Multi-Target Neuroprotection Lead Identification Through Deep Learning: AI-Driven Platforms for Enhanced Drug Discovery in Neurodegenerative Disease

Dr. Eric Verschueren, Professor of Electrical Engineering, Ghent University, Belgium

1. Introduction

exhaustive knowledge of disease at the biochemical, genetic, and physiological levels, combined with clinical trials directly on patients, is not only important but absolutely essential for achieving efficacious therapies that truly improve patient outcomes. To date, the need for this comprehensive understanding cannot be overstated as it serves as the foundation for innovation in medical treatment.

1.1. Overview of Neurodegenerative Diseases

A broad array of progressive disorders that cause the degeneration of the structure and function of the brain and spinal cord fall under the umbrella of neurodegenerative diseases. The primary risk factor for these diseases is advancing age. Neurodegenerative diseases are mainly classified into three types based on their clinical severity and anatomical structures of the nervous system: dementia, movement disorders, and motor neuron diseases. Neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, frontotemporal dementia, and multiple sclerosis are slowly emerging as a major health and socioeconomic burden around the globe. These disorders display regional pathophysiology; Alzheimer's disease primarily causes dementia and memory loss, while Parkinson's disease causes movement disorders, amyotrophic lateral sclerosis results in loss of muscle use, and Huntington's disease is notorious for the selective and progressive demise of subcortical structures. The etiology underlying these diseases is multifactorial, with genetic as well as environmental factors, which are either lifestyle or metabolic cues.

Normal aging is associated with functional degeneration in the nervous system; however, these disorders exhibit a distinct pathophysiology in which the damaged cells

are not replaced or repaired, leading to progressive and irreversible damage. Owing to the complexity of the pathogenesis, the diagnosis of these diseases is frequently delayed due to the association of varied and heterogeneous phenotypes, associated comorbid disorders, and late-stage clinical presentation. In terms of treatment modalities, research has renounced any development of a drug that can cure these diseases up until now. Recently, molecular-level therapeutic strategies are being designed that can modify the underlying pathogenic cascades progressively and hold the ability to prevent the slow progression of the diseases. Development and adoption of such a strategy require effective patient stratification and identification of meaningful responders to such a slow-acting therapy. In simpler terms, a 'singular drug fits all' approach cannot be adopted, and a more personalized medicine approach is developing to treat patients suffering from these diseases.

1.2. Challenges in Drug Discovery for Neurodegenerative Diseases

The development of therapeutics for neurodegenerative disorders poses multiple difficulties in the field of drug discovery. For example, 90% of new candidate drugs fail to reach the market, and less than 12% of drugs in clinical trials reach the market. One of the key reasons for unsuccessful drug development is the inability to translate preclinical findings to relevant clinical outcomes such as improved cognition, survival, and quality of life. In part, this is due to the current preclinical models, which do not recapitulate complex disease states. For example, neurodegenerative diseases reflect the complexity of aging; the aggregation of pathologies and comorbidities with age results in a wide range of clinical phenotypes. Moreover, common in vivo models have reduced predictive validity for drug repositioning. This leads to lengthy and resource-intensive optimization steps during drug discovery campaigns. To mitigate this, preclinical data is incorporated into target validation pipelines. However, the relevance of assayed endpoints is a significant drawback, as well as the often reductionist view towards the underlying disease mechanisms, resistant to single-target approaches.

In the clinic, there are further inherent challenges relevant to neurodegenerative diseases. The aging population will lead to an increase in individuals with age-related diseases and, with it, progression models that predict elderly patients' outcomes in the near future. Given the time, cost, and attrition rates of clinical drug trials, improved methodologies that predict and prognosticate are desperately needed. The relevance of

preclinical biomarkers can also be an issue. There are no peripheral biomarkers linked to disease mechanisms that correlate with clinical measures in humans. As a result, lumbar puncture and imaging techniques are often employed due to increased relevance to the relevant disease pathways. They have some significant limitations, however, including the cost, difficulty in obtaining longitudinal data, and confounders that affect sample quality. Regulatory and market pressures also weigh heavily on the decision-making process. Finally, the financial consequences of drug failures mean many pharmaceutical companies are reticent to engage in the development of novel therapeutics for neurological disorders. While this could expand market share for successful drugs that are brought to the market, it raises other relevant questions. If pharmaceutical companies are only engaging in drug discovery fields with existing success, will we continue to get novel therapeutics from them in this field?

