applied sciences

MDPI

Review

Holistic Approach for Artificial Intelligence Implementation in Pharmaceutical Products Lifecycle: A Meta-Analysis

Konstantin A. Koshechkin [1], Georgiy S. Lebedev [1], Eduard N. Fartushnyi [1] and Yuriy L. Orlov [1,2,*]

[1] The Digital Health Institute, I.M. Sechenov First Moscow State Medical University of the Ministry of Health of the Russian Federation (Sechenov University), 119991 Moscow, Russia
[2] Agrarian and Technological Institute, Peoples' Friendship University of Russia, 117198 Moscow, Russia
* Correspondence: y.orlov@sechenov.ru

Abstract: Recent developments in Digital Medicine approaches concern pharmaceutical product optimization. Artificial Intelligence (AI) has multiple applications for pharmaceutical products' lifecycle, increasing development speed, quality of the products, and efficiency of the therapy. Here, we systematically review the overall approach for AI implementation in pharmaceutical products' lifecycle. The published studies in PubMed and IEEE Xplore were searched from inception to March 2022. The papers were screened for relevant outcomes, publication types, and data sufficiency, and a total of 73 (1.2%) out of 6131 studies were retrieved after the selection. We extracted the data according to the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) statement. All Artificial Intelligence systems could be divided into multiple overlapping categories by implementation. For the 177 projects found, the most popular areas of AI implementation are clinical trials and pre-clinical tests (34%). In second place are novel small molecule design systems, with 33% of the total. The third most popular scope for AI implementation is target identification for novel medicines. More than 25% of the systems provide this functionality. It is interesting that most of the systems specialize in only one area (102 systems—57%). None of the systems provide functionality for full coverage of the lifecycle and function in all categories of the tasks. This meta-analysis demonstrated that Artificial Intelligence solutions in pharmaceutical products' lifecycle could find numerous implementations, and none of the available market solutions covers them all.

Keywords: health; machine learning; deep learning; neural networks; pharmaceutical products' lifecycle; pharmaceutical industry

Citation: Koshechkin, K.A.; Lebedev, G.S.; Fartushnyi, E.N.; Orlov, Y.L. Holistic Approach for Artificial Intelligence Implementation in Pharmaceutical Products Lifecycle: A Meta-Analysis. *Appl. Sci.* **2022**, *12*, 8373. https://doi.org/10.3390/app12168373

Academic Editors: Fabio La Foresta and Vangelis Karalis

Received: 13 July 2022
Accepted: 16 August 2022
Published: 22 August 2022

1. Introduction

Medical bioinformatics challenges problems of pharmaceutical product optimization based on Artificial Intelligence (AI) approaches. The lifecycle of a pharmaceutical product begins with Research and Development. In the first stage, it is necessary to search for the composition of a medicinal product and determine its biological activity based on the properties of its components, as well as the technology of future production. The development phase is usually one of the most expensive and time-consuming in bringing a new drug to market. This is followed by pre-clinical studies of drugs across in vitro and in vivo models. After passing the previous stage, an application is submitted for conducting clinical trials. At this stage of the drug product life cycle, a trial design is created, and after its approval, clinical trials are carried out. This is followed by the market authorization of the medicinal product. The developer of the medicinal product sends for examination the dossier materials containing the information collected during the previous stages and pharmaceutical product samples. The organization of production is usually carried out in parallel with the previous stages. Production issues include not only the presentation of the active substance in a consumer available dosage form but also the selection of the necessary auxiliary substances. The fundamental aspect of this system is quality assurance

at all stages of production—from the purchase and quality control of the raw materials to the final product quality control.

The next stage is storage and transportation provided by the manufacturer or by the distributor. During wholesale, the possibility of contamination is excluded, and the observance of the cold chain is ensured. This process also capacitates the delivery of drugs to medical and pharmaceutical facilities at the right time. Retail sale follows after wholesale. Medical use is the main goal of drug development and the main stage of their circulation. The use of medicines is carried out as prescribed by medical workers or independently by patients (responsible self-medication with Over-the-Counter (OTC) drugs). Unfortunately, the main effect of drugs may be accompanied by side effects (any other types of action, except for the main one, including adverse reactions and post-vaccination complications) or the lack of declared effectiveness. The medical use of drugs must be simultaneously accompanied by the monitoring of their safety and effectiveness.

The use of AI can be considered for each of the listed lifecycle stages. AI in pharmaceutical companies refers to the use of automated algorithms to perform multiple tasks which traditionally rely on human intelligence. Over the last years, the use of AI in the pharmaceutical and biotechnology industry has redefined how scientists identify new drug targets, perform drug repositioning and repurposing, generate novel molecules, conduct clinical trials, and perform all other activities across the medical product lifecycle.

One of the main reasons for the interest in AI is cost reduction for drug development. A study published by the Massachusetts Institute of Technology found that only 13.8% of drugs successfully pass clinical trials [1]. A company can expect to pay USD 1.3 billion on average for any drug to complete the entire clinical trial process and obtain Food and Drug Administration (FDA) approval [2].

The aim of our work was to systematically study the general approach to the introduction of AI in the life cycle of pharmaceutical products. In the next sections, we review AI approaches in pharmaceutical product studies, methods of the studies search, results of the search (existing AI systems), and discuss the applications, such as drug repositioning.

2. AI in Pharmaceutical Product Studies

2.1. Research and Development

In drug development, AI assists in the initial screening of drug compounds with a predictable success rate based on biological factors and rapid measurement of Ribonucleic Acid (RNA) and Deoxyribonucleic Acid (DNA) by sequencing. Precision medicine or next-generation sequencing helps to find drugs and customized medicines for individual patients faster. From the development of new molecules to the identification of new biological targets, AI plays an important role in drug identification and validation, drug discovery based on targeted, phenotypic, and multipurpose drugs, drug reassignment, and identification of biomarkers [3–6]. For example, IBM's Watson reveals connections and relationships among genes, drugs, diseases, and other entities by analyzing multiple sets of life sciences knowledge.

2.2. Pre-Clinical Studies

AI is being used to reduce the likelihood associated with pre-clinical trials. Researchers are now using AI to streamline data collection and screening for pre-clinical test recipients. Data collection and analysis is an integral part of health research, and a human researcher cannot keep up with the data available. However, with AI tools such as deep learning and machine learning, it is possible to analyze, select patterns, and link relevant data that could lead to drug discovery. AI can be used to automate the pre-clinical analysis of images and samples [7,8]. For instance, Alphabet recently launched Isomorphic Labs based on AI breakthroughs at its DeepMind AI operation, Nvidia has invested in the Clara suite of AI tools and applications, and Baidu's AI drug discovery unit has struck a major deal with Sanofi.

2.3. Clinical Studies

AI technologies have reached a level of maturity that allows them to be used in real-life settings to assist decision-makers. AI has the potential to transform the key stages of clinical trial development from study preparation to execution in order to increase trial success rates, thereby reducing the workload for pharmaceutical researchers. Determination of a suitable study candidate based on history and disease conditions, as well as additional characteristics considering infection rate, demographics, and ethnicity to represent the most affected. Real World Data (RWD) and Real World Evidence (RWE) are playing an increasing role in healthcare decision-making [9]. The use of computers, mobile devices, wearables, and other biosensors to collect and store vast amounts of health-related data is rapidly accelerating. These data have the potential to enable pharmaceutical companies to better plan and conduct clinical trials and research in healthcare settings to answer questions that were not previously possible. In addition, with the development of sophisticated, new analytical capabilities, pharmaceutical companies can better analyze these data and apply the results of the analysis to the development and approval of medicines [9–12].

2.4. Market Authorization

The registration dossier preparation procedure involves the collection of the necessary administrative documents, as well as the generalization of the results of the conducted studies of the quality, safety, and efficacy of the medicinal product. In addition to the previously described capabilities, AI technology enables text mining and automates the presentation of research results. International regulators adopt rules governing AI's Involvement in analytical research results [13–15].

2.5. Manufacturing

Pharmaceutical production is another area for the introduction of innovative technologies. The Quality by Design (QbD) approach is a methodology used to ensure product quality and is characterized by a well-defined roadmap. Artificial Neural Networks (ANN) in test drug development based on QbD have been successfully applied to support the establishment of drug specifications and constraints on process parameters, linking formulation development to in vitro performance and positive clinical results obtained in bioequivalence studies [16,17].

Continuous Manufacturing (CM) of pharmaceutical drugs is another new approach in the pharmaceutical industry. The error rate allows adequate control of the process using a Deep Neural Network (DNN); it was proven that the main critical process parameters could be identified at a higher level of process understanding. The synergy between process analytics and ANN implementation creates a superior continuous production line monitoring system and enhances knowledge about this innovative production line and the products it produces [18].

2.6. Availability and Logistics

Epidemic forecasting is one of the key examples of this topic. Machine learning and AI technologies are also being applied to monitor and predict epidemic outbreaks or seasonal diseases around the world. A predictive forecast helps plan the supply chain to obtain inventory at the right time and in the right quantity based on the predicted intensity.

Periodic market shortages of essential medicines are a global public health problem. The use of multilayer neural networks for drug shortage monitoring will allow the forecasting apparatus to be transferred from the level of empirical observations to a scientific forecasting methodology based on modern digital technologies. Analysis and forecasting, as well as modeling of the modern drug supply, are in the target state of the healthcare system [19–21].

