

Article

Using Artificial Intelligence to Advance the Research and Development of Orphan Drugs

Carla Irissarry ^{1,*} and Thierry Burger-Helmchen ^{2,*}¹ Paris School of Economics, 75014 Paris, France² Université de Strasbourg, Université de Lorraine, AgroParisTech, CNRS, INRAE, BETA, 67000 Strasbourg, France

* Correspondence: carla.irissarry@sciencespo.fr (C.I.); burger@unistra.fr (T.B.-H.)

Abstract: While artificial intelligence has successful and innovative applications in common medicine, could its application facilitate research on rare diseases? This study explores the application of artificial intelligence (AI) in orphan drug research, focusing on how AI can address three major barriers: high financial risk, development complexity, and low trialability. This paper begins with an overview of orphan drug development and AI applications, defining key concepts and providing a background on the regulatory framework of and AI's role in medical research. Next, it examines how AI can lower financial risks by streamlining drug discovery and development processes, analyzing complex data, and predicting outcomes to improve our understanding of rare diseases. This study then explores how AI can enhance clinical trials through simulations and virtual trials, compensating for the limited patient populations available for rare disease research. Finally, it discusses the broader implications of integrating AI in orphan drug development, emphasizing the potential for AI to accelerate drug discovery and improve treatment success rates, and highlights the need for ongoing innovation and regulatory support to maximize the benefits of AI-driven research in healthcare. Based on those results, we discuss the implications for traditional and AI-powered business in the drug industry.

Keywords: AI; research; orphan drug; rare disease; drug discovery and industry; business; financial risk; clinical and virtual trials; health economics; development; treatment success

Citation: Irissarry, C.; Burger-Helmchen, T. Using Artificial Intelligence to Advance the Research and Development of Orphan Drugs. *Businesses* **2024**, *4*, 453–472. <https://doi.org/10.3390/businesses4030028>

Academic Editor: Lester Johnson

Received: 25 July 2024

Revised: 28 August 2024

Accepted: 3 September 2024

Published: 9 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rare diseases are estimated to affect around 300 million people [1]. In half of the cases, they lead to sensorial, motor, or intellectual disability [2]. They represent a significant burden: 57.5% to 65% of rare diseases are associated with a reduced lifespan [3], and the cost of living with a rare disease can vary from hundreds of thousands of euros to millions depending on the disease [3]. Due to the high prevalence and burden of rare diseases, there is a moral necessity to develop adequate treatments. Yet, the vast majority of rare diseases have no known cure: in the USA, over 90% of rare diseases do not have an FDA-approved therapy [4]. As of 2021, 207 orphan drugs had a market authorization in the EU [5]. However, this does not imply that 207 rare diseases do have a cure. Treatments for rare diseases are called orphan drugs, and their exact definition varies across countries.

The specificities of rare diseases with regard to common diseases, which will be explored further in this work, are a source of difficulties in the development of orphan drugs. In recent years, many authors have focused on studying the barriers (and opportunities) related to drug development, particularly in the area of orphan drugs [6–11]. The five most frequent barriers to orphan drug development identified are (i) high financial risk, (ii) high complexity of orphan drug development, (iii) low level of trialability, (iv) lack of image improvement, and (v) perception of small non-financial benefits.

Among these barriers, the first three barriers—high financial risk, high complexity of development, and low level of trialability—will be the focus of this study. The barriers related to non-financial incentives (image improvement and small non-financial benefits) will be left out because these barriers can be overcome mainly by legislative measures, which will not be the focus of this study.

In our case, complexity is defined as “*the degree to which an innovation is perceived as being difficult to use*” [6,11] and includes factors such as a lack of knowledge about the disease, inadequate biomarkers, and inappropriate diagnostics. Trialability refers to “*the degree to which an innovation may be experimented with before adoption*” [6] and is hindered by small sample sizes in clinical trials. Although financial incentives in the USA and EU have mitigated some financial risks, uncertainty remains, especially concerning long-term revenue streams. Despite the high prices of such drugs potentially offsetting the costs of R&D, barriers like complexity and trialability continue to impede orphan drug development.

To focus on enabling innovation, we will explore artificial intelligence (AI), a tool increasingly used in various fields, including medicine. AI has driven advances in research, development, diagnosis, and prognosis, and we will examine its potential impact on overcoming these barriers. AI has been a source of advances in research and development, but also in the diagnosis and prognosis of diseases, as we will see. Some examples of these applications are the identification of diabetic retinopathy through image recognition [12] and early detection of anomalies in electrocardiograms [13].

In the second section, an overview of orphan drug development and the application of artificial intelligence in medical research will be presented. The third section will focus on the barriers of complexity and financial resources, demonstrating how AI systems can facilitate the development of new molecules. Finally, the fourth section will address the barriers of trialability and complexity, exploring how AI can overcome these challenges in clinical trials.

2. Overview of Orphan Drug Development and of the Applications of AI in Medical Research

2.1. Definitions

Rare diseases are defined as such according to their low prevalence and the severity of the disease. In the European Union, the term rare disease qualifies “*life-threatening or chronically debilitating diseases*” that affect less than 5 out of 10,000 people [14]. Rare diseases are often orphan diseases, that is, diseases for which no efficient cure is available. However, the distinction between rare diseases and orphan diseases remains important since not all orphan diseases are rare. Research on rare diseases is more difficult than research on common diseases, notably because of the lack of knowledge on disease mechanisms and the small populations of patients in clinical trials [15,16] (see also <https://www.orpha.net/>, accessed on 28 August 2024). A prevalence lower than 1/10,000 makes the research process particularly difficult [17].

The treatments available for rare diseases are called orphan drugs. The concept of orphan drugs was first coined in the USA in 1983 for the Orphan Drug Act, which gave incentives for pharmaceutical companies to develop drugs for rare diseases. This law enabled financial and technical support from the state and the FDA for clinical trials and a 7-year market exclusivity period [18]. The European Union followed with a similar initiative a few years later [19] with regulation n°141/2000 from the European Parliament. This text defines orphan drugs as medicinal products intended for the diagnosis, prevention, or treatment of a rare disease for which no other satisfactory medicinal product is approved in the Community or which represent significant improvement over the existing alternatives [14] (Regulation No. 141/2000 on orphan medicinal products, Article 1). It also states that patients with a rare disease “*should be entitled to the same quality of treatment as other patients*”, justifying the necessity to stimulate research on rare diseases.

For rare diseases and common diseases, finding a new treatment is a tedious process. The development of a drug consists of many steps: the first one is understanding the pathogenesis of a particular disease (pathogenesis is defined by the Merriam-Webster dictionary as “the origination and development of a disease”). Then, fundamental research on potential molecules can start. Once a molecule is selected, it is tested on animal subjects for quality, safety, and efficacy. It is only after these steps that a clinical trial can start to assess its safety, efficiency, and side effects on human subjects [6]. Clinical trials are divided into three phases of testing. The first phase aims at assessing the safety of the molecule: a very small sample of healthy human subjects is administered the molecule and is monitored closely for major side effects [20]. The second phase aims at finding the safest and most efficient dosage of the molecule. It is, most of the time, a double-blind study where the molecule is tested on 100–500 sick patients against either a placebo or an alternative treatment. The third phase aims at confirming the safety and efficacy of the molecule on a large sample of patients (up to 3000) from multiple countries. This phase can last up to several years. After phase three, the molecule can be submitted for approval to regulatory agencies. The approval request is examined by the Committee for Human Medicinal Products (CHMP), a body of the European Medicine Agency (EMA). The CHMP will send its recommendation for or against approval to the European Commission, which will publish the marketing authorization following the CHMP’s decision [18]. Thanks to this EU-wide procedure, an approved drug is granted market access to EU member states, Iceland, Liechtenstein, and Norway [14]. Even after the market authorization, the drug keeps being monitored to evaluate its long-term effects and less common adverse effects. Overall, the process of a clinical trial takes 10 to 12 years [19]. Thus, despite stimulating research on rare diseases, it seems that treatments will not be developed in just a few months and that finding ways to make research easier and faster would benefit this research.

In the Introduction, artificial intelligence was mentioned as a potential tool to facilitate research on rare diseases. Before diving into its potential applications for research on orphan drugs, definitions of key concepts are provided in the next paragraph.

Artificial intelligence is defined by the European Commission as “systems that display intelligent behavior by analyzing their environment and taking actions—with some degree of autonomy—to achieve specific goals” [14]. As of today, we refer to artificial intelligence systems as “narrow” AI, because these systems are typically only able to perform one specialized task. Examples of these types of AI systems are chatbots, facial recognition software, and digital assistants such as Siri or Alexa.