1.3. Role of AI and Machine Learning in Drug Discovery

The use of artificial intelligence (AI) and machine learning in drug discovery is of significant interest for addressing the complex, time- and cost-intensive issues associated with developing disease-modifying treatments for neurodegenerative diseases. They are attractive platforms for the drug discovery pipeline for several reasons. The first is their ability to acquire, process, and analyze complex biomedical data, which exceeds that routinely handled by researchers or modern databases. The second is their capability to spot patterns and create hypotheses, both being necessary and challenging bottlenecks in target and drug candidate discovery, as well as in predicting in vitro, in vivo, or clinical trial results, thus streamlining the riskiest and costliest steps of the development process. Machine learning, in particular, has shown recent substantial success for target discovery—arguably the hardest part of drug discovery. In conclusion, AI represents an innovative way of thinking, connecting several dots and filling several scientific gaps that traditional drug discovery methods have not or cannot achieve.

AI technologies possess the potential to revolutionize in silico modeling, segmenting, and disease stratification, and a few in silico projects are now being conducted with proven expertise. A number of large companies are also starting or have conducted large projects for drug discovery and development in neurodegenerative diseases that combine real-world data with AI techniques and models, showing how the AI revolution is believed to bridge current knowledge and system gaps in these diseases. It

is not the objective of this review to frame the technical aspects of the AI technologies described above in neurodegenerative diseases within the context of the so-called 'AI taxonomy.' Recognition, however, must be given to the novel and critical role machine learning has in drug discovery. It is a well-known fact that machine learning algorithms are designed to facilitate the learning of system patterns and relationships directly from data, automating the processing of vast amounts of data rapidly. In combination with robotics, which are already capable of routine high-throughput analysis of screening, toxicology, pharmacokinetics, and the like, and with adoption in drug repurposing and prototype design, this powerful technology can fundamentally reduce drug discovery costs and timings.

2. Machine Learning Applications in Identifying Novel Therapeutic Targets

The successful discovery of novel and effective drug targets for neurodegenerative diseases could revolutionize the treatment of neurological disorders. The identification of relevant and reliable potential targets is crucial in improving the likelihood of efficacy for a drug candidate. If a selected target is not involved in the disease or is not responsible for the mechanism addressed by the drug, the whole treatment could fail, producing large waste sums of money and time. Many research efforts focus on improving the drug target identification process. Many sources of omic data can be used to generate new insights into neurodegenerative diseases and provide new therapeutic nominations. Genomic studies using simple statistical linear models, such as differential expression analysis or differential gene expression analysis, amass thousands of potential new experimental targets without addressing their reliability. Researchers use gene expression data, protein expression data, or the literature collection to identify potential suitable targets.

Feature selection and feature engineering, i.e., the combination of noisy data to build novel, more powerful predictive features, are equally important steps. Several studies demonstrate the importance of accurate target identification, affirming the ability of notable performance improvements following progressive feature selection, feature engineering, and further algorithm-driven model optimization. As a widely versatile and class nondiscriminative algorithm, unsupervised learning approaches are capable of revealing insights in exploratory or discovery-driven studies through dimensionality reduction and data structure detection, successfully identifying appropriate feature

subsets or molecular correlations that can ultimately inform further, more specific target identification. In clinical reasoning, traditional supervised learning analysis is often used for categorization or regression-based prediction of noteworthy patient outcomes, such as survival, progression, or treatment response. High-dimensional, hypothesis-free algorithms are powerful tools in the identification of spared, rare, or unexpected underlying categories in complex omic systems, including the identification of novel actionable targets that may be overlooked in traditional analysis. Several recent case studies highlighted the potential in harnessing machine learning approaches for the identification of disguised or withheld therapeutic targets.

2.1. Data Sources and Types in Drug Discovery

Despite the significant advances in drug discovery for many diseases, there are limited therapeutic options for most central nervous system (CNS) disorders. This is, in part, due to the lack of knowledge regarding disease pathogenesis, heterogeneity, and complex symptoms. Therefore, it is essential to utilize a broad range of data, including clinical and biological data from tissues and patient samples, to support efforts in identifying innovative therapies for neurodegenerative diseases (NDDs) that are centered on disease modification. Moreover, this big data revolution can improve the successful translation of therapies. In drug discovery, different types of data are used to describe various biological levels, from molecular mechanisms to phenotypes. The integration of such diverse data is pivotal to our understanding of the mechanisms underlying disease and the identification of novel drug targets. Genetic data can demonstrate how variations in genes can lead to changes at the protein level, consequently causing variations at the level of cells, tissues, and ultimately, clinical symptoms. High-throughput technologies, such as genomics, proteomics, and metabolomics, can provide data on changes in the amounts and modifications of biomolecules in biological systems, as well as changes in activities and relationships. Furthermore, cellular phenotypes can also be measured, either using in vitro or in vivo not necessarily from the patients. These assays can help researchers understand the effect of genetics and proteins on cellular function and can be used for drug screening purposes. Ultimately, clinical data are used for understanding the disease patterns and treatment response. These multiple data types can be used to paint a picture of not only the disease but also of the underlying processes.

