, which implies huge challenges for the public health of the countries affected. Also, NTDs are responsible for high levels of morbidity and mortality. Although many technological developments have been used

Table 1. Information about some of the most relevant disease of neglected people

Neglected Tropical Diseases	Global prevalence / millions	Population at risk / millions	Annual mortality / thousands	Regions of high prevalence in the World
Malaria	249	ca. 4,000 (ca. 50% world population	608	Sub-Saharan Africa,[a] Asia, South and Latin America, Middle East, and Pacific Islands
Schistosomiasis	240	700	200	Africa, Middle East, South America, Caribbean, China, Southeast Asia, Cambodia, Laos, Central Africa
Tuberculosis	10.6	1,000	1.5	India, Indonesia, China, Philippines, Pakistan, Nigeria, Bangladesh, Democratic Republic of Congo
Leishmaniasis	12	1,000	30[b] 1,000[c]	India, South Asia, Sub-Saharan Africa, Latin America, Caribbean, and Mediterranean region
Chagas disease	6-8	75	0.7	Latin America, Caribbean

[a]94% of cases, 95% of deaths-78% children under 5 years old; [b]visceral leishmaniasis; [c]cutaneous leishmaniasis.

in drug discovery, in general, this does not happen with NTDs. Some reasons contribute to this situation, such as the lack of interest of pharmaceutical industries, motivated by the cost of developing new drugs[66] and the implicit low financial revenue. Furthermore, in most of the poor countries involved, there is no access to new technologies, which might trigger the development of more effective drugs for those diseases.

Although this scenario has changed over the years, with successful partnerships between international organizations, such as WHO, Drugs for Neglected Diseases Initiative (DNDi), TB (Tuberculosis) Alliance, countries' governments, and some pharmaceutical industries, the financial resources available to tackle the problem are still insufficient, as it can be seen in Brazil.[67] From 2007 to 2020, the investments by the source of funding for product-related Research and Development for neglected diseases were 64.66% by institutions of the public sector; 20.62% by philanthropic foundations, non-governmental organizations, corporate donors; 12.97% by pharmaceutical and biotechnology companies; 0.88 by academic and other research institutions; 0.80% multilateral public sector, and 0.07%, unspecified. The amount accounts for around US$ 50 million, but this corresponds to 14 years of funding, being the investment per year insufficient.[65]

Much might be done to face this challenge. Some approaches, such as drug repositioning, and more advanced processes, such as AI, either separately or together. Nevertheless, the number of works is still reduced and could be increased mainly by the pharmaceutical industry sector.[68,69] However, in the past few years, there have been many works using drug repurposing in experimental or computational high throughput screenings, searching active agents against different neglected diseases.[70] These studies may identify several compounds already used in the therapeutics.

Notwithstanding, it is important to consider that computational methods are becoming more and more relevant in medicine, due to the existence of large databases that allow their use in AI, and it looks like a promising approach for the near future. Works about AI use in neglected diseases comprise not only therapeutic but also diagnosis and predictive objectives.[71] As already discussed, despite being an advantageous approach, there are some drawbacks, being accuracy and availability of good quality databases among the major ones. These limitations recommend care, as there will be security risks, involving social variables among others. Therefore, studies to overcome these limitations must be focused on these concerns resolution, which anticipates its more rational use in the future.

Some neglected diseases are awakening more interest in the use of AI approaches alone and sometimes together with repositioning methods. Advances have been registered in malaria works and they depend on the improvement of computational algorithms, along with the enhancing knowledge of the biology and biochemistry of the parasites, which is essential to repositioning. One of the threats of malaria has been the resistance of the parasites to the first-line drugs. In addition to many others, the extrapolation of the *in vivo* studies in experimental animals to clinical trials has been still a gap in the research.[36] American trypanosomiasis, or Chagas disease was considered the first use of machine learning for a neglected tropical disease.[72] It was used to solve a formulation problem of benznidazole, one of the two drugs available for this neglected disease. The problem was the low water solubility and the use of machine learning for a neglected tropical disease allowed

to have effective higher solubility microparticles, being the oral absorption was enhanced. Formulations containing chitosan microparticles were made to improve the water solubility of benznidazole. Artificial neural networks were employed to evaluate the influence of process parameters such as encapsulation efficiency, size, yield, and dissolution. A neural network involving three layers was developed: input, hidden, and output layers. The output layer was designed with four neurons, corresponding to the size, encapsulation efficiency, yield, and dissolution rates. The input layer was related to polymer concentration, NaOH concentration, stirring rate, and spraying rate. A multi-response optimization was then applied to obtain minimum size and maximum throughput, encapsulation efficiency, and dissolution rates. The influence of predicted parameters to improve the benznidazole solubility were: polymer concentration, 1.5% (m/v), NaOH concentration, 6.0% (m/v), stirring rate, 1400.0 rpm and spraying rate, 5.0 mL min^{-1}. These conditions were reproduced in the laboratory and showed to be the ideas for preparing the formulation containing benznidazole in chitosan microparticles. Thereby, the formulation was very efficient in reducing particle size and increasing the encapsulation efficiency, yield, and dissolution rate of benznidazole.

Guerra *et al*.,[73] in 2013, developed with the aid of AI, a study about neural network application to 72 compounds of a in house data set,[74] comprising imidazolidines arylhydrazones, *N*-oxides of benzimidazole, indazole, quinoxaline and benzofuroxan 5-nitrofuryl semicarbazones, coumarins and nitrofurazones. Those researchers achieved a model using a CODES/RD Program,[75] which was considered able to get "outstanding predictive results". It is worth mentioning that by this program each molecule is codified into a set of numerical parameters based on topological information of each chemical structure. They report the study as the first model for the prediction of trypanocide activity among heterogeneous series of organic compounds based on the strategy of artificial neural networks.