2.7. Medical Use

In addition to being implemented in diagnostics, AI can be useful in clinical practice as a tool for personalizing treatment, supporting decision-making about therapy, and managing multiple drug prescriptions. Given the importance of data-rich analysis in identifying appropriate intervention targets and treatment strategies for a person with a disease, AI can play an important role in the development of personalized medicines [22,23]. Machine learning is now being used to predict COVID-19 treatment outcomes [24,25].

2.8. Pharmacovigilance

Automating the processing of pharmaceutical safety applications represents a significant opportunity to impact the single largest cost driver for a company's overall pharmacovigilance budget. The expansion of public resources and the introduction of Electronic Health Records (EHR) have allowed the use of AI methods for pharmacovigilance. Post-marketing pharmacovigilance relies on a variety of data sources such as molecular, chemoinformational, and clinical databases, as well as social media and the biomedical literature. Natural Language Processing (NLP) techniques, including word embedding and attentional mechanisms, are the preferred methods for extracting drug-Adverse Event (AE) associations in text data [26–28]. Many mid- and large-scale pharma organizations are investing in AI-enabled automated pharmacovigilance (PV) processes, applying algorithms to data entry and triage, medical review phases, and reporting. In doing so, they aim to enhance data quality and accuracy, eliminate bias, reduce the manual burden on their teams, and maintain compliance with regulatory requirements. Providing real-time monitoring of high-quality data sets also has significant implications for patient safety and protection.

3. Methods and Materials

3.1. Search Methods for Identifying Studies

Meta-analysis guidelines were used for this systematic review [29]. The first stage of the literature search was performed with two publicly available databases, IEEE Xplore and PubMed, for relevant studies published through March 2022. No language or publication year filters were applied, while non-human experiments, case reports, guidelines, conference papers, letters, editorials, and review articles were excluded.

The search for publications was also carried out in the PubMed and PMC databases in the time interval from 1 January 2010 to 1 June 2022. The search query is represented by possible paired combinations of entities from the sets [any {Drug}] AND [any {AI}] (Figure 1):

Drug{*'discovery','research','development','manufacturing','modelling','creation'*}

AI{*'machine learning','deep learning','neural net','text mining','natural language processing','artificial intelligence'*}

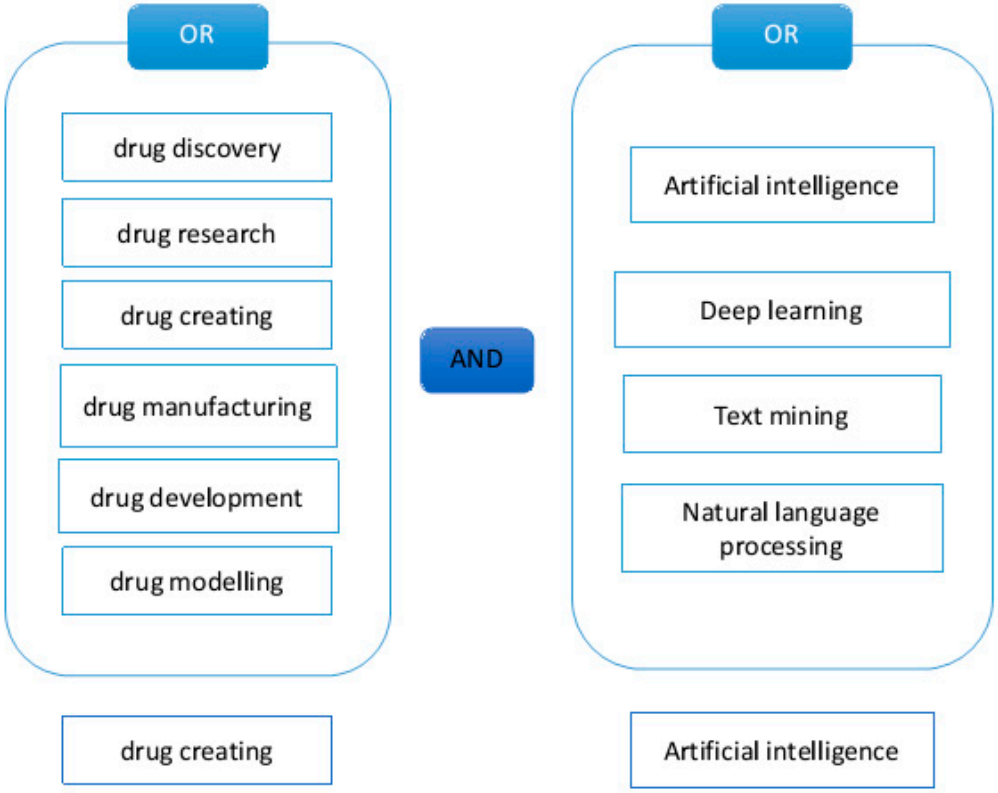

Figure 1. Graphical representation of a search query.

For the subsequent processing and classification of the selected documents according to thematic content, the Sechenov AI Information System for the Intelligent Processing of Biomedical Texts was used (Information system for intelligent processing of biomedical texts. Computer program—RU 2020666121, 4 December 2020).

To create a probabilistic thematic model, the "bag of words" hypothesis was used, where the subject and semantic content of the document is determined by the set of words included in the document and does not depend on their order or location in the document. Then, each document, with some probability, can belong to a variety of topics. A topic is a subset of tokens or n-grams, each of which is estimated by the probability of belonging to a particular topic. In a mathematical setting, any topic from a collection of selected documents is defined as a discrete probability distribution in the space of tokens or n-grams of the dictionary of the analyzed collection. In this case, each document in the collection is considered as a sample of tokens from the bag of words formed by the collection of documents. If we evaluate the probability distributions of topics on the generated collection of documents, then we can determine to which subsets of topics each document is assigned with one or another probability. Thematic modeling is based on the application of the Bayes formula, in which the distribution of words and topics is expressed as a mixture of the densities of the distributions of words and documents.

Thematic content and distribution of documents by topic form the basis of this systematic review. The thematic model forms the probabilistic distribution of documents in the collection on a set of 26 topics. Using the technology of visualization of subject models proposed [30], Figure 2 shows an example of distribution and a bag of words for documents related to topic 2, recurrent and convolutional neural networks, used in modern drug design.

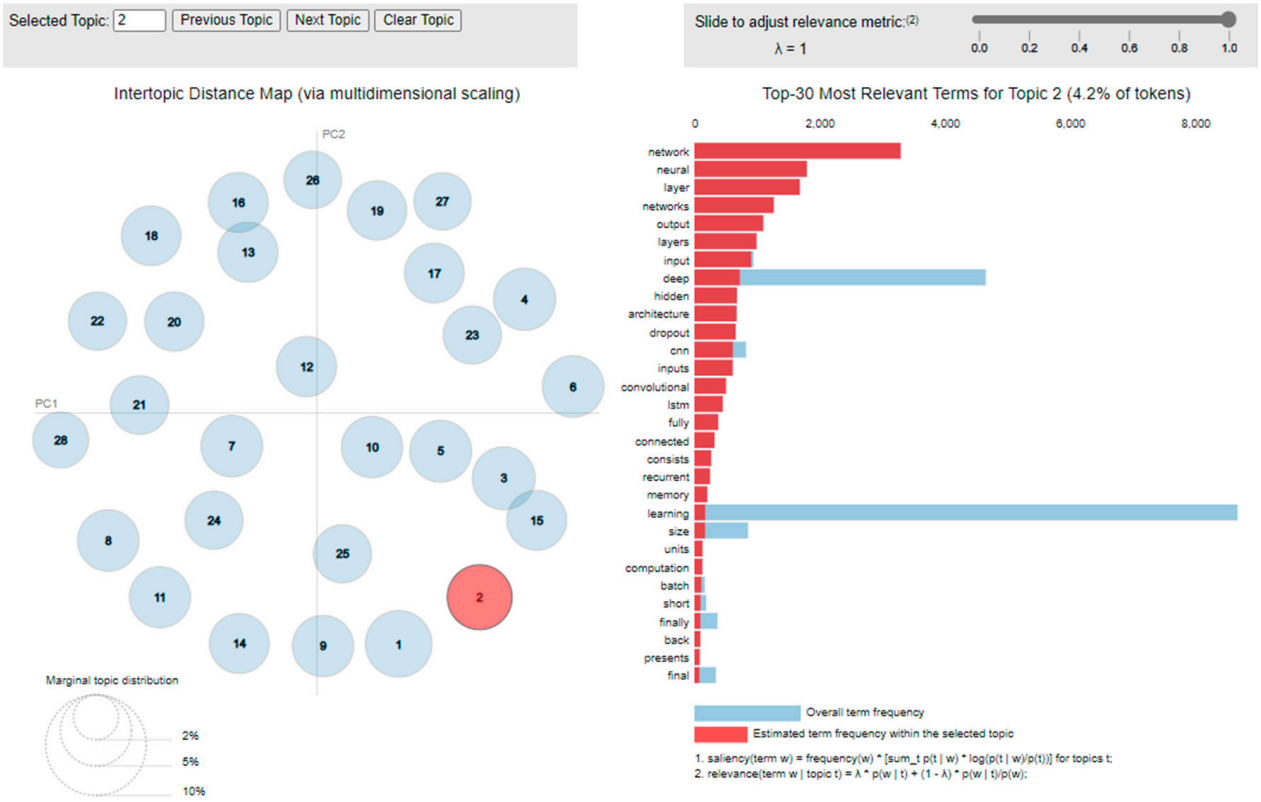

Figure 2. Visualization of a bag of words on the subject of recurrent and convolutional neural networks.