Among the vast set of AI systems, some of them are capable of “learning” or adapting to their environment thanks to machine learning. Machine learning is defined by Microsoft as “the process of using mathematical models of data to help a computer learn without direct instruction. (...) machine learning uses algorithms to identify patterns within data, and those patterns are used to create a data model that can make predictions” [21]. Machine learning is useful for identifying patterns or structures in data, for data mining, and for classifying data. To develop an algorithm that uses machine learning, the first step is to collect and compile data, then train the model with those data, and finally validate it by evaluating its performance and accuracy. A type of machine learning that is often mentioned is deep learning, which uses a type of algorithm structure called neural networks.

2.2. Regulation EC 141/2000: A Turning Point for Orphan Drug Development within the EU Regulatory Framework

The Orphan Drug Act, implemented in 1983 in the USA, had positive repercussions on the development of drugs for rare diseases in the USA. Following this momentum, the European Union followed with a similar legislation: regulation EC 141/2000 [22]. This regulation was the first European legislative text on rare diseases. It led to the market authorization of over 80 orphan drugs between 2000 and 2015 [23] and showed the beginning of a commitment from the European Union to the development of rare disease policies at a

supranational level. In fact, given that the scarcity of knowledge on a particular disease and the small number of patients are major impediments to research on rare diseases, organizing it at the European level enables the sharing of these limited resources and using them more optimally [23]. This text was followed by national strategies and legislations on orphan drugs. It was influential as it set the first step towards a European strategy for research on rare diseases, established criteria for the orphan disease designation and for the definition of rare diseases, and set several incentives for research and development. Indeed, when drugs are designated as orphan drugs under regulation EC 141/2000, they can benefit from reduced EMA fees when they submit their molecule for evaluation. They can also receive free scientific advice on the protocol of the clinical trial from the EMA, which is linked with a higher success rate of the clinical study [23]. Additionally, they receive financial incentives: they can obtain research grants from EU member states and are guaranteed 6 to 10 years of market exclusivity once the drug is approved (Regulation No. 141/2000 on orphan medicinal products, Articles 6, 8, 9). The market exclusivity incentive ensures that the costs of research and development are covered despite the small size of the market [18]. During this period, it is not possible for another manufacturer to request an orphan drug designation for a drug in the same area of application or for one that is similar in terms of its chemical structure and molecular mechanism of action [18].

Hence, the orphan drug designation comes with both financial incentives and scientific advice that can improve the chances of success of the clinical trial. To receive the orphan drug designation, a manufacturer has to send a formal request, which is examined by the Committee for Orphan Medicinal Products (COMP), a body of the EMA, in under 90 days. The committee publishes its opinion on whether the drug meets the criteria for designation according to regulation n°141/2000 [14], and the European Commission officially publishes the granting of orphan drug status within 30 days following the COMP's advice. Once the designation is approved by the Commission, the drug is included in the Community Register of orphan medicinal products for human use. However, since the orphan drug designation occurs early in the research process, the fact that a drug enters the register does not imply that it will reach the market: orphan drugs must also submit their market approval request once the clinical trial has been successful [18]. A unique feature of the market approval process can be used for orphan drugs: if a study does not reach the required statistical significance to prove the efficacy and safety of a molecule because the sample of patients in the trial was too small, but the benefit of the product is generally recognizable, a manufacturer can obtain a Marketing Authorization under Exceptional Circumstances [24] (Regulation No. 726/2004 laying down Community procedures for the authorization and supervision of medicinal products for human and veterinary use and establishing a European Medicines Agency, Article 14). This way, orphan drugs have a chance of receiving market authorization in cases where the disease is too rare to conduct statistically robust clinical trials.

2.3. Artificial Intelligence in Medical Research

Artificial intelligence already has numerous applications in common medicine. It is widely recognized as a means for innovation in pharmaceuticals and major pharmaceutical companies are currently using AI systems: in 2021, Sanofi invested USD 180 million in Owkin's, an artificial intelligence and precision medicine company [25]. In 2019, Eli Lilly sealed a multi-year partnership with the biotech Atomwise, which invented a deep learning AI technology for small-molecule drug discovery, to advance preclinical drug discovery efforts [26]. As many biotech startups are investing the field of AI applications for medical research, different kinds of applications can be distinguished: some aiming at understanding a disease better, some to support diagnosis or clinical decisions, and some to develop new treatments. Let us see some examples of applications of AI in medical research.

AI can be used to better understand some diseases using "multi-omics data". This term refers to "the biological process where different "-omics" data, such as genomics,

proteomics, transcriptomics, epigenomics, and microbiomics, are jointly collected and analyzed" [12]. Machine learning is used to obtain a comprehensive understanding of biological processes and offer a multi-view setting. The use of multi-omics data enables to develop models that explore the relationships between different omics data in order to predict a quantitative phenotype [12]. For example, a model has been developed to predict a drug's cytotoxicity. These data can also be used to build models that predict survival rates in different illnesses or to predict drug resistance in viruses [12].

In addition, artificial intelligence can be used as a support for the diagnosis of patients and clinical decisions. For example, it can be used to analyze medical images. A deep learning model was developed in 2016 to categorize patients by identifying diabetic retinopathy in photographs [27]. The feasibility of applying this to the clinical setting still has to be researched, but it is an interesting lead for the applications of AI in medicine. In order to support clinical decisions, other AI systems analyze data from Electronic Medical Records, which have both structured data like medical images and also unstructured data like doctor's notes from an examination. Using natural language processing (NLP) algorithms, it is possible to use these unstructured data [12]. The start-up Regard implements this technology in order to assist medical practitioners in diagnosing common diseases. The software that Regard calls a "medical co-pilot" can recognize about 50 common medical conditions and is made to be complementary with medical records [28].

In addition, one of the significant challenges in rare disease research is the limited availability of data due to the small patient populations. Traditional research methods often struggle to draw meaningful conclusions from such small datasets. However, AI offers innovative solutions to overcome these limitations, making it a valuable tool in the context of rare diseases. (i) Data augmentation and synthetic data generation: AI can generate synthetic data based on the limited available data, effectively increasing the size of the dataset. This process allows for more comprehensive model training and validation, even when the original dataset is small. By simulating new data points that reflect the characteristics of the disease, AI enhances the robustness and accuracy of predictive models. (ii) Transfer learning: Another AI technique that addresses the small dataset challenge is transfer learning. This method involves pre-training AI models on larger datasets from related medical domains and then fine-tuning them with the specific rare disease data. This approach allows the model to leverage knowledge from broader datasets, improving its predictive performance and reducing the dependency on large datasets specific to the rare disease. (iii) Federated learning: Federated learning enables AI models to be trained across multiple decentralized datasets from different institutions without the need to share sensitive data. This technique allows researchers to collectively utilize small datasets from various locations, increasing the overall dataset size and diversity, which enhances the AI model's ability to generalize and make accurate predictions [29].

Finally, a major application of AI is in drug development. A striking example of the potential that AI has to offer for drug development is the creation of a COVID-19 vaccine by Moderna. While pre-clinical research in vitro and on animal subjects typically lasts 1 to 10 years, Moderna was able to develop a novel mRNA vaccine for COVID-19 in just over two months once the virus' genetic sequence had been published [30]. Thanks to their prior research on mRNA technology and their extensive use of AI systems, the company could design a molecule on the computer and launch the first phase of the clinical study at an unprecedented speed. Another example is in the development of Spinraza: AI was used to simulate genetic splicing mechanisms and predict the efficacy of ASOs in modulating the SMN2 gene's expression. Traditional methods would have relied heavily on trial-and-error approaches, which are time-consuming and less targeted. AI provided a way to streamline this process by focusing on the most promising molecular targets from the outset, thereby accelerating the pathway from discovery to clinical trials.

Hence, many applications of AI are possible in health, using deep learning, natural language processing, image recognition, and other types of algorithms (Table 1). There is not one single AI system that can solve all the problems in medicine, but a multitude of

possible software types and models that can be useful for a particular function, be it diagnosing a given condition, predicting survival rates, assisting medical practitioners, or designing molecules. Some of them have already been implemented in clinical settings: Regard is already in use in multiple healthcare centers in the USA and Moderna's COVID-19 vaccine Spikevax was approved by the FDA on 31 January 2022 [31]. Other examples cited above have been developed by researchers and have yet to be implemented for patients. In the following section, we will investigate how these kinds of technologies can facilitate the development of an orphan drug.

Table 1. Summary of the main uses of AI in the medical sector (source: authors).