The possible synergy between repositioned drugs against *Trypanosoma cruzi* was explored by Planer *et al.*[76] They found promising results from FDA-approved drugs and screened as for their trypanocidal activity. Most of the active compounds showed (half maximal effective concentration (EC_{50})) in the range of either micromolar or nanomolar concentrations. Through this study they identified clemastine, an antihistamine drug, EC_{50} = 0.4 μM; fluoxetine, a selective serotonin reuptake inhibitor, EC_{50} = 4.4 μM, and pyrimethamine, an antifolate drug, EC_{50} = 3.8 μM, and other drugs as well. Figure 5 shows the structure of these three drugs mentioned above. Assayed in the murine model of *T. cruzi* infection, most of them showed lesser efficacy than when in combination with other classes of drugs. This study was extended to 24 active compounds tested *in vitro* combined with other drugs screened. They concluded that combinations of FDA-approved drugs had a synergistic effect is a promising strategy (polypharmacology) for developing treatments for Chagas disease.

Aguilera *et al.*[77] reviewed the use of polypharmacology in the treatment of Chagas disease. Polypharmacology is a method that has been increasing the efficacy and tolerance of drugs in many diseases. It consists of a combination of drugs either useful in the treatment or repositioned for the treatment of Chagas disease. The authors discuss the employment of ergosterol biosynthesis inhibitors, anti-inflammatory agents, cardiac dysfunction drugs, trypanothione biosynthesis inhibitors, and vitamins, among other drugs. Natural products were also used in the application of this strategy. The most known combination is benznidazole with ergosterol biosynthesis inhibition, namely with posaconazole (Figure 6). Even though this combination failed, polypharmacology deserves to be explored for other classes of compounds. The work[77] shows some examples of this fact.

Computational models, using LBDD (ligand-based drug design) and SBDD (structure-based drug design), with the support of AI, machine learning, were developed by Ferreira Junior,[78] using some specific softwares.[79-83] These approaches have been used for the repositioning of drugs and drug candidates, to contribute to the treatment of Chagas disease. In the LBDD approach, the machine learning algorithms random forest, Naive Bayes (NB), support vector machine, and probabilistic neural networks were used to develop computational models based on

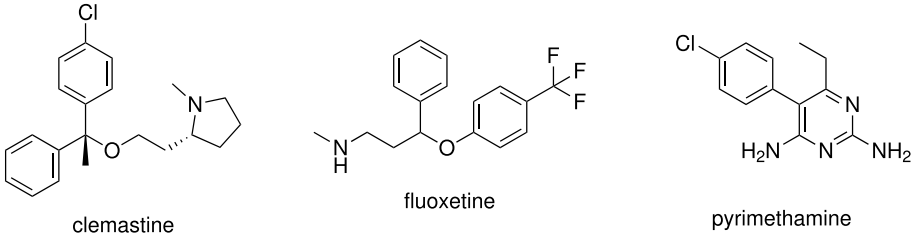


Figure 5. Drugs that can have synergistic effect in Chagas disease, according to Planer *et al.*[76]

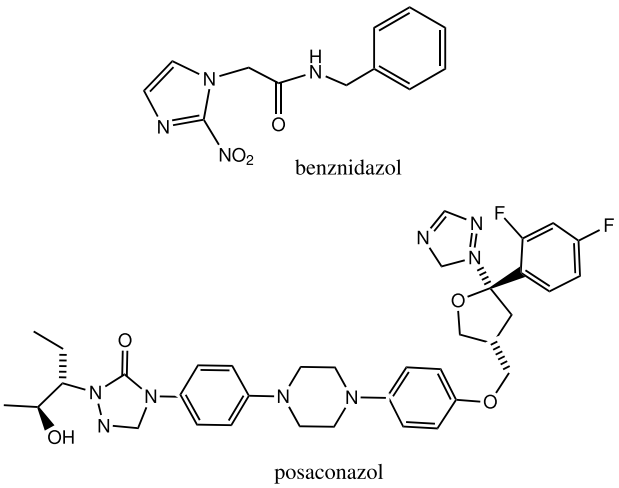


Figure 6. Benznidazole and posaconazole in pharmacological combinations.[77]

phenotypic data of in vitro *T. cruzi* inhibition. The NB-FeatMorgan model showed the best performance, with a high predictive power in the initial validations (5-fold cross-validation and external validation), which was confirmed by two types of validation. 269 drugs comprised the data set and 19 (7.1%) out of the total drugs were selected as potentially active (IC_{50} (drug concentration able to inhibit 50% of the parasite proliferation at the hos cells) < 5 μM and SI (drug selectivity index) > 10) in the literature during the virtual screening performed with this model. Additionally, 17 drugs have been experimentally tested for *T. cruzi* inhibition (Figure 7). Desloratadine, an anti-histamine drug, and flibanserin (drug originally developed as anti-depressive) showed the highest efficacy against the parasite (A_{max}, maximum activity against parasites = 91 and 92%, respectively), when tested in cells U2OS, infected with Y clone of *T. cruzi*. It is worth mentioning that the intention was to use some compounds in clinical assays and even metabolites of some drugs. However, the little information about these classes of compounds does not allow for solid results.