The processing of the resulting model was carried out in accordance with the recommendations for the preparation of systematic reports and meta-analysis (PRISMA, http://www.prisma-statement.org, accessed on 1 June 2022).

3.2. Eligibility Criteria for Considering Studies for This Review

We included studies that evaluated ML algorithms as a means of improving at least one of the pharmaceutical product lifecycles. We included studies that proposed technical solutions based on AI and described practical, real-life implementation experiences. We considered only publicly available solutions.

3.3. Study Selection

The study selection and data extraction were independently performed by two authors (KK and GL). After removing duplicates, titles and abstracts were screened for exclusion of studies with potentially non-relevant data.

3.4. Data Collection

The next stage was a full-text review. ML solutions were further classified into categories: Target identification, Drug Repositioning and Repurposing, Generation of Synthetic biology, Novel small molecules, Clinical trial, Natural language analysis, Image classification, Personalized therapy, Drug dispense control, Logistics and contracts, Submission reports generation, and Epidemiology.

4. Results

4.1. Search Results

The first two stages of the literature search yielded 2617 hits from PubMed and 1466 hits from IEEE Xplore. The volume of search results from Sechenov AI Information System was 6131 publications in 446 journals. After reviewing titles and abstracts, we

excluded 2837 studies due to duplication and 2745 studies due to unrelated abstracts or publication types. A total of 549 studies were evaluated in full text, of which 189 were excluded due to results of disinterest and 37 studies were excluded due to retrieval problems. A further 79 were excluded due to the lack of practical implementation, 54 were not relevant to the pharmaceutical products' lifecycle, and 117 were excluded as a result of data insufficiency. After further exclusion, only 154 studies were selected after the second round of literature selection. In the third round of literature review, 23 studies were added, resulting in 73 studies being extracted after literature selection, resulting in a total of 177 AI implementations included for the final analysis (Supplementary Table S1).

Figure 3 shows a block diagram showing the algorithm for processing the received documents at the stages of the systematic review. It shows the number of identified, included, and excluded records, as well as the reasons for exclusions.

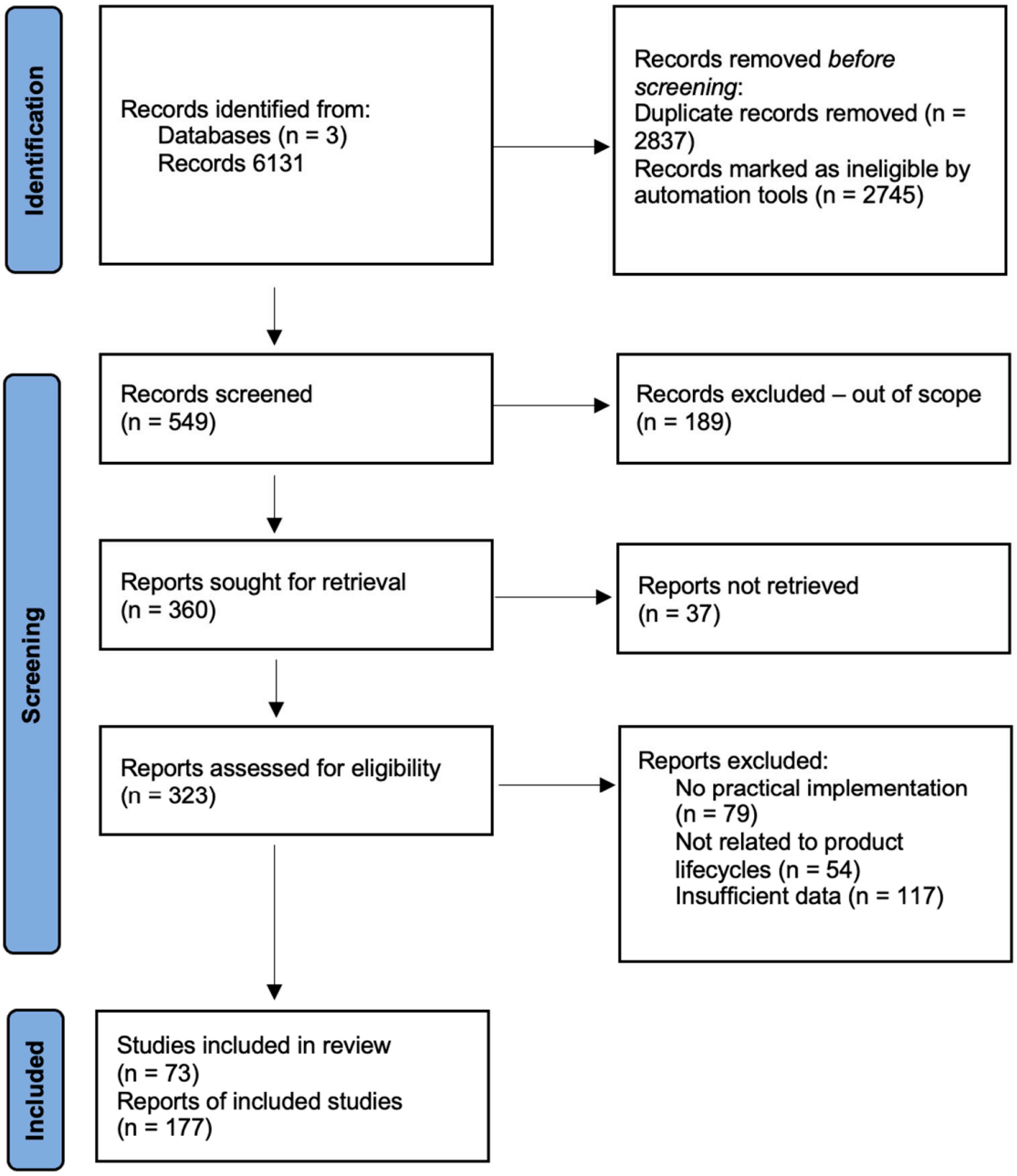

Figure 3. Identification of studies via PRISMA 2020 recommendations.

All technical solutions found are based on multidomain knowledge. For our research, we have selected those projects that have a priority on the life cycle of medicines. In general, almost any solution that uses AI in healthcare has an impact on drug therapy in one way or another since practically every medical service has a drug component. We considered those solutions which primarily modify the processes associated with drugs. Certainly, a major effort of AI implementation across pharmaceuticals is condensed inside so-called "Big Pharma" companies. In most cases, their solutions are not publicly available and were not considered in our study, or the opposite is an implementation of one or several AI projects that we discuss. In this section, we will describe several of the most influential AI systems on the market.

4.2. AI Systems

The medical company Insilico (Hong Kong, China) is focused on applying next-generation AI techniques to drug development, biomarker development, and aging research. It focuses on two machine learning methods: Generative Adversarial Networks (GAN) and Reinforcement Learning (RL) [31].

Vyasa Analytics (Newburyport, MA, USA) provides a platform for collaborative knowledge discovery and data analysis. Cortex uses AI to organize data and provide information to design teams. The catalog includes tens of millions of high-quality text sources such as PubMed, patents, and clinical trials [32].

Xbird (Berlin, Germany) uses advanced AI to understand the relationship between patient behavior, therapy, and health. This helps to increase adherence to therapy and improve treatment outcomes [33].

Gene Network Sciences (GNS) Healthcare (Somerville, MA, USA) said computer simulation systems are being used to run billions of virtual experiments (what the company calls "forward simulation") to identify biomarkers of drug efficacy and toxicity to stratify patients in clinical trials. These biomarkers can later be confirmed in a traditional wet laboratory environment [34].

Standigm (Seoul, South Korea) uses AI to identify patterns and trends related to human biology in vast biomedical databases. It uses three core patented technologies, including lead optimization for hit generation, drug repurposing, and target discovery, and then applies new design solutions to move from new disease targets to innovative molecules in record time. Applying its own innovative AI to the real world, Standigm currently has over 20 pre-clinical pipelines in various therapeutic areas [35,36].

Unlearn AI (San Francisco, CA, USA) is accelerating drug development through groundbreaking computational clinical trials. They create a digital twin that is a longitudinal, computationally generated clinical map describing what would happen if a particular patient received a placebo [37].

Deep Lens (Columbus, OH, USA) AI-based technology improves clinical outcomes in oncology. Based on feedback from experienced users around the world, they are improving the system to include AI-based image detection and workflow support, telepathology, collaboration, cloud storage, and a built-in Application Programming Interface (API) for integration by hardware and software vendors and biopharmaceutical companies [38].

Genesis Therapeutics (San Francisco, CA, USA) has created an advanced small molecule discovery platform to accelerate and optimize the development of new drugs. A combination of DNNs, biophysical modeling, and scalable computing infrastructure to achieve molecular generation and property prediction [39].

BioXcel Therapeutics (New Haven, CT, USA) is a clinical-stage biopharmaceutical company that is using AI to determine the next wave of drugs in neuroscience and immuno-oncology. Their approach leverages existing approved drugs and/or clinically evaluated product candidates together with big data and proprietary machine learning algorithms to identify new therapeutic indices. This approach has the potential to reduce the expense and time associated with drug development in diseases with substantial unmet medical needs [40].