2.2. Feature Selection and Engineering

Feature selection and feature engineering as part of the wider data preprocessing and machine learning lifecycle are central to the development of models for drug discovery tasks. While machine learning models are data-hungry, having a model with unnecessary features slows down training, impacts interpretability, and retains noise or confounding effects from irrelevant features. A deep understanding of the data and its biology, as well as previous literature and experimental findings, are required determinants of what to include in machine learning approaches. An ability to handle big data is sometimes a requirement too, as datasets in drug discovery tasks can often be heterogeneous and large.

Feature selection can occur using a host of different methods including statistical tests, algorithms, or by manual input in the form of valuable domain expertise. Once relevant features are selected, generating additional variables of topological, physicochemical, or biological interest that further influence the modeling of inputs and targets — with strong underlying biological arguments if used for between-target comparisons — is of value in preprocessing during feature engineering. For examples of feature engineering in between-target comparisons: primary response features; creating protein-protein interactions for determining regulatory or synonymous SNPs; and primary drug-tissue exposure as new features. Thus, in drug discovery, feature selection and engineering are not only important for gaining predictive accuracies, but because of the direct interpretation of manually selected and engineered features, the biological significance in both biomedical and biotechnological applications needs to be considered.

2.3. Machine Learning Algorithms for Target Identification

Machine learning algorithms are widely employed to support sponsors of drug development in the early stages of target identification. They can be classified into two main categories: supervised learning and unsupervised learning, both of which feature their pros and cons for different tasks and various types of input data, such as gene expression data, proteomics, interaction data, or phenotypic data. Although pre-selection of targets can speed up the computational approach, the known complexity of various interactions between a plethora of molecules and pathways sometimes necessitates an unbiased data-driven approach, searching for all possible connections and targets. The most frequently used supervised learning algorithms are decision trees,

support vector machines, and neural networks, whereas the unsupervised learning algorithm of choice is clustering. One of the challenges is to avoid overfitting the model and ensuring that the generated hypotheses are suitable for other clinical tests, let alone commercial or regulatory approval. Finally, the interpretability and feedback possibilities of the model are critical. It is currently possible to predict drug targets using biological and genetic methodologies that have led to the identification of novel targets successfully undergoing clinical trials.

Several reviews have been published listing the algorithms developed for target identification in drug discovery or systems medicine. The majority of promising target identification models are based on the use of machine learning or data mining algorithms and differ in the kind of input data used to perform the target predictions. The use of clinical data may also serve as a powerful data source. One of the latest reviews has compiled cutting-edge deep learning techniques for the identification of potential drug targets in oncology. Finally, introducing molecular dynamics as filters for drug development to select potential targets is also considered state of the art.

3. Predicting Disease Modifiers using AI

Relating to the prediction of so-called "disease modifiers" for neurodegenerative diseases, it can be difficult to track how the disease is progressing in order to return information about which drugs may be more effective for a given patient. As a result, the field of precision medicine in neurodegenerative diseases has so far proven difficult. However, uncovering more of these disease modifiers can give a more comprehensive representation of genetic interactions that impact the disease. By using data that tracks a patient's disease progression and their treatment response to train machine learning models, we are now in a position to predict possible disease modifiers far more effectively. One of the challenges of training models based on machine learning is that an increased amount of information typically leads to improved model performance. Due to the fact that this training process requires the integration of data from multiple sources, the use of multi-omics approaches has seen a surge in popularity. If the predictive power of disease progression or any of its individual steps can be either classified or regressed accurately, these predicted changes are likely to be related to the disease in some way.