Type of AI System Used	Main Outcome	Source
Natural Language Processing	- Exploitation of medical health records for diagnosis purposes - Regard start-up	[12] [28]
Image Identification	Identification of diabetic retinopathy	[27]
Machine Learning	- Study of multi-omics data: prediction of phenotypes, drug cytotoxicity, survival rates - Regard start-up	[12] [30]
Predictive Maintenance	- Reducing equipment failure rates and optimizing resources	[29]
Process Visualization and Simulation	- Improving decision-making in pharmaceutical manufacturing with digital twins	[32,33]
Research and Development (R&D)	- Accelerating drug discovery and development processes, including molecule identification and drug repurposing	[34,35]

3. Decreasing the Barriers of Complexity and Financial Risk: How AI Systems Can Facilitate the Development of a Molecule

3.1. Using AI to Understand the Etiology of Monogenic and Complex Diseases and Drug Repurposing

A significant barrier to the design of effective treatments against a disease is the lack of understanding of its etiology and physiopathology. According to Lee and his co-authors, a key goal of biomedical research is the detailed characterization of the molecular basis of diseases to enable diagnosis and treatment [36]. Now that the sequencing of the human genome is possible, it allows for insights into the genetic basis of a disease and for the identification of disease-associated genes. However, identifying the function of these genes is not an easy task [37]. Before the widespread use of AI, there were challenges in systematically determining the cellular effects of drugs [37]. Although high-throughput screening methods contributed significantly to this area, they often required extensive time and resources. AI systems, such as the Connectivity Map (CMAP), have further enhanced this process by providing a more efficient approach. Developed through collaboration between MIT and Harvard researchers, CMAP offers “a comprehensive catalog of cellular signatures representing systematic perturbation with genetic (reflecting protein function) and pharmacologic (reflecting small-molecule function) perturbagens.” These AI-driven signatures enable researchers to identify useful and previously unrecognized connections more effectively, reducing the potential for side effects and uncovering new therapeutic applications.

In the case of monogenic diseases, that is, diseases caused by a problem in the expression of a single gene, Mears and his co-authors tested three different ways to identify an existing drug–gene link that could be investigated as a potential treatment. They argue

that this method could be used on hundreds of rare diseases: potential targets would be any disorders caused by a recessive allele with mis-sense mutations that preserve the protein localization on the chromosome and some function [38]. Identifying a drug–gene link would serve both as a source of knowledge on the molecular process of a disease from a gene to its expression and as a way to repurpose existing drugs, thus reducing the time needed to discover a treatment. The foundation for this approach is that genes code for the synthesis of proteins by cells, and monogenic diseases can be thought of as dosage problems caused by levels of the protein that is coded by the gene. By modulating mRNA, an intermediary product between DNA and the synthesized protein, a drug can modulate gene expression [38]. To implement this method of treatment of monogenic diseases, relevant drug–gene interactions need to be identified. These interactions between existing drugs and genes can be found by literature mining or by using AI systems such as the Connectivity Map. In their study, Mears and his co-authors compared results obtained from the Connectivity Map approach, literature mining, and their own AI system. Out of a database of 70 genes and 149 existing drugs, the AI approach Connectivity Map had the highest yield of correct identification of drug–gene links: 9.3% of the links found were validated by further in vivo research. These drug–gene links can then be further tested for potential use as repurposed rare disease drugs [38]. Finally, the authors argue that the three approaches that they tested should be used in complementary ways to find as many drug–gene links as possible to advance treatments for genetic diseases.

The drug–gene link approach is promising for this specific type of rare diseases, but it may not be sufficient to develop treatments for most rare diseases, which are the result of more than one gene’s malfunction. In fact, some complex genetic diseases can present with similar symptoms or gene mutations but be caused by very different molecular mechanisms.

A complex disease is a disorder that “results from the contributions of multiple genomic variants and genes in conjunction with significant influences of the physical and social environment” [39]. Diabetes and some cancers are considered complex diseases; this is also the case for some rare diseases. For example, Sjorgen syndrome and systemic lupus erythematosus both appear to be caused by a similar upregulation of some specific genes, but patients have very different symptoms, suggesting that different biological pathways are at play in these two diseases. Complex diseases have only slight variations that make them different from one another; they share some functional and genetic changes but can require different treatments. For this reason, Lee and his co-authors argue that it is necessary to use multiple factors, not just a single gene expression, to differentiate between diseases and learn more about them [36]. In fact, the existing method to study the molecular basis of complex diseases is often a comparison between the genome of healthy and sick patients. The authors argue that studying the expression of the entire genome in individuals is a promising direction for study and for distinguishing between complex diseases. Yet, studies in these areas have been limited in disease coverage or in scale. A major difficulty is that, for a single disease, there may be thousands of different mutations to investigate and thousands of possible drugs, but only few that would work [40]. Thus, in their study, Lee et al. propose a systematic framework, URSA, that uses a large dataset of clinical samples to identify the distinctive molecular characteristics of 335 human diseases [36]. This framework is an AI system that uses machine learning to build disease-specific probabilistic models to estimate disease signals. It can not only distinguish between healthy and sick subjects, but, most importantly, between these 335 diseases. It is promising because it outperforms other, narrower approaches, such as the study of individual genes or the healthy/sick differentiation, to quantify disease signals. The output of the probabilistic model is interpretable, which means that we do not only obtain a result showing the probability of a given disease, but we can learn about the biological processes underlying each prediction. This is highly informative for the purposes of building hypotheses for treatments [36]. Particularly, it can be used as a tool to guide drug repurposing for rare diseases. Because it only requires expression data to make its

predictions, no prior knowledge on a disease is needed. It can associate a rare disease to a well-studied disease with available drugs by identifying similarities in their biological pathways.

In a study, Lee et al. [36] (Figure 1) illustrate this with two rare diseases that have very similar symptoms caused by different biological pathways. The URSA model is able to distinguish the mechanistic difference between these two disorders and to associate each disease with a different common disease whose treatment may be appropriate.

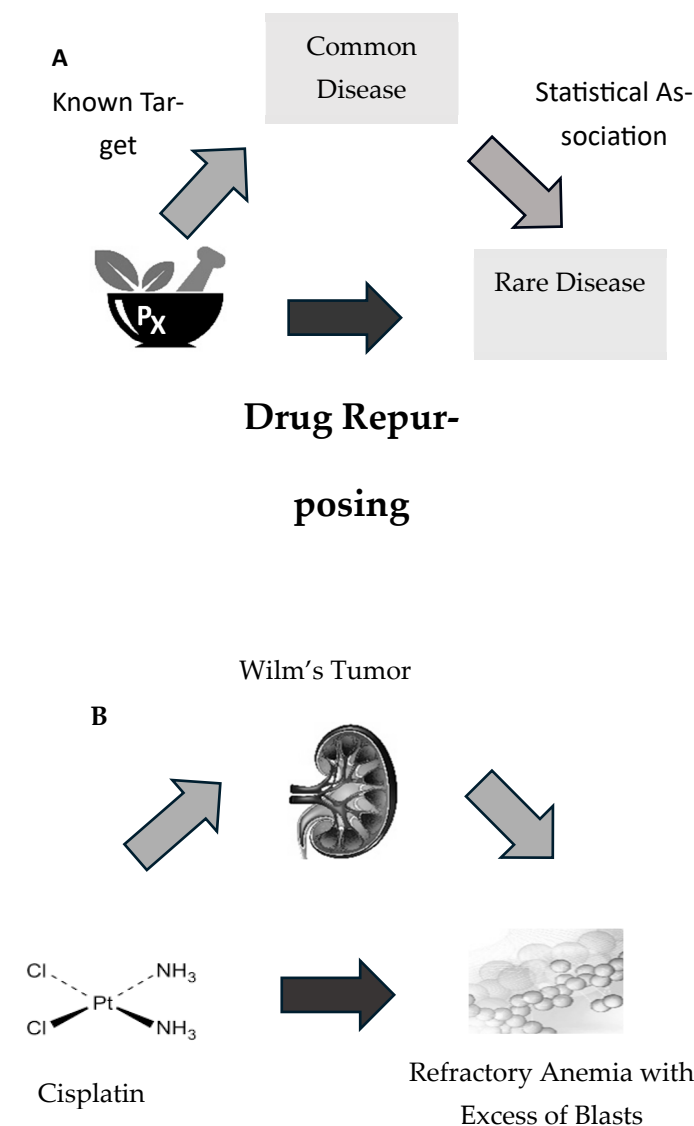


Figure 1. Target drug repurposing from data-driven disease–disease association (adapted from [36]).