InterVenn Biosciences (San Francisco, CA, USA) has combined the resolution of molecular-scale mass spectrometry with the throughput of AI to create a platform for discovering new biomarkers, enabling more effective treatments, developing accurate medical tests, and achieving new goals [41].

The MedAware (Ra'anana, Israel) platform detects and prevents medication-related errors and risks by applying advanced machine learning algorithms and outlier detection mechanisms similar to the fraud detection solutions used by financial institutions around the world [42].

PerceptiMed (Mountain View, CA, USA) uses AI technology to ensure the safety of drug dispensing in both the pharmacy and the hospital. Their solutions aim to eliminate medication errors while avoiding side effects. Machine learning models for drug identification will help pharmacists and medical staff focus on patient care rather than dispensing drugs [43].

The CareTuner (Herlev, Denmark) platform includes new mobile technology that allows you to digitally identify medicines and save a digital copy of all medicines you consume—with just a smartphone. Patients scan their medications using a mobile phone camera, and advanced computer vision and AI algorithms immediately identify the medications and match them to the patient's prescription. Feedback is provided to the patient, and a personal history of medication use is securely stored and can be shared with caregivers and healthcare professionals [44].

Savana (Madrid, Spain) is a machine learning-based service that turns clinical notes into structured patient information for physicians and pharmacists. It is based on NLP technology and, according to the developers, improves service quality as well as patient safety [45].

LabGenius (London, UK) is finding next-generation drugs through machine learning, synthetic biology, and lab automation, using the power of machine learning to create more advanced protein therapeutics [46].

InsightRX (San Francisco, CA, USA) is a software platform that incorporates the principles of quantitative pharmacology and machine learning to provide personalized insight into patient response to treatment. The company provides its technologies in the form of easy-to-use point-of-care clinical decision support tools that help guide treatment decisions. The platform guides treatment decisions at both the individual and population levels [47].

PathAI (Boston, MA, USA) currently uses machine learning to classify digital images of cells, each processed with different experimental compounds. Machine learning algorithms collect and group compounds that have similar effects before passing clean data to researchers, who can decide how to use this data in their work [48].

Cyclica (Toronto, ON, Canada) is a biotech company that combines biophysics and AI to find faster, safer, and cheaper drugs. They screen small molecule drugs for repositories of structurally characterized proteins to determine polypharmacological profiles. From here, the company identifies important protein targets and then uses AI to determine the effect of the drug on those targets. Finally, AI produces a visual result of the interaction of the drug and proteins. By understanding how small molecule drugs interact with all proteins in the body, Cyclica can develop the best solution, understand potential side effects, and identify new applications for existing drugs [49].

Healx (Cambridge, UK) is a promising start-up focused on accelerating the treatment of rare diseases, and AI is at the forefront of its activities. Its AI platform, HealNet, allows scientists to increase the production of cures for diseases while reducing time, cost, and risk. The company is not directly focused on developing new drugs to treat these conditions. Instead, they use AI technology to study existing drugs and repurpose them to treat rare diseases [50].

IBM Watson (New York, NY, USA) is a leading institution at the forefront of changing treatment decisions, using medical information and patient history to optimize treatment options. Pharmaceutical companies will use Watson to speed up the process of analyzing

and testing hypotheses that are based on "big data from disparate sources". The latter includes more than 30 million laboratory and statistical reports, as well as medical literature. Watson will be able to combine this vast information array with the results of its own research [51].

5. Discussion

5.1. Subgroup Analyses for Artificial Intelligence Implementation

All AI systems for aiding pharmaceutical products' lifecycle could be divided into multiple overlapping categories (Supplementary Table S1). For the total of 177 projects found, the most popular areas of AI implementation are clinical trials and pre-clinical tests (34%). Some of the solutions emulate the assessment process on the neural network, which works as a digital twin of the patient or test system. Other solutions help to manage received data or arrange patients applicable for the study. In second place are the novel small molecules design systems with 33% of the total. These systems try to create new substances from a feature of the target or similar drugs. The third most popular scope for AI implementation is target identification for novel medicines. More than 25% of the systems provide this functionality. It is interesting that most of the systems specialize in only one area (102 systems—57% are single-purpose). In general, they are new products of start-up companies. None of the systems provide functionality for full coverage of the lifecycle and function in all categories of the tasks. A detailed description of these categories and examples of the solutions are discussed below. The categorization of the solutions is presented in Figure 4.

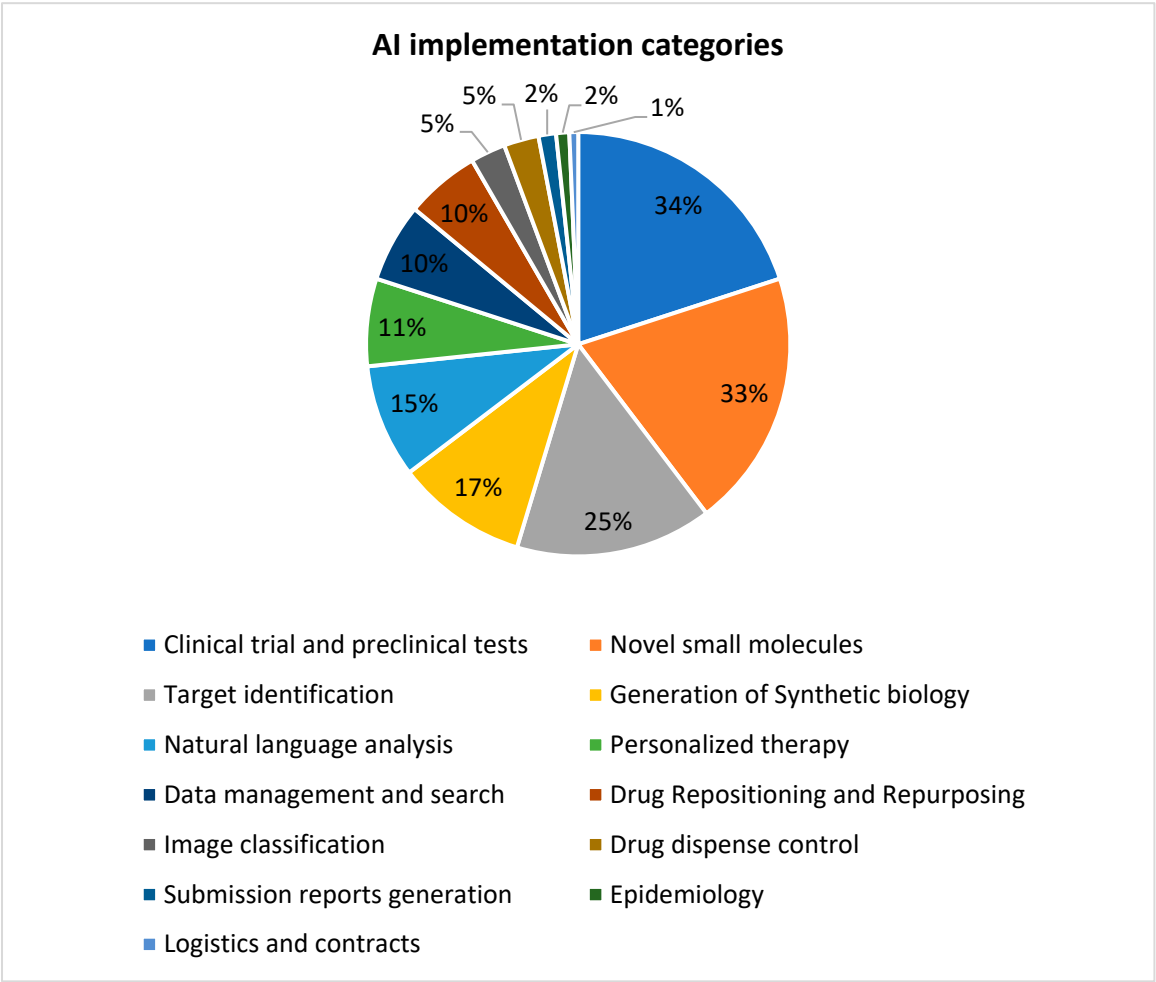

Figure 4. AI implementation categories.

5.2. Natural Language Analysis Subgroup

Natural language analysis can be divided into categories: quantitative methods, based on the simpler principle of a set of words, and more complex methods of understanding natural language. "Bag of words" is a simplified model used in NLP and information retrieval. In this model, the text is presented as an unordered set of words without regard to grammar or even word order. The simplifying assumption that these variables are independent is considered. Since these methods are mainly based on word and/or feature counts, these problems with medical annotation text become extremely relevant since all different representations of a feature must be counted as the same feature [52]. NLP allows you to parse text and extract metadata from content such as concepts, entities, keywords, categories, moods, emotions, relationships, semantic roles, and syntax. This will require the creation of a pipeline for text cleaning, automatic detection and correction of errors in fuzzy words, work with abbreviations, normalization of units of measure, normalization of date and time formats, text blurring, and text extraction. Several tools are used: Tokenizers [53], Segmenters [54], Tree Analyzer [55], and Dependency Analyzers [56]. In *Understanding Text from the Ground Up*, Xiang Zhang and Yann LeKun demonstrate that Convolutional Neural Network (CNN) can achieve outstanding performance without knowledge of words, phrases, sentences, or any other syntactic or semantic structures in relation to human language [57]. Peilu Wang and colleagues in *Part-of-Speech Tagging with Bidirectional Long Short-Term Memory Recurrent Neural Network* presented a model for POS (Part-of-Speech) tagging [58].