Nevertheless, in the SBDD approach, a workflow for identifying *T. cruzi* dihydroorotate dehydrogenase (TcDHODH) inhibitors were developed.[78] The best scoring function (ASP) has been used for predicting the binding mode of compounds in the TcDHODH and for eliminating inactive molecules. As this scoring function displayed a low/intermediate capacity to identify compounds with a high affinity at the enzyme, a machine learning model

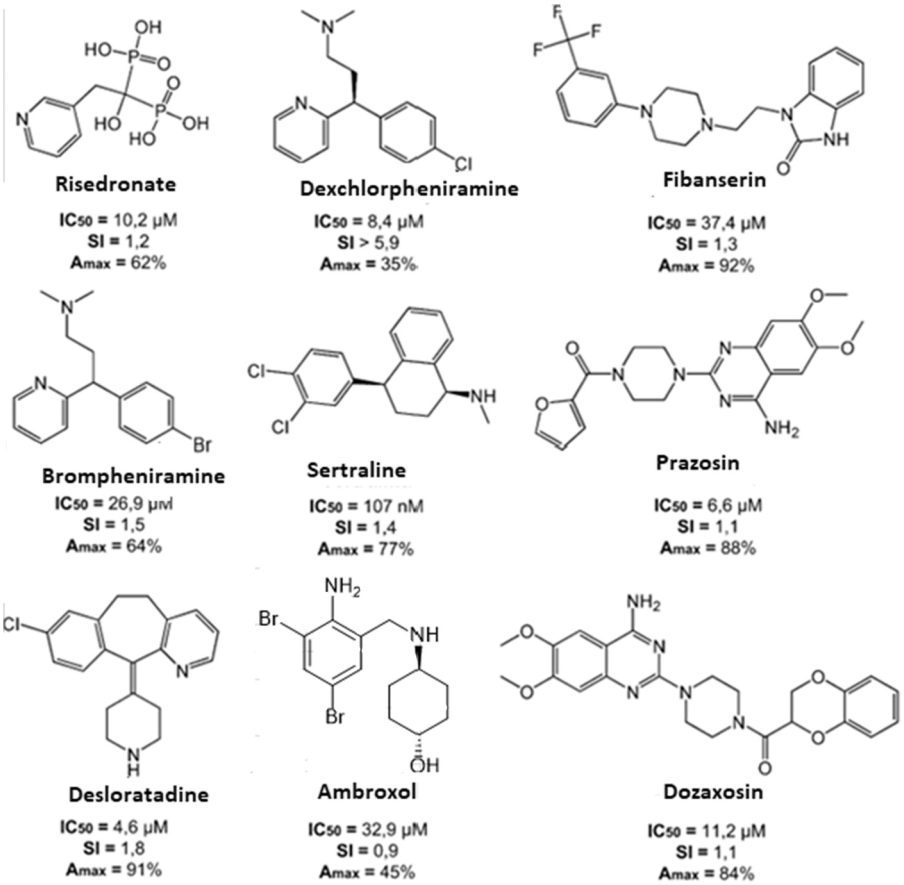
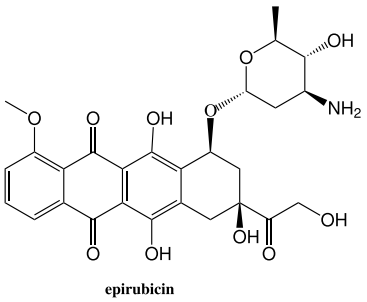


Figure 7. Drugs in trials against Chagas disease.[78] IC_{50} and CC_{50}: drug concentration able to inhibit 50% the proliferation of the parasito or of de host cell, respectively. SI: selectivity index, claculated by the relation between CC_{50} and IC_{50}. Amax: maximum activity determined as the maximum normalized.

was developed, based on orotic acid derivatives tested against the enzyme, to be employed to reclassify the results from an initial virtual screening using the ASP function. The descriptors explored in this work were the ligand-protein interaction enthalpy, the ligand lipophilicity, and the entropic energy loss associated with its complexation with the target protein. The best reclassification model, NB TcDHODH-8, exhibited high performance in the validations used. The SBDD and LBDD models developed in this work can contribute to the development of novel safe and efficacious candidates in the treatment of Chagas disease. Moreover, they can be used as starting points for novel anti-*T. cruzi* drug Research and Development initiatives and the validated models could also be used for novel virtual screening campaigns.

In 2020, Ferreira *et al*.[84] published a study using computer-assisted chemogenomics drug repositioning to identify new hits for malaria treatment. Computer-assisted chemogenomics involves a systematic screening of chemical compounds relative to biological targets in high-throughput screening (HTS) toward finding leads for a determined activity.[84] It is worth mentioning that this tool represents an integration of biological and chemical spaces. From this work, epirubicin (Figure 8), an antineoplastic antibiotic from the anthracycline chemical class, was shown to be active in the sexual and asexual stages of *Plasmodium falciparum* and *P. vivax*. Two molecular targets have been predicted by functional and computational assays as well: *Plasmodium* GyrA and a putative target in *Plasmodium*, a GlcNac-1-P-transferase (GPT) enzyme involved in protein *N*-glycosylation. Both molecular targets are involved in the metabolism of isoprenoids, also found in eukaryotic cells. The results lead to the perspective that through a detailed study of characteristics of epirubicin related to its repositioned activity it might be possible to find leads, and subsequently to optimize them to obtain potential antimalarial agents.



epirubicin

Figure 8. Epirubicin, repositioned for malaria, active in the sexual and asexual phases of *Plasmodium falciparum* and *Plasmodium vivax*.[79]

Leishmaniasis is an intricate disease associated with a complex life cycle that comprises around 20 species of the genus *Leishmania*. It is important to highlight that this complex of protozoa is responsible for health and social aspects in the poor population affected. The chemotherapy for this neglected disease is far from ideal, mainly because of the drug resistance and ineffectiveness of the drugs available. Therefore, it is of utmost importance that new methods could be applied to the search for new and better drugs for this disease. Then, the use of AI methods can be promising to change the paradigm from a classical process of finding drug candidates. With this objective, some works have focused on SBDD and LBDD with the aid of AI, using folate biosynthesis as one of the targets, as well as glycosome, involved in the parasite survival.[85] Works like this have been successful in discovering putative inhibitors. Nonetheless, there were problems, as *in vitro* and *in vivo* validation of targets that could allow minimal side effects. There is a lack of tertiary and quaternary structures of many leishmanial proteins, which has been a problem in applying SBDD in the studies with Leishmania. Notwithstanding, computational studies using AI and machine learning approaches-Alpha Fold2[86] and RoseTTA fold,[87] would allow them to develop studies even though there is no experimental data. Then, the possible characterization of the leishmanial proteins, around eight thousand, can allow the development of anti-leishmanial candidates.[85]