The use of AI for drug repurposing is not limited to research and projects for the future but is already being implemented to advance medical research. While AI shows promise in drug repurposing and advancing medical research, its application to orphan diseases with complex or unknown genetic etiology remains controversial. Given that false-positive rates in clinical trials can range from 7% to 15%, the final decision on the therapeutic effectiveness and production of AI-driven orphan drugs should not rely solely on AI and virtual trials. Instead, it is essential to involve multidisciplinary international teams, including experts from the drug industry, IT, medical statistics, legislators, lawyers, and leading figures in clinical trials, to critically assess AI-generated results before moving forward with production. The French company Owkin (Paris, France) developed

an AI system used for drug repurposing and is already implementing it for its research. Specialized in precision medicine, the company uses AI algorithms to analyze disease mechanisms and identify possible diseases that can also be targeted by an existing drug [41]. Another company, BioXcel (Paris, France), uses its own AI algorithms to advance research on neuroscience disorders and immune-oncology. Using big data, they aim, among other goals, at “*drug Re-Innovation*”: *the repurposing of approved drugs or of clinically evaluated drugs candidate to “reduce the expense and time associated with drug development in diseases with substantial unmet medical needs.”* [42]. Since half of rare diseases are neurological [43], using AI to advance medical research on neurological disorders is a meaningful channel to advance research on rare diseases, even though not all neurological diseases are rare.

The Connectivity Map, the URSA model, and proprietary AI algorithms developed by Owkin and BioXcel are possibilities that can be used to learn about a rare disease’s etiology and to find potential drugs for drug repurposing. Understanding the etiology better is a necessary step in medical research on rare diseases and in decreasing the complexity barrier identified by Moors and Faber. Drug repurposing is a way to develop treatments faster and at a lower cost than when trying to develop a new molecule from scratch. Hence, these AI systems would be able to decrease the complexity barrier of developing new treatments and the financial risk barrier.

However, they are no magic solutions and using these tools will also come with challenges. The URSA model can be extended with sufficient training data samples, but it needs large datasets. This highlights the need for publicly available healthy and disease-specific tissue expression data [36]. Moreover, the Connectivity Map is a great tool to identify potential drug–gene links, but its findings still need to be tested further in vivo: in some cases, patients with a disease in which it seems that a gene would need to be activated more and for which a drug has been identified can see their symptoms exacerbated, instead of diminished, by the activation of that gene [38]. Finally, the very project of drug repurposing is not so easy: because of looser patent rules compared to patents for newly developed molecules, drug repurposing does not appear as profitable to pharmaceutical companies as developing a drug from scratch. Thus, they may lack incentives to repurpose drugs [44]. This negative effect may be mitigated by the incentives specific to orphan drugs and further research on the profitability of drug repurposing applied to rare diseases for pharmaceutical companies would be useful in this context.

3.2. Using AI to Design Molecules from Scratch

Drug repurposing appears to be a promising way of using AI to develop new treatments for rare diseases with a reduced complexity barrier and at a lower cost. Additionally, AI systems can enable pharmaceutical companies to develop new drugs from scratch. An example in the context of rare diseases is the development of the drug Spinraza (nusinersen) for Spinal Muscular Atrophy (SMA), a rare genetic disorder. AI was instrumental in accelerating the identification of antisense oligonucleotides (ASOs) that could effectively modulate the splicing of the SMN2 gene, which is critical for SMA treatment. The traditional drug discovery process would have taken much longer to identify such precise targets. The use of AI allowed researchers to rapidly analyze genetic data, model potential therapeutic interventions, and prioritize candidate molecules, significantly reducing the time to clinical trials and eventual approval. We have heard more of such success stories since the COVID-19 pandemic. Moderna’s COVID-19 vaccine is a great example. Moderna’s approach to the vaccine used mRNA instead of a weakened virus. The company had been working on mRNA as a channel for different drugs for a decade before the COVID-19 outbreak. Its processes were largely digitized: the Harvard Business Review describes Moderna (Massachusetts, United States) as an “AI-driven company” [45]. The firm operates on the cloud to store big amounts of data. These data are integrated: lab instruments are connected to each other. All these data generated in the lab are then processed by AI algorithms. Each experiment that is carried out feeds it, which means that

the nine vaccine trials that Moderna carried out before starting to work on the COVID-19 vaccine contributed to its success [45]. Thanks to this, Moderna could design the vaccine on the computer after identifying which protein on the coat of the SARS-CoV-2 virus (the “spike”) to use to induce an immune response. The extensive use of AI in their entire research and development processes enabled Moderna to design a vaccine candidate against the COVID-19 only two days after its gene sequencing was made available. The phase 1 clinical study started merely five weeks after the reception of the gene sequencing data, while this period of identification of a molecule and pre-clinical research typically lasts 4 to 9 years for vaccines [46]. Moderna’s story is not about a rare disease, but rather a virus that caused a pandemic. However, some of its extremely rapid processes for the development of a drug could be applied to research on rare diseases. Some rare diseases are infectious and, if vaccines were to be a relevant way to prevent them, then Moderna’s mRNA technology and the methods they used to design their target molecule in two days might be used for some infectious disease vaccines.

Other companies are using AI to design new molecules from scratch, specifically targeting rare diseases. Deep Genomics uses artificial intelligence “to discover and develop better treatments for genetic diseases, both rare and with a large prevalence” [40]. Deep Genomics uses RNA therapies, like Moderna, but to target genetic conditions. AI is used to discover therapies in two ways: first, it is used to find the “target”: it enables the identification of the disease-causing mutation(s) and potential ways of fixing them. Secondly, it is used to design therapies by assessing hundreds of thousands of potential molecules that are most likely to be efficient. After these two steps, the selected candidate molecules are further tested in the lab. Like Moderna, Deep Genomics improves its AI systems with experience: the data on every molecule identified, as well as data from their clinical trials (such as biomarker data), are collected. These data are fed back into the AI system to improve its future predictions. Since the development of its proprietary AI system, Deep Genomics has updated it to refine its predictions and is working on an updated version that could target more complex genetic diseases [40].

In medicinal chemistry, the design of a new chemical entity with desired properties is called *de novo* design [47]. Among different ways to produce a *de novo* molecule, generative AI can be used to learn from known bioactive chemicals and to design novel ones autonomously, without the need to explicitly include rules for chemical transformation in the algorithm. Merk and his co-authors conclude that generative AI has a promising potential for this task. Thus, once a rare disease is better understood, possibly thanks to the help of AI systems to learn about its etiology, it could also be possible to use AI in the design of a treatment.

3.3. Can the Barriers of Complexity and Financial Risk Really Be Decreased by AI?

AI systems have the potential to reduce costs at various stages of orphan drug development, from drug discovery to clinical trials and manufacturing. By streamlining processes and improving efficiency, AI can help lower the overall financial burden. This reduction in costs can, in turn, increase the expected revenue from orphan drugs, making them more financially viable. While financial risk remains a consideration, AI’s ability to reduce these risks across multiple stages of development provides a promising avenue for alleviating such barriers. The accompanying figure illustrates where AI interventions occur throughout the drug development process. However, using AI systems in an organization may add other types of complexity to the process. In fact, complexity in an organization can be understood as when a large number of elements (such as people, technologies, or products) have many connections to one another. This type of organizational complexity can reduce efficiency within the structure. It can lead to a decrease in how understandable the system is; thus, it may impair the manageability of the organization [48]. Because of heightened interconnectedness, it becomes harder to identify the source of a problem or to remove it once identified: by removing it, one may affect other parts of the structure. By implementing AI algorithms on each part of the value chain, companies like

Moderna and Deep Genomics become complex organizations: every source of data is connected to the others, and every step in the research and development process affects and depends on the other steps. Complexity is even higher: previous and ongoing research and development projects directly affect future projects through their feedback in the AI algorithms. Thus, these companies have quite complex organizations which might make them fragile to some shocks. To gain insight into rare diseases, there may be a trade-off between keeping a moderate level of complexity within the organization and using AI systems that add to organizational complexity but are helpful to the research process.

While AI systems may help reduce development costs, thereby alleviating some financial concerns, it is important to recognize that financial incentives have already addressed many of these risks. The greater challenge often lies in overcoming the complexity of research and the limited trialability of treatments for rare diseases, where AI can also play a crucial role. Yet, the existence of a financial barrier could be nuanced: few empirical studies systematically estimate the cost of orphan drugs development and show such a barrier.

Jayasundara et al. [49] estimate that the clinical cost of developing orphan drugs is significantly lower than that of non-orphan drugs: USD 291 million per approved orphan drug compared to USD 412 million for non-orphan drugs. While this suggests that orphan drug development costs may not be as prohibitive, it is important to note that factors like disease rarity, trial complexity, and limited market size also impact costs and feasibility. Financial barriers are linked to expected revenue, which varies due to differing cost estimates. Although non-orphan drugs may have higher revenue potential due to larger markets, AI could reduce orphan drug development costs, potentially increasing the expected revenue and mitigating financial risks, if they exist.

4. Decreasing the Barriers of Low Trialability and Complexity: How AI Can Facilitate Clinical Trials

4.1. Using AI to Recruit Patients

The limited number of patients who participate in a clinical trial is an important difficulty in clinical trials for orphan drugs. Recruiting patients for trials on common diseases is already a difficult task [50], but the low prevalence of rare diseases and low rates of diagnostics make clinical trials for orphan drugs even more tedious. In fact, patient recruitment represents one-third of the overall duration of a clinical trial [50].