In [59], the effectiveness of the use of NLP methods for extracting knowledge about new chemical compounds, which are crucial for the drug discovery process, is shown. Despite the complexity of the task due to the large volume of existing patents and the complex linguistic properties of patents for chemicals, the use of NLP methods demonstrates high efficiency.

In [60], Bouhedjar and colleagues found an approach to NLP based on the embedding of DNNs. This method aims to convert the format of the Simplified Molecular Input System (SMILES) string into word embedding vectors to represent compound semantics. These vectors are fed into supervised machine learning algorithms such as LSTM, Support Vector Machine (SVM), and Random Forest (RF) to build quantitative models of structure and activity relationships on toxicity datasets.

5.3. Submission Reports Generation Implementation Subgroup

In the article *Natural language generation, paraphrasing and summarizing user feedback using recurrent neural networks*, the authors demonstrate a Recurrent Neural Network (RNN) model that can generate new sentences and summaries of documents [61]. Transducers are designed to process serial data such as natural language for tasks such as translating and summarizing text. Because the Transformer model allows for more parallelism during training, it has allowed training on larger datasets than was possible prior to its introduction. This has led to the development of pre-trained systems such as Bidirectional Encoder Representations from Transformers (BERT) and Generative Pre-trained Transformer (GPT), which have been trained using huge common language datasets such as the Wikipedia corpus and can be fine-tuned to specific language tasks [62,63]. Many pre-trained models such as GPT-2, GPT-3, BERT, XLNet, and RoBERTa demonstrate the ability of Transformers to perform a wide range of such NLP-related tasks and can find real-world applications for Document Generalization, Document Generation, and Named Entity Recognition (NER) [64,65].

5.4. Target Identification Implementation Subgroup

Drug targets are molecular structures whose abnormal activity associated with a disease can be altered by drugs, improving the health of patients. ML algorithms that have been widely used in drug discovery include RF, Naive Bayes (NB), and Support Vector

Machine (SVM), among other methods. Extracting knowledge from this large amount of unstructured information is one of the relevant methods [66].

5.5. Novel Molecules Implementation Subgroup

First, AI generated millions of potential new molecules that meet certain specifications. A machine learning platform was then used to predict which compounds would be active on hundreds to thousands of proteins. A third layer of algorithms, called active learning, was then applied, which automatically prioritized which decisions researchers needed to make and test. These processes allowed the scientists to accurately design the drug while simultaneously completing many AI-assisted design tasks.

Principal Component Analysis (PCA) can be used to reduce the size of the extracted features. Considering these properties, the molecular representations used in the AI drug development algorithm include the molecular fingerprint, the SMILES string, potential energy measurements (e.g., from ab initio calculations), molecular plots with different weights for atoms or bonds, Coulomb matrices, molecular fragments or bonds, 3D coordinates of atoms, the density of electrons around a molecule, or combinations thereof. These inputs are used in the training phase of the DNN and can be processed by different DNNs in different phases, namely the generation and prediction phase. This procedure can facilitate RL. In a typical study, the DNN generative stage takes SMILES input and trains to produce chemically acceptable SMILES strings, while the predictive stage learns about molecular properties. Although the two stages are initially trained separately using supervised learning algorithms, bias can be applied to the outcome when the two stages are trained jointly by rewarding or punishing certain features [4,67,68].

5.6. Drug Repositioning and Repurposing

The COVID-19 pandemic has caused the evaluation of new chemicals and biologicals as potential treatments for SARS-CoV-2 infection. Repositioning and using existing drugs approved for unrelated conditions is being widely promoted as a therapeutic approach for COVID-19 because it provides faster, more effective responses to the virus. One can use dictionary-based text mining in combination with specialized AI or machine learning, originally called statistical pattern recognition, such as Bidirectional Biomedical Language Representation Model (BioBERT) [69,70].

A fully connected Feed-Forward Neural Network (FNN) is an architecture in which artificial neurons are connected layer by layer from input features to output targets. A weight is associated with each connection and is optimized by minimizing the loss in predicting output targets through backpropagation across training samples. FNN can be used to classify drugs into pharmaceutical therapeutic classes based on drug transcriptome profile vectors [71]. AI classification can quickly discover drugs that can fight both emerging diseases (such as COVID-19) and existing diseases [72,73].

5.7. Generation of Synthetic Biology Implementation Subgroup

As an example for this AI implementation subgroup, we consider a model based on a CNN and a Multilayer Perceptron (MLP) for RNA analysis. It considers fingerprint sequences as one-dimensional images or lines of nucleotide bases and identifies base patterns and potential interactions between these bases to predict RNA function (so-called "toehold switch sequences") [74]. Another model is based on NLP and considers each RNA-based sequence as a "phrase" made up of patterns of "words", ultimately learning how certain words are put together to make a coherent phrase [74,75].

5.8. Clinical Trials

AI algorithms can speed up the development of clinical trials by automatically identifying suitable subjects, ensuring the correct distribution of study participants, and providing an early warning system for clinical trials that do not produce meaningful results [70]. AI algorithms can be used to check data validity (invalid input), completeness (structure

and amount of missing data), corruption (representativeness of data), and timeliness (data reflecting current practice). Another way to use AI in clinical trials is based on RWD. These projects can include patients in a study, profile them (for example, using DNA sequencing, proteomics, metabolomics, etc.), and then use RWD to match the drugs in the study with the pathologies identified in the profiling. Patient profile matching strategies in these studies may be based on AI and OD analysis of large relevant datasets. AI and ML can be additionally used to support an electronic version of study data monitoring, thereby ensuring the correctness of the data and patient safety, thus reducing the need for costly site monitoring of the study. In addition, EHR data can be combined with other types of RWD, such as genomics, and patient-reported issues can be obtained with AI and ML techniques to create a more complete picture for drug discovery and biomarkers [76]. OpenAI has created a powerful processing model, GPT-3, which has the potential to generate human-like text on demand. It is proposed to use GPT-3 for clinical operations, for example, to generate an FDA-regulated electronic Case Report Form (eCRF) from a protocol [77]. The GPT-3 language model API allows users in clinical trials to create multiple applications that run on the same underlying technology, such as creating a study or making changes to a protocol much faster.

5.9. Image Classification Subgroup

Recently, the most used image classification algorithm is CNN. CNNs are adapted versions of neural networks that combine multilayer neural networks with specialized layers capable of extracting the most important and important features for the classification of an object. CNNs can automatically detect, generate, and study image features. This significantly reduces the need to manually label and segment images to prepare them for machine learning algorithms. They also have an advantage over MLP networks because they can work with non-convex loss functions [78,79].

5.10. Personalized Therapy and Drug Dispense Control

Personalized medicine or more effective treatment based on individual health data combined with predictive analytics is also a hot area of research and is closely related to better disease assessment. Currently, supervised learning rules in this area allow doctors, for example, to choose from a more limited set of diagnoses or assess a patient's risk based on symptoms and genetic information. For example, collecting information about the DNA sequence obtained from a person to make a genetic diagnosis of a disease can greatly help AI-based analysis [80]. The development of multifunctional machine learning platforms for extracting, aggregating, managing, and analyzing clinical data can help clinicians by effectively separating subjects to understand specific scenarios and optimize the decision-making process. The introduction of AI into healthcare is a compelling vision that can lead to significant improvements to achieve the goals of providing more personalized and population-based medicine in real-time at a lower cost [81].

5.11. Logistics and Contracts

In many cases, medical professionals face problems related to drug orders. These could be connected with illegible handwriting or wrong drug selection. Some drug names could be confused due to similarity. Dosages could be incorrectly calculated. It could be human errors or problems with the computerized system in use. AI and ML could be used to process big data from EHR and verify contracts, orders, and drug dispensing and administration [82–84].

5.12. Epidemiology Implementation Subgroup

ML and AI technologies are also applied to monitor and forecast epidemics around the world, based on data collected from satellites, historical information on the web, and real-time updates on social networks. SVM and ANN have been used, for example, to

predict malaria epidemics, taking into account data such as temperature, average monthly precipitation, and total number of positive cases [85,86].

As a result of processing the corpus of documents corresponding to this topic in [87], about 20 machine learning technologies used in modern pharmaco-epidemiology were identified: deep learning of neural networks of various topologies (mainly CNN), RF algorithms, K-clustering and K-random neighbors, classification, regression and decision trees, Bayesian methods, SVM, discriminant analysis, hierarchical clustering, etc.

Machine learning technologies and methods were used to solve problems predicting clinical reactions and responses to the applied pharmacotherapy, determining the required dosage for pharmacotherapy for a particular patient, predicting drug consumption, predicting the degree of adverse reactions to drug use, predicting the length of stay of patients receiving clinical therapy, etc.

It was shown that in more than 50% of comparisons, AI showed better results than traditional pharmacoepidemiological methods. However, there is a high heterogeneity of efficiency among various AI methods. Many methods have hardly been investigated. Ultimately, these results suggest that further research is needed, focusing on head-to-head comparisons of traditional pharmacoepidemiological methods with machine learning methods in different research scenarios and with different data sources.