Schistosomiasis is a type of worm infection, which affects more than 200 million neglected people worldwide.[65] There is only one available drug, praziquantel, which has been shown to not be completely effective in this disease. This shows the urgent need to develop new drugs for this worm disease. Many approaches can be used with this aim, including AI methods.[88] The existence of databases, mainly a result of the phenotypic screening and target-based approaches has been decisive for the development of AI works, giving rise to QSAR models, which allow the search for new drugs using Virtual screening, and other methods such as SBDD, LBDD, to discover new bioactive compounds for the treatment of schistosomiasis. To discover drugs that could be active in many stages of *Schistosoma mansoni*, one of the causative agents of schistosomiasis, Andrade *et al.*[88] have implemented a proteome-wide alignment screen of a dataset of 2,114 proteins. They discovered that paroxetine, a selective serotonin reuptake inhibitor (Figure 9), has a potent activity in schistosomula viability, with an EC_{50} = 2.5 μM, after 72 h of exposition, and in male and female worm motility, EC_{50} = 5.1 μM and EC_{50} = 5.1 and 9.9 μM, respectively. In addition, they are selective as schistosomicide. After this experimental observation, they confirmed the molecular basis of this activity, applying molecular modeling studies with an *S. mansoni* serotonin transporter.

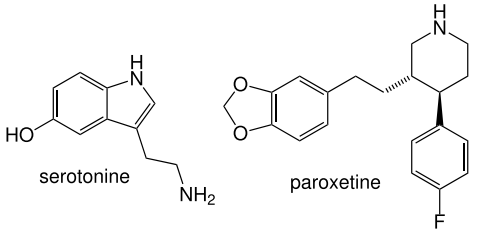


Figure 9. Paroxetine, a selective serotonin reuptake inhibitor like fluoxetine, as anti-*S. mansoni* agent.[81]

Tuberculosis has been responsible for 1,4 million deaths and 10 million cases each year, mainly in neglected poor people.[65] One of the main challenges of tuberculosis is drug resistance, which decreases the effectiveness of medicines used to treat this infectious disease. However, also the differential diagnosis is very relevant, once it can differentiate this infection from other pulmonary diseases.[89] Otherwise, latent tuberculosis infection has been the main cause of active tuberculosis and the use of Machine Learning technology is promising in diagnosis differentiation. Although there are advantages to using this technology, there are some drawbacks that limit this usage. Data is one of the limitations and the employment of different types of data sources can help to have a standard clinical diagnosis, which leads to better treatment of the latent infection. Using this technique in the early stages of the disease would have the advantage of adequately discriminating between latent and active infection, reducing the progression of the former to the latter.[89] AI, with different tools, such as deep learning and radiomics, helps doctors to differentiate drug resistance infections, providing better and specific medicines for treatment.[90] It is relevant to emphasize that *in silico* platforms for drug repositioning,[90] included initiatives towards tuberculosis, in addition to other diseases, such as malaria and dengue. Integration of protein drug targets and approved drugs, which means phenotype and physiology approaches are interrelated in the conception of specific algorithms. The 35% top-ranking predictions identified novel repurposed therapies for many diseases, including for tuberculosis. To summarize, drug repositioning together with AI is also a perspective. It is relevant to solve the gaps still present in the management of TB with the aid of AI, and this includes technical, regulatory, and educational aspects among others.[91] These knowledge gaps, if solved, could provide better ways to manage TB in many aspects.[92]

We presented herein the neglected diseases that have been the highest challenges for poor neglected populations. Nevertheless, with the increase of migration worldwide, countries not involved in epidemic areas have been threatened by those diseases.

6. Concluding Remarks

The evolution of Medicinal Chemistry has been astonishing, especially in the third millennium, with the great development of the areas intimately linked with this fascinating field of science, which searches for the health improvement of the people.

It is worth noting that there is an increase in the availability of databases about the work on drugs for NTDs. This relative change in the scenario related to these diseases allows for the exploration of computational methods for searching for drugs in this area. Hence, the application of AI, especially machine learning, in the modeling and prediction of biological activity toward discovering new drugs for NTD, has been employed much more now, changing the paradigm of drug discovery/design.[66] However, several drawbacks yet deserve to be solved, considering many advanced tools have been introduced over the years.

Different areas of drug design and discovery have benefited from the advanced technology involved. Notwithstanding, there is much to do for orphan/rare diseases and for neglected tropical diseases, which affect more than 1 billion people in the world, causing high morbidity and mortality. It is worth noting that neglected diseases affect primarily poor and neglected people.

Malaria, tuberculosis, Chagas disease, leishmaniasis, and schistosomiasis are the most studied with this objective and for this reason, they have been briefly discussed in terms of works driven with those tools.

If we consider that AI has been used in most areas of science, technology, and medicine, there is space to search for new therapeutic agents for neglected diseases. Despite the interest in developing works toward finding new leads for these diseases, only some of them are awakening for the potentialities of AI together with Drug Computational chemogenomics, which can find promising leads, is still poorly explored in drug repositioning for NTDs.[93]

According to our point of view, AI and drug repositioning, either together or in an isolated form, have been most used in academic environments and in research institutes as well. It is worth mentioning, conversely, that the translation between preclinical to clinical aspects should be mostly managed by the industries. The high costs and the infrastructure necessary for introducing new drugs in the therapeutic normally requires the relevant association with this sector, sometimes intermediated by world organisms of research. This has been performed with a high level of success by DNDi, WHO, just to mention some international organisms. For this very reason, we believe that this integration must be strongly emphasized and probably the

governments display an urgent role towards facing this challenge, probably with public health politics. Maybe the scenario changes in the near and promising future.