Our predictions are designed to be as general as possible, meaning that they are able to include the prediction of complex molecular characteristics, like those reached by drugs and genes in order to make either gene expression or omics predictions. The disease modifiers identified for prediction using genetics can, to a large extent, be used to predict other forms of omics, such as metabolomics. This is because these forms of omics are ultimately driven by the biology of an individual's genomic genetic makeup. By performing an integrative deep learning approach involving multiple omics, it is now possible to predict changes in gene expression. In particular, integrating RNA-seq or other expression-based data of varying depth and breadth with other multifactorial data sources such as proteomics, lipidomics, metabolomics, and/or imaging, with the aim to discover predictive differences of one or more of the omics that might reveal more about a patient's condition, represents a particularly valuable future direction for this work. Deep learning time series models in particular are increasingly used in medical image analysis and for hybrid transcriptomics lately. These disease progression investigation efforts are the crème de la crème of bioinformatic differential expression analysis, providing training data that could boost and optimize an integrative predictive model. It is also worth noting that highly confounding clinical variables such as age and gender can now be incorporated into the model as a feature to be of improved use. As long as the potential outcomes of particular genes or disease progression transcriptomics can be gathered, it will be a future goal to develop an integrative prognostic or diagnostic disease model that can also inform about the effect size if the highly confounding demographic-related variables can be included.

3.1. Understanding Disease Modifiers in Neurodegenerative Diseases

One fascinating aspect of neurodegenerative diseases resides in the fact that multiple factors can influence disease onset and severity. Sometimes referred to as "disease modifiers," these are factors external to the pathogenic mutation itself but could nonetheless play a crucial role in the course and progression of the disease. Drug discovery for neurodegenerative diseases should therefore take into account these disease modifiers, as we now understand them to have a great impact on patient prognosis. These factors can take the form of biological, genetic, or environmental factors. A few prominent examples of biological modifiers include systemic inflammation, metabolism, vascular factors, or multiple system atrophy pathology. Genetic modifiers give rise to what is referred to as "genetic penetrance" or "genetic

risk," that is, a set of genes that modify the effect of a primary, causative pathogenic mutation. The concept of "environmental modifiers" is broad and includes lifestyle, occupation, and also, presumably, the composition of the human gut microbiome.

A set of general therapeutic strategies can be coined to target these disease modifiers. We could aim to maintain the action of "protective modifiers" or target "faulty protective mechanisms" upregulating the potential disease modifiers' effect on slowing disease progression. Intervention could also be aimed at the level of "inhibiting factors accelerating disease progression," halting the pathological effect of those potential disease modifiers. Finally, we could also focus on inhibiting "complex disharmony" factors acting at the clinical level to slow disease progression. Overall, it is just starting to become clear to clinicians that much of the bewildering variability in natural history experienced by patients with neurodegenerative diseases could reside in those same biological, environmental, and genetic disease modifiers that are now being identified. As previously mentioned, drug development at the clinical trial level is also greatly suffering from the variability in natural history, leading to the need for ever-larger patient cohorts along with even longer follow-up for appropriate intervention demonstration. Unless this issue is addressed to be preempted by using precise genetic and biomarker stratification, it is likely that to address such a great degree of variability efficiently, it will be necessary to develop ever more heterogeneous neurodegenerative disease portfolios in the short term for optimal intervention targeting.

The challenge of aimed intervention at the level of individual disease modifiers will be as much in the area of identification of those unnecessary hallmarks contributing to disease heterogeneity as well as the design of innovative precision medicine approaches to optimize multifactorial therapeutic intervention targeting. Furthermore, it is worth mentioning that it is increasingly understood that disease modifiers can also be therapeutically relevant per se, such as families carrying genes of large absolute risk. New models of these modifiers are also now emerging at the cellular and molecular level. The ability to accurately pinpoint potential therapeutic targets amongst genetic factors or molecular pathways known to be promoted by disease modifiers is a key endpoint of neurodegenerative disease modifiers' discovery. Therefore, what needs to guide their identification is the anticipated ability to therapeutically effect them in

screening innovative preclinical models as well as in the continuum of the drug discovery process reducing uncertainty.

3.2. Data Integration and Multi-Omics Approaches

Deciphering the molecular mechanisms involved in the neurodegenerative process is a tremendous challenge that can be handled by integrating data sources and endophenotypes at different levels of complexity. Currently, the multi-omics approach, combining genomics, transcriptomics, proteomics, and metabolomics, can provide a comprehensive ultra-dimensional view of cellular processes that may result in different types of influence or even conflict. Establishing or proving interactions among different "omics" will allow us to understand the influence of each process level and to shed light on their possible convergence into a common pathological and biological pathway in the peripheral nervous system. The possibility of network-based integrations across biological layers and levels can reveal the interaction among processes and the up- or down-regulation of a "cellular component" and/or "selective PTMs" among diseased or healthy conditions.