Usually, clinical trials recruit patients using patient registries or disease-specific registries. A patient registry is defined by the EMA as an “Organised system that collects uniform data (clinical and other) to identify specified outcomes for a population defined by a particular disease, condition or exposure. The term ‘patient’ highlights the focus of the registry on health information. It is broadly defined and may include patients with a certain disease, pregnant or lactating women or individuals presenting with another condition such as a birth defect or a molecular or genomic feature”. A disease registry is defined as a “Patient registry whose members are defined by a particular disease or disease-related patient characteristic regardless of exposure to any medicinal product, other treatment or particular health service”. When patients are recruited from registries, the recruitment process is heavily labor-intensive [51]. Medical health records need to be screened in order to check which patients fit the inclusion and exclusion criteria of the study; then, the remaining candidates have an interview with healthcare practitioners to discuss the trial and give, or refuse, their consent to participate in the study. The whole process is time-consuming and expensive. Another problem of patient registries is that the only people in a given registry are the patients who have been registered by healthcare practitioners, and some patients can be missing due to misdiagnosis or lack of awareness about the existence of a registry. Thus, there may be a selection bias of certain characteristics or symptoms that may cause other patients with different characteristics to be left out of the registry and hence also left out of clinical trials.

To speed up the recruitment process, Geva and his co-authors [52] developed a computable phenotype algorithm to conduct data mining on electronic health records in a hospital in order to recruit patients for a trial on pulmonary hypertension, a rare disease. A computational phenotype refers to a set of statistically computable electronic health record data that enables patients of interest to be identified. Within a pediatric hospital in Boston, they compared the patients recruited for the clinical trial through a traditional registry to the patients that their algorithm retrospectively identified through the data mining of electronic health records of the hospital. Their model sought to “*identify patients with a high probability of having the phenotype of interest (...) and thus who can be used as subjects for future studies without further clinician or researcher review*”. While the traditional registry identified 179 patients in the hospital in Boston, the computable phenotype model retrospectively identified 413 patients who may have been eligible for the study. Patients identified by the algorithm and who were not in the registry were younger, more medically complex, and the subject pool had a higher proportion of deceased patients. The two populations were phenotypically different, which shows that some patient populations are not captured in registries. Those patients are important for the understanding of the disease and their identification can enable the clinical trial to be more robust since the pool of patients would be more diverse. Finally, the authors concluded that registry-based methods of recruitment and data mining of electronic health records are complementary methods to identify potential participants for a clinical trial. Another study [51] investigated using an AI screening system for patients’ trial eligibility in real time, which is closer to a real-life setting than the retrospective screening of patients performed by Geva [52]. The implementation of a real-time automated patient screening system successfully recommended potential candidates and reduced the screening time by one-third [51].

Some AI systems for patient recruitment have been developed by the industrial sector: IBM developed a clinical trials matching system that uses natural language processing to process patient and trial data from unstructured sources and match patients to clinical trials. It promoted awareness of clinical trial opportunities and increased enrolment rates in a lung cancer trial [53]. Moreover, the clinical research organization IQVIA (Illkirch-Graffenstaden, France) offers the service IQVIA Core®, a direct-to-patient recruitment process that leverages AI to target the right patients to recruit for clinical trials [54].

Using AI systems to recruit patients more efficiently and quickly and to produce more diverse patient cohorts sounds promising. Yet there are some barriers to the advancement of these technologies. The studies on patient screening mentioned above were carried out in a single hospital. In real clinical trial contexts, the trials are carried out in multiple hospitals, which means that patient data can be dispersed in different sources. Electronic health records from different hospitals have different formats, which could cause interoperability issues. The problem would be especially salient in the absence of clear instruction on how data should be processed prior to being analyzed by the AI system: each hospital might have different practices and would enter very different data into the system. Additionally, all AI algorithms developed in studies might not translate well into the real clinical setting: if an algorithm was built using retrospective data, this creates barriers to its translation into a real clinical trial setting where the staff may want to use it during the recruitment process, not after.

4.2. Using AI to Ensure a Smooth Run of the Clinical Trial

Another way to apply AI systems to clinical trials for orphan drugs is to use its applications throughout the process of each phase. Namely, AI systems can contribute to monitoring patients’ adherence to the study protocol.

For a clinical trial to work, patients must adhere to the study protocol, that is, they must take the right dose at the right time, go to the check-in visits with the healthcare staff, and report everything that is required, like their symptoms. In fact, one of the main causes of the high failure rate of clinical trials is the lack of technical infrastructure to ensure reliable adherence control, patient monitoring, and clinical endpoint detection systems [50].

Data need to be collected reliably and efficiently. Patient adherence needs to last for the entire duration of their protocol, yet the average drop-out rate of patients across clinical trials is about 30% [50]. This poses a problem, because more patients have to be recruited while the study is being rolled out to replace the patients that have withdrawn. This is a time-consuming and costly process that takes resources away from the monitoring of patients in the study, which decreases the quality of the statistical results of the study and sometimes delays the whole clinical trial [50]. The issue is especially relevant to orphan drugs: since rare diseases have a low prevalence, it is overall harder to recruit patients. Hence, recruiting more patients to replace the patients that have withdrawn during the study can take even more time and resources and delay the clinical trial even more than for more common diseases. Finding solutions to decrease the drop-out rate would decrease the trialability barrier and the complexity barrier, and AI systems provide new opportunities in this context.

The first way to ensure a reliable collection of patient data is to use opportunities offered by the Internet of Things combined with AI: wearable sensors and video monitoring can automate the collection of patient data, and AI systems can analyze such data in real time to detect relevant events [50]. If needed, healthcare practitioners could be contacted in case, for example, of major adverse effect collected by the wearable sensors. All the data collected can serve as evidence, or lack thereof, of the patient's adherence to the study protocol. This could be a more reliable system than the one which is currently used, which consists in patients self-monitoring their symptoms and entering the data themselves, which can be subject to error and forgetfulness.

AI systems can also be used to dynamically predict the risk of drop-out for a specific patient: by automatically collecting data, as stated above, if a patient seems to not be taking their doses regularly, or to show any early warning signs of non-adherence to the study protocol, probabilistic models can form a prediction of the likelihood of drop-out and warn healthcare practitioners if this probability exceeds a given threshold. Then, the staff of the clinical trial can contact the patient and address the root cause of the patient's signs of low adherence, instead of missing the clue and having the patient drop out a few weeks or months later [50]. This preventive approach to study withdrawal could be a good way to decrease the drop-out rate of studies and enable a smoother run of clinical trials.

As AI systems increasingly contribute to the design of new molecules and virtual trial simulations, it is essential to establish robust control and validation methods to ensure their reliability. In the near future, a combination of advanced computational verification, real-world data integration, and phased validation approaches could be assumed to assess AI-generated outcomes. Given that orphan drugs often target diseases leading to sensorial, motor, and intellectual disabilities, which complicate traditional clinical trials, these AI-driven approaches must be supplemented with rigorous real-world evidence and adaptive trial designs. Such methods would help to mitigate the inherent unreliability of clinical trials in this context and ensure that AI-designed treatments are both safe and effective.

As the field of AI in drug design and virtual trials progresses, establishing control mechanisms will be critical to the safe and effective application of these technologies. One possible method involves the use of "in silico" validation, where AI-generated results are cross-checked with existing biological and pharmacological data before moving into physical trials. Additionally, integrating AI-generated data with real-world evidence (RWE) from patient registries and health records could provide an additional layer of validation. For orphan drugs, which often involve patients with severe disabilities, adaptive trial designs that allow for ongoing modifications based on intermediate results could be particularly beneficial. These methods, combined with traditional regulatory oversight, will be essential to ensure that AI-designed molecules meet the necessary safety and efficacy standards before they are introduced to the market.

Such a solution has been developed and commercialized by AiCure (Patient Connect—Remote Patient Monitoring Solution) [55]. Their phone application, AiCure Patient

Connect, visually and automatically confirms patient identity, medication, and medication ingestion. Patients receive automated reminders and precise dosing instructions and are then identified. When the study protocol is not abided by, healthcare providers are informed and can engage with participants via a chat function in the application. This system also provides the possibility to register symptoms and development. These data, along with the AI-driven data collection of adherences to the protocol, can give important information to the healthcare staff working on the clinical trial. This solution has been tested in a study on treatment adherence to direct oral anticoagulants [56], which is a therapy that requires a high level of adherence to be effective. Introducing the use of the application AiCure Patient Connect to the therapeutic protocol was found to lead to an absolute improvement of 67% in patients taking their therapy. The result, obtained in a clinical setting with sick patients, shows the potential of AiCure to increase adherence to treatment protocols, including in the context of clinical trials.