Acknowledgments

We want to thank CNPq for supporting the fellowship for E.I.F.

Author Contributions

The contribution has been equitable for all the authors.

Fernanda B. Rodrigues *holds a bachelor's degree in pharmaceutical sciences and biochemistry from the University of São Paulo. She is currently working on Merck S.A. in Brazil as a Junior Sales Assistant in the Lab Water division. Her Final Course Assignment was related to the drug repositioning in the treatment of leishmaniosis through artificial intelligence. This has stimulated her interest in the area of Medicinal Chemistry including application of artificial intelligence in drug repositioning for neglected diseases.*

Marcos Antonio Ferreira Junior *is currently a small molecule drug discovery coordinator at Aché Laboratórios. He holds a master's degree from the School of Pharmaceutical Sciences, University of São Paulo and a bachelor's degree in pharmacy from the State University of Campinas. His research focused on the development of machine learning models to identify drug repurposing opportunities for the treatment of Chagas disease. His interests include drug discovery, medicinal chemistry and in silico modelling.*

Soraya S. Santos *holds a bachelor degree in Pharmacy from Catholic University of Santos, in Santos, Master Degree in Drugs and Medicine Graduate Course, at School of Pharmaceutical Sciences, University of São Paulo, and PhD at the same Course. She has been Assistant Professor at Paulista University, at Santos, since 2013. She was lecturer in Chemical-Pharmaceutical Technology, between 2017 to 2018, and in Medicinal Chemistry, between 2020 and 2021, both at School of Pharmaceutical Sciences University of São Paulo. She had a four-month training at Tulane University, New Orleans, in 2015, developing a work on dendrimers. She is interested in drug design in Medicinal Chemistry and in Pharmacy.*

Jeanine Giarolla *is Assistant Professor at School of Pharmaceutical Sciences, University of São Paulo. She is pharmacist from the Pontifical Catholic University of Campinas, holds her Master's and Doctorate in Sciences, in the Drug and Medicines Program from University of São Paulo, PhD training in University of Central Lancashire and Post-Doctorate in Medicinal Chemistry from the University of São Paulo. Currently, she is visiting Professor at University of Minnesota, Department in Medicinal Chemistry, Institute for Therapeutics Discovery and Development. She is interested in Medicinal Chemistry, especially prodrug design to improve physicochemical properties of drugs/bioactive compounds. She is interested in Medicinal Chemistry, especially prodrug design to improve physicochemical properties of drugs/bioactive compounds.*

Elizabeth Igne Ferreira *is pharmacist from the School of Pharmaceutical Sciences, University of Sao Paulo. She got the PhD in Science, area of Organic Chemistry, at Institute of Chemistry, University of São Paulo. She made a short pos-doctoral training at University of Illinois at Chicago, in Molecular Modeling, supervised by Prof A. J. Hopfinger. Since 2001, she has been Full Professor, now retired after around 49 years of teaching at School of Pharmaceutical Sciences, University of São Paulo, where she developed works on Medicinal Chemistry, area of Drugs for Neglected Diseases. She had been Editor-in-Chief of Brazilian Journal of Pharmaceutical Sciences for 25 years. Since 2014, she has been member of INCt-INOFAR, under the coordination of Prof Eliezer J. Barreiro.*

References

1. Wermuth, C. G.; Ganellin, C. R.; Lindberg, P.; Mitscher, L. A.; *Pure Appl. Chem.* **1998**, *70*, 1129. [Crossref]
2. Barreiro, E. J.; *Princípios de Química Medicinal, A Evolução Cronológica da Química Medicinal,* http://www.lassbio.icb.ufrj.br/download/cursos_semin/SEFAR_2010_PARTE1.pdf, accessed in July 2024.
3. Schneider, P.; Walters, W. P.; Plowright, A. T.; Sieroka, N.; Listgarten, J.; Goodnow, R. A.; Fisher, J.; Jansen, J. M.; Duca, J. S.; Rush, T. S.; Zentgraf, M.; Hill, J. E.; Krutoholow, E.;