Primary data sources, referring to "omics" gastrocnemius profiling in large cohorts of human and animal samples and subjected to technical analyses, are made up of data in different formats. Transcriptomics data are processed with either microarrays or RNA-seq technologies, providing data in the form of numerical probe sets or transcripts. Assessing data properties and their mutual correlation, dealing with the complexity of platforms and cross-technology harmonization represent some of the technical challenges that might improve the quality of multi-omics data and the reliability of integrative bioinformatic processes in network medicine analysis. A convergent multi-omics multi-evaluation system could summarize the effect and strength of the network modification and provide new "molecular" and functional insights in directing drug discovery and validation processes, unraveling potential effective drugs and therapeutic strategies. Large-scale multi-omics datasets combined with powerful machine learning algorithms and AI-based methods have improved our ability to reveal connections between biological systems, progressing the identification of new drug targets or modifiers for certain diseases and enhancing our insight into the interaction among the multiple processes intertwining in nervous system homeostasis. Therefore, the multimodal systematic analysis of nervous system degenerative diseases represents a

critical point, and the application of AI and the multi-omics approach generates novel information for the prediction of early and accurate pathogenesis and disease progression with potential benefits for precise treatments in these pathologies.

3.3. Machine Learning Models for Predicting Disease Modifiers

Predicting disease modifiers is one of the key ongoing tasks in AI research for neurodegenerative diseases. It can be addressed as supervised and unsupervised learning, with possible algorithmic variants like regression, classification, and ensemble methods. The choice of an accurate and suitable prediction model to use depends highly on the nature and availability of data, as well as the specific aim of the study. The choice of features for model design can range from small molecules and gene expression levels to higher-order data representation. Disease modifiers are usually presented in the form of gene-drug, gene-drug-disease, and drug-pathway combinations, associations, or interactions. Recent studies demonstrate promising results from some machine learning techniques in this complex prediction task for different diseases.

Two recent case studies have used ensemble models to predict drug-disease associations in selected disease types. They discuss high area under the curve values of over 0.75 and 0.71 in distinct datasets. One challenge these studies confront is unbalanced data and failure to outperform reference set power-law distributions. Deep learning has been used in recent studies to predict disease modifier associations in several hurdles. It sometimes has a competitive or relatively high positive prediction performance of around 0.96 or higher in area under the curve scores, showcasing the methods' ability to learn latent information from the molecular data. A main downside of deep learning and neural networks, in particular, lies within their low-level interpretability. This can hamper the usage of proposed models in clinical practice, drug discovery, and development. Regularization is commonly used in the studies to tackle overfitting, whereby refactoring models to fit the disease modifier prediction objective is one approach. Research in this field relies on database use and updating, as demonstrated by incremental model retraining and testing as new datasets are published. As a result, the precision of proposed algorithms might shift with the availability of new data. These techniques open research avenues into high-performance algorithms that have significant applications in clinical practice, drug repurposing, and development. The

progression of improved AI models is driving knowledge gains into designing methods to study the role of genes, alleles, and pathways in neurodegenerative disorders.

4. Case Studies and Success Stories

This section aims to report several noteworthy AI-driven platforms, which have yielded commercially interesting drug candidates for different neurodegenerative diseases alone and in collaboration. These case studies are intended to demonstrate some of the recent success stories where different advanced AI methodologies have been effectively applied to improve the drug discovery process. This section also illustrates a number of important aspects that pharmaceutical researchers are normally interested in discussing, such as the following: 1) how the AI approach has helped (e.g., in drug target identification, lead compound optimization, and drug repurposing), 2) methodological choices and their impact on the project (such as data types, systems studied, drug targets, and layers of collaboration between AI and biomedical research communities), 3) technical and/or biological challenges encountered and solutions identified. The potential benefits for the patient and healthcare of these case studies in the neurodegenerative disease field and beyond, as well as the challenges and limitations, will be the focus of the section. Exciting new frontier technologies such as AI are now being introduced at large scale in drug discovery workflows across major pharmaceutical companies, several of which continue to report encouraging progress. In 2020, two pharmacokinetic modifiers for the neurodegenerative disorder Friedreich's ataxia were nominated for clinical development by pharmaceutical companies, after repurposing using network meta-analyses. These commercial developments alone clearly illustrate the significant traction with both industrial and professional healthcare organizations exemplified by the current upsurge in AI systems and technology that are being used in clinical medicine and in the design of biomedical systems. Given the rapidly increasing numbers of commercial AI platforms, requiring detailed explanation, only platforms with a strong emphasis and applications for neurodegenerative diseases were included. These platforms have used several different advanced AI methodologies that have recently demonstrated translational potential, ranging from novel AI-driven drug target prediction tools to AI drug discovery, and will be discussed in the case study section to enable us to critically interrogate which methodologies are advancing in the field, and what future work is deemed as translationally relevant by commercial investors in AI.