A major difficulty in the wide-scale implementation of such a solution is the question of personal data protection. Personal data would need to be collected and anonymized in a way conform to the European General Data Protection Regulation (GDPR) [57], especially because medical data are sensitive, confidential data. The access to and the use and control of patient data by a private company like AiCure poses the first challenge, and proper regulations must apply to ensure data protection. Sensitive patient health data owned by a private company have previously been transferred to another jurisdiction where other data regulations apply: an example is the controversy that happened when Google took over DeepMind while DeepMind was being used in the United Kingdom by the NHS Foundation Trust to implement AI in the management of acute kidney injury. Since Google is an American company, health data from British patients were transferred to the USA and hence fell under the American regulation [58]. If AI systems were to be used to monitor patients in clinical trials, careful attention should be given to what company's app will monitor them, which jurisdiction it is from, and what data protection rules would apply. Patients who consent to participate in a clinical trial do not necessarily consent to the transfer of their data to private entities in another country where their sensitive data are potentially less protected, and they should be made fully aware of how their data will be owned and used in order to consent not only to the participation in the development of a drug, but also to the management of their private data. Murdoch and his co-authors claim that patient data should be regulated to remain in the jurisdiction from which it is obtained [58].

There is also the question of privacy breaches that may lead some patients using such AI systems for a clinical trial to be identified and face consequences. In fact, several studies show that it is possible to re-identify people quite effectively, even from a database where all the information has been anonymized and scrubbed of all identifiers [58]. Further research and policy should focus on ways to mitigate this risk, for example through the use of generative models that can generate realistic but synthetic patient data to train machine learning algorithms. On the legal side, contracts should be made carefully between private entities, public healthcare services, and patients. A patient's right to not only withdraw from the clinical trial, which is already implemented, but also to withdraw their data, should be clearly communicated and exercisable. States should work on implementing an adequate regulatory framework to protect citizens and their data. Table 2 proposes a summary of the potential uses of AI systems for clinical trials (+ signs indicate the relative importance of each point).

Table 2. Summary of potential uses of AI systems for clinical trials (sources: [50] and authors).

Design Themes	Cohort Composition/Patient Recruitment					Patient Monitoring			
Features	Suitability		Eligibility	Empowerment	Motivation	Adherence	Endpoint Detection	Retention	
Methodology	Clinical trial enrichment Biomarker verification			Clinical trial matching			Automatic event logging Drop-out risk forecast and intervention		
Functionality	Reduced population heterogeneity	Prognostic enrichment	Predictive enrichment	Automatic eligibility assessment	Simplification of trial description	Automatic trial recommendation	Disease diary Disease episodes, Medication administration, Health monitoring	Study protocol diary Medication administration, Record-keeping	Patient coaching Proactive intervention to prevent Drop-out
AI techniques	Machine learning/Deep learning Reasoning			Machine learning/Deep learning Reasoning Human-machine interfaces		Machine learning/Deep learning Human-machine interfaces		Machine learning/Deep learning Reasoning Human-machine interfaces	
Data	EMR Omics Medical literature Clinical domain knowledge		Clinical trial databases Trial announcements Medical literature Eligibility databases Social media EMR		Internet of Things and wearables Speech Video				
Outcomes	Optimized cohort composition ++ More effective trial planning and faster launch + Maximized chances for successful outcome ++ Faster and less expensive trials +		Optimized cohort composition ++ More effective trial planning and faster launch ++ Maximized chances for successful outcome + Faster and less expensive trials ++		Maximized chances for successful outcome + Faster and less expensive trials ++		Maximized chances for successful outcome + Faster and less expensive trials ++		

++ Increase very significantly, + increase significantly.

5. Conclusions

This article explored how AI systems can significantly contribute to the development of orphan drugs by reducing financial costs, simplifying the research process through advancing learning about rare diseases, and facilitating clinical trials. In the European Union, the development of orphan drugs is supported by regulation n°141/2000, which defines orphan drugs and establishes a regulatory framework for market approval and incentives. While these incentives, such as financial aid and scientific advice, play a crucial role, AI offers a complementary approach by addressing barriers that these incentives alone cannot fully overcome.

The widespread use of AI in medical research presents a promising opportunity for patients with rare diseases, increasing the likelihood of finding treatments tailored to their specific conditions. However, alongside clinical effectiveness, it is essential to prioritize patient and family satisfaction and safety. The success of AI-assisted orphan drugs depends not only on improving clinical outcomes but also on enhancing the quality of life and ensuring safety. AI can also improve patient recruitment for clinical trials, ensuring that more patients gain access to appropriate studies.

Legislators have a pivotal role in promoting the adoption of AI in rare disease research. Establishing open databanks with genomic and tissue data from both healthy individuals and those affected by rare diseases would be a valuable resource for AI development. Protecting patient data, regulating data ownership, and empowering patients to control their data are equally important. The General Data Protection Regulation (GDPR) in the EU provides a foundation, but regulations must evolve to keep pace with advancements in AI technology. Policymakers should leverage existing academic research on data protection to address potential gaps and ensure that AI systems advance public health while safeguarding individual privacy.

The impact on the economics of this industry and the firm is potentially tremendous. Of course, it has always been difficult to evaluate the economic impact of new technologies in the domains of pharmaceutical and drug development [59,60]. However, as mentioned in this work, the increase in speed and reducing the costs of orphan drugs can easily be transplanted to more general drug development [61]. Therefore, firms involved in orphan drug development can significantly benefit from the implementation of AI systems within their organizations. To leverage these benefits, it is recommended that they implement AI systems into their processes, which can be achieved without necessarily increasing the complexity of the system but can even eventually decrease the complexity [62]. This can involve systematic data collection and the use of cloud computing or of any other relevant infrastructure. By embracing AI within the limits of adequate organizational complexity, firms can potentially gain better clinical trial processes with shorter timelines and reduced costs. However, such technological development should not be achieved regardless of ethics and the more general purpose of the innovation developed [63]. It has been shown that the decrease in cost will make the financing of some drugs much easier [64], decreasing barriers to entry and changing the rules of the business game [65].

Considering the complexities of orphan diseases, particularly those with genetic or multifactorial etiologies, the integration of AI into drug development must be approached with caution. AI's potential for generating false positives necessitates a robust, multidisciplinary evaluation framework. We propose that the final decision on the therapeutic effectiveness of AI-driven orphan drugs, post virtual and clinical trials, should be made by multidisciplinary international teams. These teams should consist of experienced specialists from the drug industry, IT professionals, medical statisticians, legislators, legal experts, and renowned leaders in clinical trials. Their collective expertise would provide a comprehensive assessment, ensuring that AI-generated therapies are both scientifically sound and ethically viable before they reach the market [66].

While AI-assisted orphan drugs offer new avenues for treatment, their success must be evaluated not only through clinical outcomes but also through the lens of patient and family satisfaction. The unique challenges faced by patients with rare diseases and their

relatives make it essential to consider their perspectives when assessing treatment efficacy. Satisfaction encompasses both the quality-of-life improvements provided by the treatment and the assurance of safety, particularly given the innovative nature of AI-assisted therapies. As such, patient-reported outcomes and satisfaction surveys should be integrated into the evaluation process for AI-driven orphan drugs, ensuring that these treatments meet the holistic needs of patients and their families [67].

In terms of potential further research, the technical, ethical, and regulatory challenges associated with the use of AI in healthcare need more exploration. The implementation of AI algorithms that extract data from electronic health records on a large scale is hindered by persistent concern surrounding interoperability among different hospitals and countries. Research efforts should be directed towards developing solutions that enable seamless data exchange and utilization. Additionally, further research should be directed towards algorithms' performance when algorithms trained on retrospective data are applied to new data for prospective purposes. Developing strategies to improve algorithm performance is important for the successful implementation of AI systems in orphan drug development.