Kohler, M.; Blaney, J.; Funatsu, K.; Luebkemann, C.; Schneider, G.; *Nat. Rev. Drug Discovery* **2020**, *19*, 353. [Crossref]
4. Lombardino, J.G. Lowe III, *Nat. Rev./Drug Disc.* **2004**, *3*, 853. [Crossref]
5. Neumeyer, J. L. In *Foye´s Principles of Medicinal Chemistry*, 7th ed.; Lemke, T. L.; Williams, D. A.; Roche, V. F.; Zito, S. W., eds.; Wolters Kluvers Walters Lippincott, Williams & Wilkins: Baltimore, USA, 2013.
6. Padhy, B. M.; Gupta, Y. K.; *J. Postgrad. Med.* **2011**, *57*, 153. [Crossref]
7. Schuler, J.; Falls, Z.; Mangione, W.; Hudson, M. L.; Bruggemann, L.; Samudrala, R.; *Drug Discovery Today* **2022**, *27*, 49. [Crossref]
8. Laboratório de Avaliação e Síntese de Substâncias Bioativas (LASSBio), http://www.lassbio.icb.ufrj.br/, accessed in June 2024.
9. Instituto Nacional de Ciência e Tecnologia em Fármacos e Medicamentos (INCT-INOFAR) CNPq, http://www.inct-inofar.ccs.ufrj.br/index.php/en/, accessed in June 2024.
10. LASSBio, Quimioteca do LASSBio, http://www.lassbio.icb.ufrj.br/quimioteca.html, accessed in June 2024.
11. Pedreira, G. B. J.; Franco, L. S.; Barreiro, E. J.; *Curr. Top. Med. Chem.* **2019**, *19*, 1679. [Crossref]
12. Menegatti, R.; Cunha, A. C.; Ferreira, V. F.; Perreira, E. F. R.; ElNabawi, A.; Eldefrawi, A. T.; Albuquerque, E. X.; Neves, G.; Rates, S. M. K.; Fraga, C. A. M.; Barreiro, E. J.; *Bioorg. Med. Chem*. **2003**, *11*, 4807. [Crossref]
13. Ashburn, T. T.; Thor, K. B.; *Nat. Rev. Drug Discovery* **2004**, *3*, 673. [Crossref]
14. Pushpakom, S.; Iorio, F.; Eyers, P. A.; Escott, K. J.; Hopper, S.; Wells, A.; Doig, A.; Guilliams, T.; Latimer, J.; McNamee, C.; Norris, A.; Sanseau, P.; Cavalla, D.; Pirmohamed, M.; *Nat. Rev. Drug Discovery* **2019**, *18*, 41. [Crossref]
15. Jourdan, J. P.; Bureau, R.; Rochais, C.; Dallemagne, P.; *J. Pharm. Pharmacol*. **2020**, *72*, 1145. [Crossref]
16. Li, Y. Y.; Jones, S. J.; *Genome Med*. **2012**, *4*, 27. [Crossref]
17. Low, Z. Y.; Farouk, I. A.; Lal, S. K.; *Viruses* **2020**, *12*, 1058. [Crossref]
18. Hua, Y.; Dai, X.; Xu, Y.; Xing, G.; Liu, H.; Lu, T.; Chen, Y.; Zhang, Y.; *Eur. J. Med. Chem*. **2022**, *234*, 14239. [Crossref]
19. Ferreira, E. I.; Barreiro, E. J. In F*undamentals of Pharmaceutical Medicinal Chemistry*; Ferreira, E. I.; Giarolla, J.; Filho, R. P.; Barreiro, E. J., eds.; Editora Manole: Santana de Parnaíba, Brazil, 2022.
20. Parvathaneni, V.; Kulkarni, N. S.; Muth, A.; Gupta, V.; *Drug Discovery Today* **2019**, *24*, 2076. [Crossref]
21. Wouters, O. J.; McKee, M.; Luyten, J.; *JAMA* **2020**, *323*, 844. [Crossref]
22. Pillayar, T.; Meenakshisundaram, S.; Manickam, M.; Sankaranarayanan, M.; *Eur. J. Med. Chem*. **2020**, *195*, 112275. [Crossref]
23. Begley, C. G.; Ashton, M.; Baell, J.; Bettess, M.; Brown, M. P.; Carter, B.; Charman, W. N.; Davis, C.; Fisher, S.; Frazer, I.; Gautam, A.; Jennings, M. P.; Kearney, P.; Keeffe, E.; Kelly, D.; Lopez, A. F.; McGuckin, M.; Parker, M. W.; Rayner, C.; Roberts, B.; Rush, J. S.; Sullivan, M.; *Sci. Transl. Med.* **2021**, *13*, eabd5524. [Crossref]
24. Oprea, T. I.; Bauman, J. E.; Bologa, C. G.; Buranda, T.; Chigaev, A.; Edwards, B. S.; Jarvik, J. W.; Gresham, H. D.; Haynes, M. K.; Hjelle, B.; Hromas, R.; Hudson, L.; Mackenzie, D. A.; Muller, C. Y.; Reed, J. C.; Simons, P. C.; Smagley, Y.; Strouse, J.; Surviladze, Z.; Thompson, T.; Ursu, O.; Waller, A.; Wandinger-Ness, A.; Winter, S. S.; Wu, Y.; Young, S. M.; Larson, R. S.; Willman, C.; Sklar, L. A.; *Drug Discovery Today: Ther. Strategies* **2011**, *8*, 61. [Crossref]
25. Bayraktar, A.; Li, X.; Kim, W.; Zhang, C.; Turkez, H.; Shoaie, S.; Mardinoglu, A.; *J. Transl. Med*. **2023**, *21*, 332. [Crossref]
26. Daina, A.; Michielin, O.; Zoete, V.; *Swiss ADME*; SiB Swiss, Inst. Bioinformatics, Molecular Modeling Group, Lausanne, Swiss, 2017.
27. Gan, Y.; Liu, T.; Feng, W.; Wang, L.; Li, L. I.; Ning, Y.; *Oncol. Res*. **2023**, *31*, 333. [Crossref]
28. Elfaky, M. A.; Elbaramawi, S. S.; Eissa, A. G.; Ibrahim, T. S.; Khafagy, E. S.; Ali, M. A. M.; Hegazy, W. A. H.; *Appl. Microbiol. Biotechnol*. **2023**, *107*, 3763. [Crossref]
29. da Rosa, T. F.; Serafin, M. B.; Foletto, V. S.; Franco, L. N.; de Paula, B. R.; Fuchs, L. B.; Calegari, L.; Hörner, R.; *Curr. Microbiol*. **2023**, *80*, 160. [Crossref]
30. Turing, A. M.; *Mind* **1950**, *59*, 433.
31. Kaul, V.; Enslin, S.; Gross, S. A.; *Gastrointest. Endosc*. **2020**, *92*, 807. [Crossref]
32. Boniface, P. K.; Sano, C. M.; Ferreira, E. I.; *Curr. Drug Targets* **2020**, *21*, 681. [Crossref]
33. Struble, T. J.; Alvarez, J. C.; Brown, S. P.; Chytil, M.; Cisar, J.; DesJarlais, R. L.; Engkvist, O.; Frank, S. A.; Greve, D. R.; Griffin, D. J.; Hou, X.; Johannes, J. W.; Kreatsoulas, C.; Lahue, B.; Mathea, M.; Mogk, G.; Nicolaou, C. A.; Palmer, A. D.; Price, D. J.; Robinson, R. I.; Salentin, S.; Xing, L.; Jaakkola, T.; Green, W. H.; Barzilay, R.; Coley, C. W.; Jensen, K. F.; *J. Med. Chem.* **2020**, *63*, 8667. [Crossref]
34. Chassagnon, G.; Vakalopolou, M.; Paragios, N.; Revel, M. P.; *Eur. Radiol.* **2020**, *30*, 2021. [Crossref]
35. Vatansever, S.; Schlessinger, A.; Wacker, D.; Kaniskan, H. U.; Jin, J.; Zhou, M. M.; Zhang, B.; *Med. Res. Rev*. **2021**, *41*, 1427. [Crossref]
36. Aggrawal, S.; Karmakar, A.; Krishnakumar, S.; Paul, S.; Singh, A.; Banerjee, N.; Laha, N.; Ball, G. B.; Srivastava, S.; *Curr. Top. Med. Chem.* **2023**, *23*, 551. [Crossref]
37. Lei, S.; Lei, X.; Liu, L.; *Front. Pharmacol.* **2022**, *21*, 13. [Crossref]
38. Gupta, R.; Srivastava, D.; Sahu, M.; Tiwari, S.; Ambasta, R. K.; Kumar, P.; *Mol. Diversity* **2021**, *25*, 1315. [Crossref]
39. Carracedo-Reboredo P.; Liñares-Blanco, J.;