5.13. Computational Drug Discovery

Computer simulations and machine learning for drug discovery are rapidly gaining popularity. The search for new drugs based on machine learning leads to a significant reduction in combinations and a reduction in the diversity of drug candidate compounds, which is significantly different from the traditionally used methods. The development of drugs and the discovery of new compounds in drugs is carried out by the structural synthesis of new compounds and modeling, the inapplicable condition of which is the difference between the drugs being created and those already released into circulation or at the stage of clinical trials. Classical drug design is based on time-consuming manual procedures in which the researcher performs a series of pharmacokinetic and/or molecular dynamics experiments. The use of machine learning allows you to set the required molecular properties, find, extract, and combine new molecules with fragmentary computer modeling. For this, various types of molecular properties are used—these are the physicochemical properties of the compounds and the three-dimensional structures of targets with information about the points or regions of binding. Such computer-generated drug candidates, generated using computational algorithms, previously unknown chemical compounds are subsequently synthesized by traditional chemical technologies and provided as new molecules in bioassays to test predicted interactions with appropriate targets.

When developing drugs using ML methods, the pipeline of operations performed is divided into five stages:

- Initial generation of molecules with the desired structural properties using conventional labeled interaction data similar to classical drug design;
- Identification and clustering of different types of created molecules, which have identified properties, are generated using calculations with a coefficient of randomization to increase the diversity. This work is done by generative models;
- Construction of vectors of features for molecules, generated by calculations, and their transfer to the interaction prediction model as input data;
- Computer modeling and obtaining output parameters for each molecule generated by calculations, which allows us to estimate the probability and efficiency of interaction with the target;
- The problem is solved iteratively until the search for an extremum—the maximum estimate of the probability.

This technology is used to build both generative and test models. An example of automated implementation of the described technology is presented in [88].

5.14. Principal Results and Comparison with Prior Work

This systematic review synthesizes the available evidence for AI technology implementation in pharmaceutical products' lifecycle. Start-up activity in this area is booming and getting close to practical results. It is highly probable that AI technology will become an indivisible component of the pharmaceutical industry. Prior works were focused on definitive parts of the product lifecycle, for example, drug development [84,85] or clinical trials [10,14]. In our study, we aimed to show a holistic approach for AI implementation in pharmaceutical products' lifecycle, which can be achieved by integration of different AI solutions into practice, as well as scope extension for ongoing start-ups to cover more activities.

5.15. The Potential Future Works in the Research Area

Modern trends in the development of science and technology are characterized by the widespread introduction of intelligent systems (technical systems that implement cognitive functions) into a wide variety of spheres of human activity. This approach demonstrates an increase in the efficiency of processes in high-tech sectors of the economy. This is especially clearly demonstrated in such areas as public safety and medicine, voice interactive services (computer hearing), natural text recognition, banking systems (credit scoring, biometric authentication systems), and others.

The pharmaceutical industry is unable to be on the first lines of AI implementation due to strict regulations and safety reasons demanding more time. However, overall AI potential archived in other areas will be transferred to the pharmaceutical industry. Performing a similar review in several years will probably show us additional insights. Another avenue for future work is to make a deeper analysis of each separate area of AI implementation.

6. Conclusions

AI solutions in pharmaceutical products' lifecycle could find numerous implementations, and none of the available market solutions cover them all. NLP allows document summarization, document generation, and NER based on novel BERT and GPT. It can be used for RWE-based trials, reports, and summaries generation. RF, NB, and SVM, as well as other methods, could be used for a large amount of unstructured information analysis for new drug target identification.

DNNs, RL, and PCA are most useful for novel molecule generation in silico and their activity prediction. Drug repositioning and repurposing could be done with text mining, coupled with FNN. Generation of synthetic biology is based mostly on NLP implementations for RNA-based sequencing. Clinical trials utilize RWD and RWE approaches with AI, NLP, and NER support. Image classification with CNNs can automatically discover, generate, and learn features of images which is useful in pre-clinical and clinical trial results processing. Personalized therapy could be aided with NN patient risk prediction and multiple factors analysis, including genetics. Drug dispensing control is based on EMR analysis for counterindications and drug combination interactions. Additionally, ML and AI technologies could be used for monitoring and predicting epidemic outbreaks around the world to align pharmaceutical development.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app12168373/s1, Table S1: AI companies.

Author Contributions: Conceptualization, K.A.K. and E.N.F.; Data curation, K.A.K., E.N.F. and G.S.L.; Formal analysis, Investigation, K.A.K., E.N.F. and G.S.L.; Funding acquisition, K.A.K., G.S.L. and Y.L.O.; Methodology, G.S.L.; Project administration, E.N.F. and G.S.L.; Resources, E.N.F. and G.S.L.; Supervision, G.S.L.; Visualization, K.A.K. and Y.L.O.; Writing—original draft, K.A.K.; Writing—review and editing, Y.L.O. and K.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: The publication was prepared with the support of the RUDN University Scientific Projects Grant System, project R.3-2022-ins (Y.L.O.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Abbreviations

AE	Adverse Event
AI	Artificial Intelligence
BERT	Bidirectional Encoder Representations from Transformers)
CM	Continuous Manufacturing
CNN	Convolutional Neural Network
DNN	Deep Neural Network
eCRF	electronic Case Report Form
EHR	Electronic Health Records
FNN	Feed-forward Neural Network
GPT	Generative Pre-trained Transformer
LSTM	Long Short-Term Memory
MLP	Multilayer Perceptron
NB	Naive Bayes
NER	Named Entity Recognition
NLP	Natural Language Processing
PCA	Principal Component Analysis
POS	Part-of-Speech
QbD	Quality-by-Design
RF	Random Forest
RL	Reinforcement Learning
RNN	Recurrent Neural Network
RWD	Real World Data
RWE	Real World Evidence
SMILES	Simplified Molecular Input Line-Entry System
SVM	Support Vector Machine

References

1. Wong, C.H.; Siah, K.W.; Lo, A.W. Estimation of clinical trial success rates and related parameters. *Biostatistics* **2019**, *20*, 273–286. [CrossRef] [PubMed]
2. Smith, B.D.; Tarricone, R.; Vella, V. The role of product life cycle in medical technology innovation. *J. Med. Mark. Device Diagn. Pharm. Mark.* **2013**, *13*, 37–43. [CrossRef]
3. Réda, C.; Kaufmann, E.; Delahaye-Duriez, A. Machine learning applications in drug development. *Comput. Struct. Biotechnol. J.* **2019**, *18*, 241–252. [CrossRef] [PubMed]
4. Chan, H.C.S.; Shan, H.; Dahoun, T.; Vogel, H.; Yuan, S. Advancing Drug Discovery via Artificial Intelligence. *Trends Pharmacol. Sci.* **2019**, *40*, 592–604. [CrossRef] [PubMed]
5. Liu, B.; He, H.; Luo, H.; Zhang, T.; Jiang, J. Artificial intelligence and big data facilitated targeted drug discovery. *Stroke Vasc. Neurol.* **2019**, *4*, 206–213. [CrossRef] [PubMed]
6. Bohr, H. Drug discovery and molecular modeling using artificial intelligence. *Artif. Intell. Healthc.* **2020**, 61–83. [CrossRef]
7. Bali, J.; Garg, R.; Bali, R. Artificial intelligence (AI) in healthcare and biomedical research: Why a strong computational/AI bioethics framework is required? *Indian J. Ophthalmol.* **2019**, *67*, 3–6. [CrossRef]
8. Tang, T.; Zawaski, J.A.; Francis, K.N.; Qutub, A.A.; Gaber, M.W. Image-based Classification of Tumor Type and Growth Rate using Machine Learning: A preclinical study. *Sci. Rep.* **2019**, *9*, 1–10. [CrossRef]
9. Sherman, R.E.; Anderson, S.A.; Dal Pan, G.J.; Gray, G.W.; Gross, T.; Hunter, N.L.; LaVange, L.; Marinac-Dabic, D.; Marks, P.W.; Robb, M.A.; et al. Real-World Evidence—What Is It and What Can It Tell Us? *N. Engl. J. Med.* **2016**, *375*, 2293–2297. [CrossRef]
10. Harrer, S.; Shah, P.; Antony, B.; Hu, J. Artificial Intelligence for Clinical Trial Design. *Trends Pharmacol. Sci.* **2019**, *40*, 577–591. [CrossRef]