4.1. Examples of AI-Driven Drug Discovery Platforms

Several AI-driven platforms aimed at improving drug discovery for a variety of diseases have been developed. Here we present a few examples of such platforms focusing on neurodegenerative diseases that were developed recently or are currently in development. These platforms, driven by machine-learning algorithms and abundant omics data, can transport compound identification, repurposing or target identification, animal model identification, optimization, or virtual testing. As they can bring considerable technological and pharmaceutical advancement, the extent of their development and their goals are highlighted here.

AI-based Target Product Profile and Molecule Identification via Evidence Recommendation presents a methodological framework that streamlines the pipeline for putative target identification using genetic, transcriptomic, and proteomic patient-level data and consequent drug repositioning. It is useful for pharmaceutical industry entrepreneurs and researchers who have a certain therapeutic market in mind but lack a suitable target or cannot select a validated target from a multitude of candidate possibilities. The tools were found to be successful in various early-stage projects and have led to the development of potential treatment approaches and drugs.

MetaPath, an AI-based drug development platform, presents an approach to drug development acceleration on the translational level. The methods used lead to a reduction of several years in drug development, almost supporting drug development failure. The framework constitutes three main similarities to the identified algorithms that integrate different biological data inputs, lowering thresholds for medio-personal profiling constraint as a preclinical test, disease-mechanism use case integration, and accurate validation performance using clinico-biological readouts. Neither of the algorithms has been validated at the clinical testing level in the case of drugs for neurodegenerative diseases. However, the relevant algorithm has already been recognized; on the other hand, both AI engineers and biological researchers are currently working together to develop a translational model of drugs for Alzheimer's disease. Key features of these examples are described in the following sections together with an evaluation of their considerations and limitations.

4.2. Impact of AI in Accelerating Drug Discovery

Having sufficient and effective treatments for diseases requires screening large sets of compounds, validating efficacy through in vitro and in vivo studies, and eventually testing them in clinical trials with human subjects. This so-called drug development pipeline is time-consuming, marked by high risks and has relatively low success rates. AI algorithms can increase the efficiency of drug discovery in both industry and academia, leading to a reduction in lead time for drug approval and a larger increase in success rates. Typical metrics for efficiency gains include a decrease in the mean time to outcome for the drug development pipeline, costs or R&D investments associated with therapeutic discovery, and larger numbers of approved drugs, novel compounds, or repurposed therapeutics. In vivo, several examples exist to measure the speed or costs for the identification of novel compounds or curated drug targets for neurodegenerative diseases, or for faster regulatory approval for drug treatments.

Several case studies showcase the power and promise of new AI-based methodologies for drug repurposing and novel compound discovery. In hyperlipidemic patients who developed the typical symptoms of Alzheimer's due to an unhealthy lifestyle, the AI-based repurposing of statins led to novel compounds associated with a higher efficacy and lower probability of adverse events. Similarly, the faster identification of the JAK-STAT pathway as a therapeutic target in amyotrophic lateral sclerosis and the discovery of B2-adrenergic receptor agonists that can be repurposed for MND led initially to small clinical trials that have shown a high dose of this drug to be effective in aborting the progression of muscle weakness within the first 6-12 months. Finally, the AI-based repurposing of ritonavir has shown its ability to slow down motor and cognitive decline in Parkinson's patients in an early phase II trial. Every one-year delay in the onset of dementia or physical disability, as a consequence of incipient or full-blown drug treatments, respectively, may delay nursing home admissions, which can save long-term care funding. Clearly, by prioritizing the compounds predicted by the AI-based system, this treatment could reach the market faster; it currently takes about 10 years and millions in funding for a new indication with promising phase II clinical data to make it to the market, via so-called phase III registration trials.

5. Challenges and Future Directions

We believe that the AI community has a crucial role to play in addressing the formidable challenges currently facing the drug discovery field to expedite the development of effective drugs for neurodegeneration.

AI Ethics and Regulation

AI ethics and responsible research innovation must be considered at the core of efforts to repurpose AI for the benefit of human health. Research must take into consideration issues like data privacy and how to ensure that AI systems are transparent and accountable. Regulation of AI must also be debated and discussed, focusing on the protection of intellectual property and how to define new, ethical standards for AI-driven technologies that are innovative. There are also considerable knowledge gaps and a lack of system interoperability limiting the integration of AI technologies into standard drug discovery pipelines, pointing to the need for increased interdisciplinary research efforts between computer and information scientists and experimental neuroscientists alongside drug discovery partners from the pharmaceutical industry.