Rodríguez-Fernández, N.; Cedrón, F.; Novoa, F. J.; Carballal, A.; Maojo, V.; Pazos, A.; Fernandez-Lozano, C.; *Comput. Struct. Biotechnol. J.* **2021**, *19*, 4538. [Crossref]
40. Luo, H.; Li, M.; Yang, M.; Wu, F. X.; Li, Y.; Wang, J.; *Briefings Bioinf.* **2021**, *22*, 1604. [Crossref]
41. Mottini, C.; Napolitano, F.; Li, Z.; Gao, X.; Cardone, L.; *Semin. Cancer Biol.* **2021**, *68*, 59. [Crossref]
42. Vanhaelen, Q.; Mamoshina, P.; Aliper, A. M.; Artemov, A.; Lezhnina, K.; Ozerov, I.; Labat, I.; Zhavoronkov, A.; *Drug Discovery Today* **2017**, *22*, 210. [Crossref]
43. Jarada, T. N.; Rokne, J. G.; Alhajj, R.; *J. Cheminform.* **2020**, *12*, 46. [Crossref]
44. Wang, Y.; Yang, Y.; Chen, S.; Wang, J.; *Briefings Bioinf.* **2021**, *22*, bbab048. [Crossref]
45. Liu, R.; Wei, L.; Zhang, P.; *Nat. Mach. Intell.* **2021**, *3*, 68. [Crossref]
46. March-Vila, E.; Pinzi, L.; Sturm, N.; Tinivella, A.; Engkvist, O.; Chen, H.; Rastelli, G.; *Front. Pharmacol.* **2017**, *8*, 298. [Crossref]
47. Amiri, R.; Razmara, J.; Parvizpour, S.; Izadkhah, H.; *BMC Bioinf.* **2023**, *24*, 442. [Crossref]
48. Zeng, X.; Zhu, S.; Lu, W.; Liu, Z.; Huang, J.; Zhou, Y.; Fang, J.; Huang, Y.; Guo, H.; Li, L.; Trapp, B. D.; Nussinov, R.; Eng, C.; Loscalzo, J.; Cheng, F.; *Chem. Sci.* **2020**, *11*, 1775. [Crossref]
49. Rao, M.; McDuffie, E.; Sachs, C.; *Toxics* **2023**, *11*, 875. [Crossref]
50. Eisenstein, M.; *Nature* **2015**, *527*, S2. [Link] accessed in June 2024
51. Gligorijević, V.; Malod-Dognin, N.; Pržulj, N.; *Proteomics* **2016**, *16*, 741. [Crossref]
52. Chen, Y.; Elenee Argentinis, J. D.; Weber, G.; *Clin. Ther*. **2016**, *38*, 688. [Crossref]
53. Luo, Y.; Zhao, J.; Yang, J.; Zhang, Y.; Kuang, W.; Peng, J.; Chen, L.; Zeng, J.; *Nat. Commun.* **2017**, *8*, 573 [Crossref]
54. Sadeghi, S. S.; Keyvanpour, M. R.; *Curr. Comput.-Aided Drug Des.* **2020**, *16*, 354. [Crossref]
55. Ko, Y.; *Appl. Sci.* **2020**, *10*, 5076. [Crossref]
56. Wilkinson, M. D.; Dumontier, M.; Aalbersberg, I. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J. W.; da Silva Santos, L. B.; Bourne, P. E.; Bouwman, J.; Brookes, A. J.; Clark, T.; Crosas, M.; Dillo, I.; Dumon, O.; Edmunds, S.; Evelo, C. T.; Finkers, R.; Gonzalez-Beltran, A.; Mons, B.; *Sci. Data* **2016**, *3*, 160018. [Crossref]
57. Issa, N. T.; Stathias, V.; Schürer, S.; Dakshanamurthy, S.; *Semin. Cancer Biol.* **2021**, *68*, 132. [Crossref]
58. Tanoli, Z.; Vähä-Koskela, M.; Aittokallio, T.; *Expert. Opin. Drug Discovery* **2021**, *16*, 977. [Crossref]
59. Zhou, Y.; Wang, F.; Tang, J.; Nussinov, R.; Cheng, F.; *Lancet Digital Health* **2020**, *2*, e667. [Crossref]
60. Bender, A.; Cortes-Ciriano, I.; *Drug Discovery Today* **2021**, *26*, 1040. [Crossref]
61. Angermueller, C.; Pärnamaa, T.; Parts L.; Stegle, O.; *Mol. Syst. Biol.* **2016**, *12*, 878. [Crossref]
62. Yu, J.-L.; Dai, Q.-Q.; Li, G.-B.; *Drug Discovery Today* **2022**, *27*, 1796. [Crossref]
63. Bender, A.; Cortés-Ciriano, I.; *Drug Discovery Today* **2021**, *26*, 511. [Crossref]
64. Aittokallio, T.; *Expert Opin. Drug Discovery* **2022**, *17*, 423. [Crossref]
65. World Health Organization (WHO); *Global Observatory on Health R&D*; WHO: Geneva, Switzerland, 2022. [Link] accessed in June 2024
66. Winkler, D. A.; *Front Chem*. **2021**, *9*, 614073. [Crossref]
67. Melo, G. B. T.; Ângulo-Tuesta, A.; Silva, R. E. N.; Santos, T. S.; Uchimura, L. Y. T.; Obara, M. T.; *PLoS Neglected Trop. Dis.* **2023**, *1717*, e0011134. [Crossref]
68. Klug, D. M.; Gelb, M.H.; Pollastri, M. P.; *Bioorg. Med. Chem. Lett*. **2016**, *26*, 2569. [Crossref]
69. Yin, Z.; Wong, S. T. C.; *Emerging Top Life Sci.* **2021**, *5*, 803. [Crossref]
70. Hernandez, H. W.; Soeung, M.; Zorn, K. M.; Ashoura, N.; Mottin, M; Andrade, C. H.; Caffrey, C. R.; Siqueira-Neto, J. L.; Ekins, S.; *Pharm. Res.* **2019**, *36*, 27. [Crossref]
71. Wiwanitkit, S.; Wiwanitkit, V.; *Asian Pac. J. Trop. Biomed.* **2016**, *6*, 350. [Crossref]
72. Leonardi, D.; Salomóna, C. J.; Lamas, M. C.; Olivieri, A. C.; *Int. J. Pharm*. **2009**, *367*, 140. [Crossref]
73. Guerra, A.; Gonzalez-Naranjo, P.; Campillo, N. E.; Cerecetto, H.; *Curr. Comput-Aided Drug Des.* **2013**, *9*, 130. [Crossref]
74. Cerecetto, H.; González, M.; *Mini-Rev. Med. Chem.* **2008**, *8*, 1355. [Crossref]
75. Stud, M.; *CODES*, v1.0 (revision 3) ed.; Manfred Stud in Instituto de Química Médica, CSIC, Madrid, Spain.
76. Planer, J. D.; Hulverson, M. A.; Arif, J. A.; Ranade, R. M.; Don, R.; Buckner, F. S.; *PLoS Neglected Trop. Dis.* **2014**, *8*, e2977. [Crossref]
77. Aguilera, R.; Alvarez, G.; Cerecetto, H.; González, M. *Curr. Med. Chem*. **2019**, *26*, 4476. [Crossref]
78. Ferreira Junior, M. A.: *Reposicionamento de Fármacos para o Tratamento de Doença de Chagas: Desenvolvimento de Modelos Computacionais Baseados em Ligantes Ativos Fenotipicamente Contra o Trypanosoma Cruzi e na Inibição da Diidroorotato Desidrogenase*; MSc Dissertation, Universidade de São Paulo, São Paulo, Brazil, 2021. [Link] accessed in June 2024
79. Berthold, M.; Erickson, T.; Kaloustian, M.; Staccia, B.; Ohl, P.; *KNIME* v3.6.0; University of Konstanz, Konstanz, Switzerland. [Link] accessed in June 2024
80. *Python*, v3.8.2; Python Software Foundation, United States, 2020. [Link] accessed in June 2024
81. CDK: The Chemistry Development Kit, https://cdk.github.io/, accessed in June 2024.