Author Contributions: Conceptualization, C.I.; methodology, C.I.; validation, C.I. and T.B.-H.; formal analysis, C.I.; investigation, C.I.; resources, C.I.; data curation, C.I.; writing—original draft preparation, C.I.; writing—review and editing, C.I. and T.B.-H.; supervision, T.B.-H.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: Carla Irissarry is a master student in applied economics at the Paris School of Economics (PSE), and a Sciences Po and University of Strasbourg alumna. Her main interests are health economics and public economics. This research was conducted during her affiliation with the Faculty of Economics and Management of the University of Strasbourg under the supervision of Professor Thierry Burger-Helmchen. Thierry Burger-Helmchen has benefited from the support of the Fondation de l' Université de Strasbourg through the research project AOC "Agilité Organisationnelle et Créativité", as well as from the ANR ECO-INNOVATE project managed by Sophie Bollinger. (University of Strasbourg).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nguengang Wakap, S.; Lambert, D.M.; Olry, A.; Rodwell, C.; Gueydan, C.; Lanneau, V.; Murphy, D.; Le Cam, Y.; Rath, A. Estimating Cumulative Point Prevalence of Rare Diseases: Analysis of the Orphanet Database. *Eur. J. Hum. Genet.* **2020**, *28*, 165–173. <https://doi.org/10.1038/s41431-019-0508-0>.
2. DGOS. *Les Maladies Rares*; Ministère de la Santé et de la Prévention: Paris, France 2023.
3. Ferreira, C.R. The Burden of Rare Diseases. *Am. J. Med. Genet. Part A* **2019**, *179*, 885–892. <https://doi.org/10.1002/ajmg.a.61124>.
4. National Organization for Rare Disorders. *Orphan Drugs in the United States: An Examination of Patents and Orphan Drug Exclusivity*; National Organization for Rare Disorders: Danbury, CT, USA, 2021.
5. European Medicine Agency. *Annual Report on the Use of the Special Contribution for Orphan Medicinal Products*; EMA: London, UK, 2023.
6. Moors, E.H.M.; Faber, J. *Orphan Drugs: Unmet Societal Need for Non-Profitable Privately Supplied New Products*; Elsevier: Amsterdam, The Netherlands, 2007; Volume 19.
7. Althobaiti, H.; Seoane-Vazquez, E.; Brown, L.M.; Fleming, M.L.; Rodriguez-Monguio, R. Disentangling the Cost of Orphan Drugs Marketed in the United States. *Healthcare* **2023**, *11*, 558. <https://doi.org/10.3390/healthcare11040558>.
8. Nuijten, M.; Capri, S. Valuation of Medical Innovation Handling with Uncertainty and Risk. *J. Mark. Access Health Policy* **2024**, *12*, 199–208. <https://doi.org/10.3390/jmahp12030016>.
9. Douglas, C.M.W.; Aith, F.; Boon, W.; de Neiva Borba, M.; Doganova, L.; Grunebaum, S.; Hagendijk, R.; Lynd, L.; Mallard, A.; Mohamed, F.A.; et al. Social Pharmaceutical Innovation and Alternative Forms of Research, Development and Deployment for Drugs for Rare Diseases. *Orphanet J. Rare Dis.* **2022**, *17*, 344. <https://doi.org/10.1186/s13023-022-02476-6>.

10. Zimmermann, B.M.; Eichinger, J.; Baumgartner, M.R. A Systematic Review of Moral Reasons on Orphan Drug Reimbursement. *Orphanet J. Rare Dis.* **2021**, *16*, 292. <https://doi.org/10.1186/s13023-021-01925-y>.
11. Belousova, O.A.; Groen, A.J.; Ouendag, A.M. Opportunities and Barriers for Innovation and Entrepreneurship in Orphan Drug Development. *Technol. Forecast. Soc. Change* **2020**, *161*, 120333. <https://doi.org/10.1016/j.techfore.2020.120333>.
12. Wang, F.; Preininger, A. AI in Health: State of the Art, Challenges, and Future Directions. *Yearb. Med. Inform.* **2019**, *28*, 16–26. <https://doi.org/10.1055/s-0039-1677908>.
13. Yao, X.; Rushlow, D.R.; Inselman, J.W.; McCoy, R.G.; Thacher, T.D.; Behnken, E.M.; Bernard, M.E.; Rosas, S.L.; Akfaly, A.; Misra, A.; et al. Artificial Intelligence-Enabled Electrocardiograms for Identification of Patients with Low Ejection Fraction: A Pragmatic, Randomized Clinical Trial. *Nat. Med.* **2021**, *27*, 815–819. <https://doi.org/10.1038/s41591-021-01335-4>.
14. European Parliament. *Regulation 141/2000 Regulation (EU) No 141/2000 of the European Parliament and of the Council of the 16 December 1999 on Orphan Medicinal Products*; European Parliament: Strasbourg, France, 1999.
15. Bell, S.A.; Tudur Smith, C. A Comparison of Interventional Clinical Trials in Rare versus Non-Rare Diseases: An Analysis of ClinicalTrials.gov. *Orphanet J. Rare Dis.* **2014**, *9*, 170. <https://doi.org/10.1186/s13023-014-0170-0>.
16. Wästfelt, M.; Fadeel, B.; Henter, J.-I. A Journey of Hope: Lessons Learned from Studies on Rare Diseases and Orphan Drugs. *J. Intern. Med.* **2006**, *260*, 1–10. <https://doi.org/10.1111/j.1365-2796.2006.01666.x>.
17. Aymé, S. État de La R&D Dans Les Maladies Rares, 10 Ans Après Le Règlement Européen Du Médicament Orphelin. *Presse Méd.* **2010**, *39*, 5–7. <https://doi.org/10.1016/j.lpm.2010.02.001>.
18. Sprecher, F. Orphan Drugs and Paediatric Medicinal Products. *Pharm. Policy Law* **2013**, *15*, 173–194. <https://doi.org/10.3233/PPL-130369>.
19. Maresova, P.; Klimova, B.; Kuca, K. Financial and Legislative Aspects of Drug Development of Orphan Diseases on the European Market—A Systematic Review. *Appl. Econ.* **2016**, *48*, 2562–2570. <https://doi.org/10.1080/00036846.2015.1125433>.
20. Roche. What Is a Clinical Trial and How Does It Work? Available online: <https://www.roche.com/innovation/process/clinical-trials/about> (accessed on 13 March 2023).
21. Microsoft Azure. What Is Machine Learning? Available online: <https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-machine-learning-platform> (accessed on 13 March 2023).
22. Maresova, P.; Klimova, B.; Kuca, K. Legislation, Regulation and Policies Issues of Orphan Drugs in Developed Countries from 2010 to 2016. *J. Appl. Biomed.* **2018**, *16*, 175–179. <https://doi.org/10.1016/j.jab.2018.04.002>.
23. Rodwell, C.; Aymé, S. Rare Disease Policies to Improve Care for Patients in Europe. *Biochim. Biophys. Acta (BBA) Mol. Basis Dis.* **2015**, *1852*, 2329–2335. <https://doi.org/10.1016/j.bbadis.2015.02.008>.
24. *Regulation 726/2004 Regulation (EU) No 726/2004 of the European Parliament and of the Council Laying Down Community Pro-Cedures for the Authorisation and Supervision of Medicinal Products for Human and Veterinary Use and Establishing a European Medicines Agency*; European Parliament: Strasbourg, France, 2004.
25. Sanofi. *Sanofi Invests \$180 Million Equity in Owkin's Artificial Intelligence and Federated Learning to Advance Oncology Pipeline*; Sanofi: Paris, France, 2021.
26. Atomwise. *Atomwise Announces Multi-Year, Multi-Target Agreement with Lilly for AI- Driven Drug Discovery Program*; Atomwise: San Francisco, CA, USA, 2019.
27. Gulshan, V.; Peng, L.; Coram, M.; Stumpe, M.C.; Wu, D.; Narayanaswamy, A.; Venugopalan, S.; Widner, K.; Madams, T.; Cuadros, J.; et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *J. Am. Med. Assoc.* **2016**, *316*, 2402–2410. <https://doi.org/10.1001/jama.2016.17216>.
28. Shu, C. *Regard Serves as a "Medical Co-Pilot" for Busy Physicians*; TechCrunch: San Francisco, CA, USA, 2022.
29. Miozza, M.; Brunetta, F.; Appio, F.P. Digital Transformation of the Pharmaceutical Industry: A Future Research Agenda for Management Studies. *Technol. Forecast. Soc. Change* **2024**, *207*, 123580. <https://doi.org/10.1016/j.techfore.2024.123580>.
30. Iansiti, M.; Lakhani, K.R.; Mayer, H.; Herman, K. *Moderna (A)*; Harvard Business School: Boston, MA, USA, 2020; Volume 9–621–032.
31. *Moderna COVID-19 Vaccines*; FDA: Silver Spring, MD, USA, 2023.
32. Chen, Y.; Yang, O.; Sampat, C.; Bhalode, P.; Ramachandran, R.; Ierapetritou, M. Digital Twins in Pharmaceutical and Biopharmaceutical Manufacturing: A Literature Review. *Processes* **2020**, *8*, 1088. <https://doi.org/10.3390/pr8091088>.
33. Coito, T.; Martins, M.S.E.; Firme, B.; Figueiredo, J.; Vieira, S.M.; Sousa, J.M.C. Assessing the Impact of Automation in Pharmaceutical Quality Control Labs Using a Digital Twin. *J. Manuf. Syst.* **2022**, *62*, 270–285. <https://doi.org/10.1016/j.jmsy.2021.11.014>.
34. Garza Ramos, A.; Daim, T.; Gaats, L.; Hutmacher, D.W.; Hackenberger, D. Technology Roadmap for the Development of a 3D Cell Culture Workstation for a Biomedical Industry Startup. *Technol. Forecast. Soc. Change* **2022**, *174*, 121213. <https://doi.org/10.1016/j.techfore.2021.121213>.
35. Wu, Q.; He, Q. DIY Laboratories and Business Innovation Ecosystems: The Case of Pharmaceutical Industry. *Technol. Forecast. Soc. Change* **2020**, *161*, 120336. <https://doi.org/10.1016/j.techfore.2020.120336>.
36. Lee, Y.; Krishnan, A.; Oughtred, R.; Rust, J.; Chang, C.S.; Ryu, J.; Kristensen, V.N.; Dolinski, K.; Theesfeld, C.L.; Troyanskaya, O.G. A Computational Framework for Genome-Wide Characterization of the Human Disease Landscape. *Cell Syst.* **2019**, *8*, 152–162. <https://doi.org/10.1016/j.cels.2018.12.010>.
37. BROAD Institute. *Connectivity Map (CMAP)*; Broad Institute: Cambridge, MA, USA, 2015.