Future Trends in AI-Driven Drug Discovery

Trends in AI applications within drug repurposing point to an increasing use of hybrid AI and experimental strategies, circumventing the constraints associated with integrating complex predictive tools that may not account for unmodeled biological mechanisms. Future exploration of novel AI methods for drug-driven indications is also likely, including emerging technologies like reinforcement learning and exploiting deep learning techniques for natural language processing to obtain, for example, a neural network approach to identify and categorize drugs on the basis of chemical structure and side effects from electronic patient and public reviews. Possible new AI-informed applications of therapeutic interventions in neurodegenerative diseases include the identification of novel treatments for specific patient subtypes through the use of biomarkers in parallel with preventative drugs for high-risk neurodegenerative disease populations. This may include drug combinations designed with different adaptive drug optimization or co-adaptive therapy technologies that are starting to emerge in oncology or neurophysiological optimization of inductors of autologous stem cells to compensate for failing interneuronal and hormonal homeostatic control. Promising

drug-target combinations might be further optimized by in vitro experiments on the organoid models of individuals' brains. Ongoing engagement of regulatory agencies for AI and neurodegenerative diseases may help design ethical standards and regulations further.

5.1. Ethical Considerations in AI-Driven Drug Discovery

While these technologies have the potential to generate groundbreaking scientific discoveries and jump-start drug discovery programs for complex diseases at an unprecedented rate, we must also carefully consider the ethical implications associated with such an undertaking. What responsibility falls on the shoulders of researchers who are able to process large amounts of patient data, risking privacy and disclosure issues? Is educational information coming from social media universally approved or respected within diverse populations, or does it raise issues of bias? When reaching out to patients for participation in large, algorithm and data-driven therapies, is obtaining or not obtaining consent acceptable? Collecting and profiling public responses seems fair, but using this data to inform new products or devices raises numerous ethical dilemmas, including data privacy and ownership.

It is important that AI-informed healthcare is driven not only by the algorithmic capability but also by the ethical principles that we (or should) embrace as healthcare providers. Fairness, transparency, and accountability should not just be trendy catchwords thrown around at industry conferences; they instead need to be charters by which the AI model community and adjacent commercial entities construct and build AI models. Indeed, much of the challenge with AI models is to ensure that the healthcare access playing field is not slanted towards those with access to medical care, computing resources, or both. Thus, there is a corner position for the AI researcher that goes beyond developing an algorithm: they must also be ethicists and always ask if the AI algorithmic instrument or model they build favors certain patient population groups over others. Will the coinciding treatment approaches be equally fair across gender or minority status? Health advocacy is an extremely delicate balance of fair treatment, access, and AI research for the benefit of all stakeholders. When trials fail, there is always another patient affected. Potential regulatory responses to the use of AI-based decision-making in clinical research questions involve asking whether the AI decision instrument should also have guarantees of accountability, robustness, and justifiability, objectives not well

defined in AI research. Perhaps it is time for ethicists, AI modelers, regulators, and patients to participate in a dialogue designed to solve these questions before we over-rotate to contribute commercial and clinical resources. What new guidelines should emerge to enable the ethical development of AI-informed drug discovery programs? The last issue, in my opinion, is education. Practicing individuals in both the biotech and pharma industries must receive business training and understand the direct implications on patient care and the healthcare world.

5.2. Integration of AI with Traditional Drug Discovery Approaches

From the previous sections, it is clear that AI offers a new way of addressing several unsolved challenges in neurodegenerative drug discovery. From the first version of an automated de novo drug design program, it is still an ongoing discussion whether all these AI advancements can be used to replace traditional drug discovery approaches. For now, the most promising and popular approach seeks to integrate AI with traditional drug discovery to complement research based on established procedures. The major strength of including AI technology lies in its potential to make traditional drug discovery more efficient. As already noted, AI is a very good multitask learning and machine learning tool; it can effectively predict molecules that will not work at an early stage, and it has also been effective for the identification and prediction of possible benefits of repurposing non-disease-related medication to combat neurodegenerative diseases.