82. RDKit: Open-source cheminformatics, 2018, https://www.rdkit.org/ accessed in June 2024.
83. Novamechanics Enalos, https://www.enaloscloud.novamechanics.com/ accessed in June 2024.
84. Ferreira, L. T.; Rodrigues, J.; Cassiano, G. C.; Tavella, T. A.; Tomaz, K. C. P.; Baia-da-Silva, D. C.; Souza, M. F.; Lima, M. N. N.; Mottin, M.; Almeida, L. D.; Calit, J.; Puça, M. C. S. B.; Melo, G. C.; Bargieri, D. Y.; Lopes, S. C. P.; Lacerda, M. V. G.; Bilsland, E.; Sunnerhagen, P.; Neves, B. J.; Andrade, C. H.; Cravo, P. V. L.; Costa, F. T. M.; *Antimicrob. Agents Chemother.* **2020**, *64*, e02041-19. [Crossref]
85. Soni, M.; Pratap, J. V.; *Pathogens* **2022**, *11*, 950. [Crossref]
86. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Michalina Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Nature 2021, 596, 583. [Crossref] https://alphafold.ebi.ac.uk
87. Rosetta, Google Deep Mind, EMBL-EBI, University of Washington, Washington, USA.
88. Andrade, C. H.; Neves, B. J.; Melo-Filho, C. C.; Rodrigues, J.; Silva, D. C.; Braga, R. C.; Cravo, P. V. L.; *Curr. Med. Chem.* **2019**, *26*, 4355. [Crossref]
89. Liang, S.; Ma, J.; Wang, G.; Shao, J.; Li, J.; Deng, H.; Wang, C.; Li, W.; *Front. Med.* **2022**, *9*, 935080. [Crossref]
90. Li, L.-S.; Yang, L.; Zhuang, L.; Ye, Z.-Y.; Zhao, W.-G.; Gong, W.-P.; *Mil. Med. Res.* **2023**, *28*, 58. [Crossref]
91. Mullins, J. G. L.; *Biochem. Soc. Trans.* **2022**, *50*, 747. [Crossref]
92. Ferreira, L. G.; Andricopulo, A. D.; *Drug Discovery Today* **2016**, *21*, 1699. [Crossref]
93. Dohál, M.; Porvazník, I.; Solovič, I.; Mokrý, J.; *Front. Microbiol.* **2023**, *14*, 1225438. [Crossref]

Submitted: January 20, 2024
Published online: August 12, 2024