38. Mears, A.J.; Schock, S.C.; Hadwen, J.; Putos, S.; Dymont, D.; Boycott, K.M.; MacKenzie, A. Mining the Transcriptome for Rare Disease Therapies: A Comparison of the Efficiencies of Two Data Mining Approaches and a Targeted Cell-Based Drug Screen. *Npj Genom. Med.* **2017**, *2*, 1–8. <https://doi.org/10.1038/s41525-017-0018-3>.
39. National Human Genome Research Institute. *Complex Disease*; Genome.Gov: Bethesda, MD, USA, 2022.
40. Deep Genomics Therapeutic Approach. Available online: https://www.deepgenomics.com/therapeutic_focus/www.deepgenomics.com/therapeutic_focus/ (accessed on 23 April 2023).
41. Owkin. *Discover—Advancing Drug Discovery*; OWKIN: New York, NY, USA.
42. BioXcel AI Platform—BioXcel Therapeutics Inc. Available online: <https://www.bioxceltherapeutics.com/ai-platform/> (accessed on 15 March 2023).
43. The Lancet. Neurology Rare Diseases: Maintaining Momentum. *Lancet Neurol.* **2022**, *21*, 203. [https://doi.org/10.1016/S1474-4422\(22\)00046-1](https://doi.org/10.1016/S1474-4422(22)00046-1).
44. Sexton, J. *Repurposing Generic Drugs Can Reduce Time and Cost to Develop New Treatments—But Low Profitability Remains a Barrier*; The Conversation: Melbourne, Australia, 2022.
45. Iansiti, M.; Lakhani, K.R. Competing in the Age of AI. *Harv. Bus. Rev.* **2020**, *98*, 60–67.
46. Broom, D. *5 Charts That Tell the Story of Vaccines Today*; World Economic Forum: Cologny, Switzerland, 2020.
47. Merk, D.; Friedrich, L.; Grisoni, F.; Schneider, G. De Novo Design of Bioactive Small Molecules by Artificial Intelligence. *Mol. Inform.* **2018**, *37*, 1700153. <https://doi.org/10.1002/minf.201700153>.
48. Reeves, M.; Levin, S.; Fink, T.; Levina, A. Taming Complexity. Making Sure the Benefits of Any Addition to an Organization's Systems Outweight Its Costs. *Harv. Bus. Rev.* **2020**, *113*, 1–11.
49. Jayasundara, K.; Hollis, A.; Krahn, M.; Mamdani, M.; Hoch, J.S.; Grootendorst, P. Estimating the Clinical Cost of Drug Development for Orphan versus Non-Orphan Drugs. *Orphanet J. Rare Dis.* **2019**, *14*, 12. <https://doi.org/10.1186/s13023-018-0990-4>.
50. Harrer, S.; Shah, P.; Antony, B.; Hu, J. Artificial Intelligence for Clinical Trial Design. *Trends Pharmacol. Sci.* **2019**, *40*, 577–591. <https://doi.org/10.1016/j.tips.2019.05.005>.
51. Ismail, A.; Al-Zoubi, T.; El Naqa, I.; Saeed, H. The Role of Artificial Intelligence in Hastening Time to Recruitment in Clinical Trials. *BJR Open* **2023**, *5*, 20220023. <https://doi.org/10.1259/bjro.20220023>.
52. Geva, A.; Gronsbell, J.L.; Cai, T.; Cai, T.; Murphy, S.N.; Lyons, J.C.; Heinz, M.M.; Natter, M.D.; Patibandla, N.; Bickel, J.; et al. A Computable Phenotype Improves Cohort Ascertainment in a Pediatric Pulmonary Hypertension Registry. *J. Pediatr.* **2017**, *188*, 224–231. <https://doi.org/10.1016/j.jpeds.2017.05.037>.
53. Leventakos, K.; Helgeson, J.; Mansfield, A.; Deering, E.; Schwecke, A.; Adjei, A.; Molina, J.; Hocum, C.; Parikh, K.; Halfdanarson, T.; et al. P1.16-14 Effects of an Artificial Intelligence (AI) System on Clinical Trial Enrollment in Lung Cancer. *J. Thorac. Oncol.* **2019**, *14*, S592. <https://doi.org/10.1016/j.jtho.2019.08.1240>.
54. IQVIA. Patient Recruitment. Available online: <https://www.iqvia.com/solutions/research-and-development/clinical-trials/patient-recruitment> (accessed on 15 June 2023).
55. AiCure. Patient Connect—Remote Patient Monitoring Solution. AiCure. 2023. Available online: <https://aicure.com/solutions/patient-connect> (accessed on 1 January 2024).
56. Labovitz, D.L.; Shafner, L.; Reyes Gil, M.; Virmani, D.; Hanina, A. Using Artificial Intelligence to Reduce the Risk of Nonadherence in Patients on Anticoagulation Therapy. *Stroke* **2017**, *48*, 1416–1419. <https://doi.org/10.1161/STROKEAHA.116.016281>.
57. *Regulation 2016/679 Regulation (EU) No 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)*; European Parliament: Strasbourg, France, 2016.
58. Murdoch, B. Privacy and Artificial Intelligence: Challenges for Protecting Health Information in a New Era. *BMC Med. Ethics* **2021**, *22*, 122. <https://doi.org/10.1186/s12910-021-00687-3>.
59. McGrath, R.G.; Nerkar, A. Real Options Reasoning and a New Look at the R&d Investment Strategies of Pharmaceutical Firms. *Strateg. Manag. J.* **2004**, *25*, 1–21. <https://doi.org/10.1002/smj.358>.
60. Burger-Helmchen, T. Option Chain and Change Management: A Structural Equation Application. *Eur. Manag. J.* **2009**, *27*, 176–186. <https://doi.org/10.1016/j.emj.2008.08.004>.
61. Napolitano, G.; Has, C.; Schwerk, A.; Yuan, J.-H.; Ullrich, C. Potential of Artificial Intelligence to Accelerate Drug Development for Rare Diseases. *Pharm. Med.* **2024**, *38*, 79–86. <https://doi.org/10.1007/s40290-023-00504-9>.
62. Heraud, J.-A.; Kerr, F.; Burger-Helmchen, T. *Creative Management of Complex Systems*; Wiley-ISTE: Hoboken, NJ, USA, 2019; ISBN 978-1-84821-957-1.
63. Neukam, M.; Bollinger, S. Encouraging Creative Teams to Integrate a Sustainable Approach to Technology. *J. Bus. Res.* **2022**, *150*, 354–364. <https://doi.org/10.1016/j.jbusres.2022.05.083>.
64. Stefani, U.; Schiavone, F.; Laperche, B.; Burger-Helmchen, T. New Tools and Practices for Financing Novelty: A Research Agenda. *Eur. J. Innov. Manag.* **2019**, *23*, 314–328.
65. Sinha, M.S.; Stern, A.D.; Rai, A.K. Four Decades of Orphan Drugs and Priorities for the Future. *N. Engl. J. Med.* **2024**, *391*, 100–102. <https://doi.org/10.1056/NEJMp2401487>.

-
66. Costa, E.; Isgro, A.; de Montalembert, M.; Leufkens, H.G.M.; Ware, R.E.; De Franceschi, L. Successes and Pitfalls in Orphan Drug Development for Sickle Cell Disease. *Blood Adv.* **2024**, *8*, 2455–2465. <https://doi.org/10.1182/bloodadvances.2023011730>.
 67. Agyralides, G. The Future of Medicine: An Outline Attempt Using State-of-the-Art Business and Scientific Trends. *Front. Med.* **2024**, *11*, 1391727. <https://doi.org/10.3389/fmed.2024.1391727>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.