11. Chen, Z.; Liu, X.; Hogan, W.; Shenkman, E.; Bian, J. Applications of artificial intelligence in drug development using real-world data. *Drug Discov. Today* **2020**, *26*, 1256–1264. [CrossRef] [PubMed]
12. Crown, W.H. Real-World Evidence, Causal Inference, and Machine Learning. *Value Health* **2019**, *22*, 587–592. [CrossRef] [PubMed]
13. Medicines Agency. HMA-EMA Joint Big Data Taskforce Phase II Report: "Evolving Data-Driven Regulation" 1. Available online: https://www.ema.europa.eu/en/documents/other/hma-ema-joint-big-data-taskforce-phase-ii-report-evolving-data-driven-regulation_en.pdf (accessed on 1 August 2022).
14. Rivera, S.C.; Liu, X.; Chan, A.-W.; Denniston, A.K.; Calvert, M.J.; Darzi, A.; Holmes, C.; Yau, C.; Moher, D.; Ashrafian, H.; et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension. *Nat. Med.* **2020**, *26*, 1351–1363. [CrossRef] [PubMed]
15. Liu, X.; Rivera, S.C.; Moher, D.; Calvert, M.J.; Denniston, A.K.; Chan, A.-W.; Darzi, A.; Holmes, C.; Yau, C.; Ashrafian, H.; et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *Nat. Med.* **2020**, *26*, 1364–1374. [CrossRef] [PubMed]
16. Simões, M.F.; Silva, G.; Pinto, A.C.; Fonseca, M.; Silva, N.E.; Pinto, R.M.; Simões, S. Artificial neural networks applied to quality-by-design: From formulation development to clinical outcome. *Eur. J. Pharm. Biopharm.* **2020**, *152*, 282–295. [CrossRef]
17. Gams, M.; Horvat, M.; Ožek, M.; Luštrek, M.; Gradišek, A. Integrating Artificial and Human Intelligence into Tablet Production Process. *AAPS PharmSciTech* **2014**, *15*, 1447–1453. [CrossRef]
18. Zhang, W.-W.; Li, L.; Li, D.; Liu, J.; Li, X.; Li, W.; Xu, X.; Zhang, M.J.; Chandler, L.A.; Lin, H.; et al. The First Approved Gene Therapy Product for Cancer Ad-*p53*(Gendicine): 12 Years in the Clinic. *Hum. Gene Ther.* **2018**, *29*, 160–179. [CrossRef]
19. Obermeyer, Z.; Emanuel, J. Predicting the Future—Big Data, Machine Learning, and Clinical Medicine. *N. Engl. J. Med.* **2016**, *375*, 1216–1219. [CrossRef]
20. Arsalan, T.; Koshechkin, K.; Lebedev, G. Scientific Approaches to the Digitalization of Drugs Assortment Monitoring Using Artificial Neural Networks. *Smart Innov. Syst. Technol.* **2020**, 391–401. [CrossRef]
21. Report | Drug Shortages: Root Causes and Potential Solutions | FDA [Electronic Resource]. Available online: https://www.fda.gov/drugs/drug-shortages/report-drug-shortages-root-causes-and-potential-solutions (accessed on 1 August 2022).
22. Schork, N.J. Artificial Intelligence and Personalized Medicine. In *Precision Medicine in Cancer Therapy*; Springer: Cham, Switerlands, 2019; Volume 178, pp. 265–283. [CrossRef]
23. Lebedev, G.; Fartushnyi, E.; Fartushnyi, I.; Shaderkin, I.; Klimenko, H.; Kozhin, P.; Koshechkin, K.; Ryabkov, I.; Tarasov, V.; Morozov, E.; et al. Technology of Supporting Medical Decision-Making Using Evidence-Based Medicine and Artificial Intelligence. *Procedia Comput. Sci.* **2020**, *176*, 1703–1712. [CrossRef]
24. Ferrari, D.; Milic, J.; Tonelli, R.; Ghinelli, F.; Meschiari, M.; Volpi, S.; Matteo, F.; Giacomo, F.; Vittorio, I.; Dina, Y.; et al. Machine learning in predicting respiratory failure in patients with COVID-19 pneumonia—Challenges, strengths, and opportunities in a global health emergency. *PLoS ONE* **2020**, *15*, e0239172. [CrossRef] [PubMed]
25. Kim, H.-J.; Han, D.; Kim, J.-H.; Kim, D.; Ha, B.; Seog, W.; Lee, Y.-K.; Lim, D.; Hong, S.O.; Park, M.-J.; et al. An Easy-to-Use Machine Learning Model to Predict the Prognosis of Patients with COVID-19: Retrospective Cohort Study. *J. Med. Internet Res.* **2020**, *22*, e24225. [CrossRef] [PubMed]
26. Basile, A.O.; Yahi, A.; Tatonetti, N.P. Artificial Intelligence for Drug Toxicity and Safety. *Trends Pharmacol. Sci.* **2019**, *40*, 624–635. [CrossRef]
27. Schmider, J.; Kumar, K.; Laforest, C.; Swankoski, B.; Naim, K.; Caubel, P.M. Innovation in Pharmacovigilance: Use of Artificial Intelligence in Adverse Event Case Processing. *Clin. Pharmacol. Ther.* **2018**, *105*, 954–961. [CrossRef] [PubMed]
28. Mockute, R.; Desai, S.; Perera, S.; Assuncao, B.; Danysz, K.; Tetarenko, N.; Gaddam, D.; Abatemarco, D.; Widdowson, M.; Beauchamp, S.; et al. Artificial Intelligence within Pharmacovigilance: A Means to Identify Cognitive Services and the Framework for Their Validation. *Pharm. Med.* **2019**, *33*, 109–120. [CrossRef] [PubMed]
29. Shamseer, L.; Moher, D.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A.; PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ* **2015**, *350*, g7647. [CrossRef]
30. Sievert, C.; Shirley, K. *LDAvis: A Method for Visualizing and Interpreting Topics*; Association for Computational Linguistics: Stroudsburg, PA, USA, 2014; pp. 63–70. [CrossRef]
31. Aliper, A.; Plis, S.; Artemov, A.; Ulloa, A.; Mamoshina, P.; Zhavoronkov, A. Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. *Mol. Pharm.* **2016**, *13*, 2524–2530. [CrossRef]
32. Vyasa Analytics Introduces Synapse, A Novel "Smart Table" Software Application Powered by Deep Learning AI Technologies. Available online: https://www.prweb.com/releases/vyasa_analytics_introduces_synapse_a_novel_smart_table_software_application_powered_by_deep_learning_ai_technologies/prweb16677978.htm (accessed on 1 August 2022).
33. AI-Powered Healthcare. Available online: https://xbird.io/ (accessed on 1 August 2022).
34. Xconomy: Gene Network Sciences Using Supercomputing to Match Patients with a Drug That Works—Page 2 of 2. Available online: https://xconomy.com/boston/2010/04/27/gene-network-sciences-using-supercomputing-to-match-patients-with-a-drug-that-works/2/ (accessed on 1 August 2022).
35. Using AI to Accelerate Drug Discovery. Available online: https://www.nature.com/articles/d42473-020-00354-y (accessed on 1 August 2022).