Indeed, numerous publications have reported drugs that were developed as a result of such a hybrid strategy, at least in the test tube. The integration of machine learning and decision tree classifiers with systems biology has resulted in the identification of cyclosporine as an effective Alzheimer's disease medication. In another Alzheimer's disease-centric example, an ensemble of random forests and support vector machine models was used in conjunction with cell-based phenotypic high-content screening, mRNA gene chip analysis, and SILAC-based protein identification. Integration of several AI methodologies with high-content screening has resulted in a 29% Akaike information criterion certainty in Alzheimer's disease drug discovery. Other research demonstrates the strengths of tradition and AI in synergy, both of which show that at this stage, research is highlighting the benefit of the integration of AI into traditional approaches rather than replacing them. There are challenges here, such as the old-

fashioned stigma attached to changes and the fundamental research changes that are hard to ignore. In order to successfully combine AI and traditional drug discovery research, not only is organizational change required, but it is also important to create a place where experts from various fields can work together. However, there is a need for further discussion on best practices as the model changes, as well as on models of ethical behavior that may arise from the discussion of integration challenges.

5.3. Future Trends and Opportunities

Course Exploration III: Future Trends and Opportunities As AI becomes more embedded in drug discovery, particularly in neurodegenerative diseases, there are a number of key trends emerging in the field. These trends are in the areas of technology and scientific advancement, the enabling and embedding of data and process standards, collaboration, and data management, as well as the adaptation of regulatory systems and the widening of the value chain in drug discovery and development. From a technology point of view, online areas that are emerging take advantage of further advances in deep learning, perhaps hybridizing them with developments in the area of natural language processing. In order to establish these methods as a key part of any investment, providing a high level of scientific understanding is also essential; this will become a greater focus for early AI investments.

Technological trends will see an increase in the sophistication of the underlying data analytic methods used, with a greater embedding of quantitative and systems pharmacology tools, as well as advanced forms of data merging based on AI technologies. Building on these predictive technologies, we expect to see a number of tools available in the precision medicine and patient stratification spaces, beginning to experiment with approaches that integrate prediction with mechanisms in disease progression. The future presents an emerging picture of how AI can pave the way for a new generation of drugs and medicine designed to address the unique molecular signature of a patient's disease. By specializing treatments on an industrial scale in this way, AI can help maximize the earlier chances of intervention and optimize treatment pathways to clinical trials for individual patient groups. AI can help pharmaceutical companies improve the chances of a drug working in a clinical trial by doing a better job of selecting patients likely to respond, by improving patient data flow from clinical trials. Thanks to genetic testing and biobanks, it can make it easier to understand the

biology of those patients who are responding to a specific drug and also help determine the right dose of a new drug more quickly and efficiently. Finally, by helping to develop AI, pharmaceutical companies can not only discover new drugs but also improve clinical trial design in order to help recruiters select the right patients for a new drug. A similar approach to patient selection has the potential to greatly improve outcome monitoring, as evidenced by the activity of a number of AI start-ups focusing on innovative trial design and analytics. Alongside this mathematically enabled approach to science and trials, AI may also play a role in drug repositioning, potentially identifying new drug candidates by their impact on the future nature and progression of diseases. Pharmacokinetics, dynamics, and metabolism are emerging powers to predict the efficacy and safety of new compounds, requiring a commitment to providing high-quality preclinical compound characterization, coupled with sound mathematical modeling when providing predictive power. Critically, although predictive modeling and AI can minimize the risk of late-stage clinical trial failure, robust preclinical and mixture data remain crucial for accurately modeling metabolism across multiple relevant animal models. Practical realities mean much early work has centered on simple compounds, efficacy prediction, as well as more traditional uses in absorption and distribution.

6. Conclusion

In this paper, we have described how AI-driven platforms can enhance the efficiency of drug discovery endeavors focusing on neurodegenerative diseases. These are complex disorders with multiple causative and modulatory factors, making it challenging to develop treatments. The combination of systems biology approaches and AI is therefore crucial to essentially extrapolating new mechanisms and suggesting known drugs to repurpose. These options are created based on the complex intertwining signatures of the already known genetic and cellular pathophysiology. This review shows how AI can enable the efficient prediction of disease-modifying pathways, molecular map reorganization, systems-wide identification of druggable cascades, and the democratic use of multiple data and multi-omics layers to predict pathophysiological trajectories. Digital avatars or patient trajectories, AI-based prediction of patient journeys can act as early screens suggesting novel therapeutic targets in a range of disease-modifying pathways.

In summary, AI has great power to augment the ability to predict an array of mechanisms dysregulated in disease and normal pathways to target disease protection and progression. To continue this great work, multidisciplinary efforts must continue to share data, a range of modalities, and data that truly relate the lives of patients to bring our AI-enhanced patient journeys up to date with data that can support decision-making in clinical circles. Additionally, there is a necessary conversation to be had around the ethics of this intervention and the responsible use of AI in drug discovery; whole new realms of multidisciplinary teamwork have already begun in institutional regulatory affairs and will continue to grow to embrace this great opportunity going forward.