36. Standigm and SK Chemicals Repurpose FDA-approved Drug into Rheumatoid Arthritis Candidate and Apply for Patent through Their Open Innovation Partnership, Business Wire. Available online: https://www.businesswire.com/news/home/20210107005121/en/Standigm-and-SK-Chemicals-Repurpose-FDA-approved-Drug-into-Rheumatoid-Arthritis-Candidate-and-Apply-for-Patent-through-their-Open-Innovation-Partnership (accessed on 1 August 2022).
37. Unlearn—Home. Available online: https://www.unlearn.ai (accessed on 1 August 2022).
38. Deep Lens BioPharma and CROs. Available online: https://www.deeplens.ai/deep-lens-healthcare-biopharma-cro (accessed on 1 August 2022).
39. Feinberg, N.; Sur, D.; Wu, Z.; Husic, B.E.; Mai, H.; Li, Y.; Saisai, S.; Jianyi, Y.; Bharath, R.; Pande, V.S. PotentialNet for molecular property prediction. *ACS Cent. Sci.* **2018**, *4*, 1520–1530. [CrossRef]
40. Product Pipeline: BioXcel Therapeutics, Inc. (BTAI). Available online: https://www.bioxceltherapeutics.com/our-pipeline/ (accessed on 1 August 2022).
41. InterVenn Biosciences | Clinical Glycoproteomics + Artificial Intelligence. Available online: https://intervenn.com (accessed on 1 August 2022).
42. Transforming Patient Safety—Medaware. Available online: https://www.medaware.com (accessed on 1 August 2022).
43. Grønning, N. Data Management in a Regulatory Context. *Front. Med.* **2017**, *4*, 114. [CrossRef]
44. Turn Patient Insights into Actionirody. Available online: https://irody.com (accessed on 1 August 2022).
45. Savana—Transform the Free Text of Your Clinical Records into Big Data. Available online: https://www.savanamed.com (accessed on 1 August 2022).
46. LabGenius. Available online: https://labgeni.us (accessed on 1 August 2022).
47. InsightRX—Precision Dosing Done Right. Available online: https://www.insight-rx.com (accessed on 1 August 2022).
48. Novartis CEO Who Wanted to Bring Tech into Pharma Now Explains Why It's So Hard. Available online: https://www.forbes.com/sites/davidshaywitz/2019/01/16/novartis-ceo-who-wanted-to-bring-tech-into-pharma-now-explains-why-its-so-hard/ (accessed on 1 August 2022).
49. Drug Discovery with an AI-Augmented Platform—Cyclica. Available online: https://www.cyclicarx.com (accessed on 1 August 2022).
50. Artificial Intelligence Startup Healx Gets $10 Million to Find Cures for Rare Diseases—Healthcare Weekly. Available online: https://healthcareweekly.com/healx-artificial-intelligence-pharma (accessed on 1 August 2022).
51. IBM Watson Health | AI Healthcare Solutions | IBM. Available online: https://www.ibm.com/watson-health (accessed on 1 August 2022).
52. Carvalho, J.P.; Curto, S. Fuzzy Preprocessing of Medical Text Annotations of Intensive Care Units Patients. Available online: https://www.inesc-id.pt/publications/10251/pdf (accessed on 1 August 2022).
53. LingPipe Home. Available online: http://www.alias-i.com/lingpipe/ (accessed on 1 August 2022).
54. The Stanford Natural Language Processing Group. Available online: https://nlp.stanford.edu/software/lex-parser.shtml (accessed on 1 August 2022).
55. Meng, H.; Luk, P.-C.; Xu, K.; Weng, F. GLR parsing with multiple grammars for natural language queries. *ACM Trans. Asian Lang. Inf. Process.* **2002**, *1*, 123–144. [CrossRef]
56. De Marneffe, M.-C.; Maccartney, B.; Manning, C.D. Generating Typed Dependency Parses from Phrase Structure Parses. 2006. Proc. of LREC. 6. Available online: https://www.researchgate.net/publication/200044364_Generating_Typed_Dependency_Parses_from_Phrase_Structure_Parses (accessed on 1 August 2022).
57. Zhang, X.; Lecun, Y. Text Understanding from Scratch. 2016. Available online: https://arxiv.org/abs/1502.01710 (accessed on 1 August 2022).
58. Wang, P.; Qian, Y.; Soong, F.K.; He, L.; Zhao, H. Part-of-Speech Tagging with Bidirectional Long Short-Term Memory Recurrent Neural Network. 2015. Available online: https://arxiv.org/abs/1510.06168 (accessed on 1 August 2022).
59. He, J.; Nguyen, D.Q.; Akhondi, S.A.; Druckenbrodt, C.; Thorne, C.; Hoessel, R.; Afzal, Z.; Zhai, Z.; Fang, B.; Yoshikawa, H.; et al. ChEMU 2020: Natural Language Processing Methods Are Effective for Information Extraction from Chemical Patents. *Front. Res. Metrics Anal.* **2021**, *6*, 654438. [CrossRef] [PubMed]
60. Bouhedjar, K.; Boukelia, A.; Nacereddine, A.K.; Boucheham, A.; Belaidi, A.; Djerourou, A. A natural language processing approach based on embedding deep learning from heterogeneous compounds for quantitative structure–activity relationship modeling. *Chem. Biol. Drug Des.* **2020**, *96*, 961–972. [CrossRef] [PubMed]
61. Tarasov, S.D. Natural Language Generation, Paraphrasing and Summarization of User Reviews with Recurrent Neural Networks. Available online: http://www.dialog-21.ru/digests/dialog2015/materials/pdf/TarasovDS2.pdf (accessed on 1 August 2022).
62. Google AI Blog: Open Sourcing BERT: State-of-the-Art Pre-training for Natural Language Processing. Available online: https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html (accessed on 1 August 2022).
63. Better Language Models and Their Implications. Available online: https://openai.com/blog/better-language-models/ (accessed on 1 August 2022).
64. Nambiar, A.; Heflin, M.; Liu, S.; Maslov, S.; Hopkins, M.; Ritz, A. Transforming the Language of Life: Transformer Neural Networks for Protein Prediction Tasks. In Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, BCB 2020, New York, NY, USA, 21–24 September 2020; Volume 5, pp. 1–8. [CrossRef]

65. bioRxiv Preprints Can Now be Submitted Directly to Leading Research Journals. Available online: https://phys.org/news/2016-01-biorxiv-preprints-submitted-journals.html (accessed on 1 August 2022).
66. Danger, R.; Segura-Bedmar, I.; Martínez, P.; Rosso, P. A comparison of machine learning techniques for detection of drug target articles. *J. Biomed. Inform.* **2010**, *43*, 902–913. [CrossRef] [PubMed]
67. Bohr, A.; Memarzadeh, K. The rise of artificial intelligence in healthcare applications. In *Artificial Intelligence in Healthcare*; Academic Press: Cambridge, MA, USA, 2020; pp. 25–60. [CrossRef]
68. Piroozmand, F.; Mohammadipanah, F.; Sajedi, H. Spectrum of deep learning algorithms in drug discovery. *Chem. Biol. Drug Des.* **2020**, *96*, 886–901. [CrossRef]
69. Lee, J.; Yoon, W.; Kim, S.; Kim, D.; Kim, S.; So, C.H.; Kang, J. BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* **2020**, *36*, 1234–1240. [CrossRef]
70. Levin, J.M.; Oprea, T.I.; Davidovich, S.; Clozel, T.; Overington, J.P.; Vanhaelen, Q.; Cantor, C.R.; Bischof, E.; Zhavoronkov, A. Artificial intelligence, drug repurposing and peer review. *Nat. Biotechnol.* **2020**, *38*, 1127–1131. [CrossRef]
71. Zhou, Y.; Wang, F.; Tang, J.; Nussinov, R.; Cheng, F. Artificial intelligence in COVID-19 drug repurposing. *Lancet Digit. Health* **2020**, *2*, e667–e676. [CrossRef]
72. AI Driven Repositioning and Repurposing Summit 2021. Available online: https://pharmaphorum.com/events/ai-driven-repositioning-and-repurposing-summit-2021/ (accessed on 1 August 2022).
73. Mohanty, S.; Rashid, M.H.A.; Mridul, M.; Mohanty, C.; Swayamsiddha, S. Application of Artificial Intelligence in COVID-19 drug repurposing. *Diabetes Metab. Syndr. Clin. Res. Rev.* **2020**, *14*, 1027–1031. [CrossRef]
74. Deep Learning Takes on synthetic Biology. Available online: https://wyss.harvard.edu/news/deep-learning-takes-on-synthetic-biology/ (accessed on 1 August 2022).
75. Radivojević, T.; Costello, Z.; Workman, K.; Martin, H.G. A machine learning Automated Recommendation Tool for synthetic biology. *Nat. Commun.* **2020**, *11*, 1–14. [CrossRef]
76. Shah, P.; Kendall, F.; Khozin, S.; Goosen, R.; Hu, J.; Laramie, J.; Ringel, M.; Schork, N. Artificial intelligence and machine learning in clinical development: A translational perspective. *npj Digit. Med.* **2019**, *2*, 1–5. [CrossRef]
77. Overview | The Conference Forum. Available online: https://theconferenceforum.org/conferences/webinar-open-ai-gpt3/overview/ (accessed on 1 August 2022).
78. State-Of-The-Art Image Classification Algorithm: FixEfficientNet-L2 | by Hucker Marius | Towards Data Science. Available online: https://towardsdatascience.com/state-of-the-art-image-classification-algorithm-fixefficientnet-l2-98b93deeb04c (accessed on 1 August 2022).
79. Cellular Image Classification for Drug Discovery | by Huseyn Gasimov | Intelec AI | Medium. Available online: https://medium.com/intelec-ai/cellular-image-classification-for-drug-discovery-4ef55741151c (accessed on 1 August 2022).
80. Lebedev, G.; Fartushniy, E.; Shaderkin, I.; Klimenko, H.; Kozhin, P.; Koshechkin, K.; Ryabkov, I.; Tarasov, V.; Morozov, E.; Fomina, I.; et al. Creation of a Medical Decision Support System Using Evidence-Based Medicine. *Int. Conf. Intell. Decis. Technol.* **2020**, 413–427. [CrossRef]
81. Ahmed, Z.; Mohamed, K.; Zeeshan, S.; Dong, X. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database* **2020**, *2020*, baaa010. [CrossRef] [PubMed]
82. Artificial Intelligence to Eradicate Prescription Errors. Available online: https://www.whatnextglobal.com/post/artificial-intelligence-to-eradicate-prescription-errors (accessed on 1 August 2022).
83. Harika Anupama, A.G.; Srilekha, G.; Uma Priya, G. Review on Artificial Intelligence (AI) in Drug Dispensing and Drug Ac-countability. Available online: https://globalresearchonline.net/journalcontents/v61-2/07.pdf (accessed on 1 August 2022).
84. Real-World Applications of Artificial Intelligence to Improve Medication Management Across the Care Continuum—Electronic Health Reporter. Available online: https://electronichealthreporter.com/real-world-applications-of-artificial-intelligence-to-improve-medication-management-across-the-care-continuum/ (accessed on 1 August 2022).
85. Artificial Intelligence and the Opioid Epidemic—Applications for Relapse Treatment, Abuse Prevention, and More | Emerj. Available online: https://emerj.com/ai-sector-overviews/artificial-intelligence-opioid-epidemic/ (accessed on 1 August 2022).
86. Malaria Outbreak Prediction Model Using Machine Learning. Available online: https://pdf4pro.com/amp/view/malaria-outbreak-prediction-model-using-machine-learning-474cb5.html (accessed on 1 August 2022).
87. Sessa, M.; Khan, A.R.; Liang, D.; Andersen, M.; Kulahci, M. Artificial Intelligence in Pharmacoepidemiology: A Systematic Review. Part 1—Overview of Knowledge Discovery Techniques in Artificial Intelligence. *Front. Pharmacol.* **2020**, *11*, 1028. [CrossRef] [PubMed]
88. Baskin, I.; Winkler, D.; Tetko, I.V. A renaissance of neural networks in drug discovery. *Expert Opin. Drug Discov.* **2016**, *11*, 785–795. [CrossRef]