Trends in Pharmacological Sciences

CellPress REVIEWS

Special Issue: Rise of Machines in Medicine

Review

Artificial Intelligence for Clinical Trial Design

Stefan Harrer,[1,*] Pratik Shah,[2] Bhavna Antony,[1] and Jianying Hu[3]

Clinical trials consume the latter half of the 10 to 15 year, 1.5–2.0 billion USD, development cycle for bringing a single new drug to market. Hence, a failed trial sinks not only the investment into the trial itself but also the preclinical development costs, rendering the loss per failed clinical trial at 800 million to 1.4 billion USD. Suboptimal patient cohort selection and recruiting techniques, paired with the inability to monitor patients effectively during trials, are two of the main causes for high trial failure rates: only one of 10 compounds entering a clinical trial reaches the market. We explain how recent advances in artificial intelligence (AI) can be used to reshape key steps of clinical trial design towards increasing trial success rates.

Highlights

Suboptimal patient selection and recruiting techniques, paired with the inability to monitor and coach patients effectively during clinical trials, are two of the main causes for high trial failure rates.

High failure rates of clinical trials contribute substantially to the inefficiency of the drug development cycle, in other words the trend that fewer new drugs reach the market despite increasing pharma R&D investment. This trend has been observed for decades and is ongoing.

AI techniques have advanced to a level of maturity that allows them to be employed under real-life conditions to assist human decision-makers.

AI has the potential to transform key steps of clinical trial design from study preparation to execution towards improving trial success rates, thus lowering the pharma R&D burden.

Artificial Intelligence Can Turn Eroom's Law into Moore's Law

It takes on average 10–15 years and USD 1.5–2.0 billion to bring a new drug to market. Approximately half of this time and investment is consumed during the clinical trial phases of the drug development cycle. The remaining 50% of R&D expenditure covers preclinical compound discovery and testing, as well as regulatory processes (Figure 1). Although pharma and biotechnology companies have continuously increased R&D investment for decades, the number of new drugs gaining regulatory approval per billion USD spent has halved approximately every 9 years [1]. Reversing **Moore's law** (see Glossary) from the world of semiconductor technology, this trend has been termed Eroom's Law. It is ongoing[i] [2] and poses a severe threat to the existing clinical development business model: in the post-**blockbuster drugs** era a lack of go-to-market efficiency of that magnitude is not sustainable. One of the main stumbling blocks in the drug development pipeline is the high failure rate of clinical trials. Less than one third of all Phase II compounds advance to Phase III [3]. More than one third of all Phase III compounds fail to advance to approval [4]. Because these crucial checkpoints do not occur until far into the second half of the R&D cycle – with the most complex Phase III trials carrying ~60% of the overall trial costs (Figure 1) – the resulting loss per failed clinical trial lies in the order of 0.8–1.4 billion USD[ii], thus constituting a significant write-off of the total R&D investment.

Two of the key factors causing a clinical trial to be unsuccessful are patient cohort selection and recruiting mechanisms which fail to bring the best suited patients to a trial in time, as well as a lack of technical infrastructure to cope with the complexity of running a trial – especially in its later phases – in the absence of reliable and efficient adherence control, patient monitoring, and **clinical endpoint** detection systems. AI (Box 1) can help to overcome these shortcomings of current clinical trial design. Machine learning (ML), and deep learning (DL) in particular (Box 2), are able to automatically find patterns of meaning in large datasets such as text, speech, or images. Natural language processing (NLP) can understand and correlate content in written or spoken language, and human–machine interfaces (HMIs) (Box 2) allow natural exchange of information between computers and humans. These capabilities can be used for correlating large and diverse datasets such as electronic health records (EHRs), medical literature, and trial databases for improved patient–trial matching and recruitment before a trial starts, as well as for monitoring patients automatically and continuously during the trial, thereby allowing improved adherence control and yielding more reliable and efficient endpoint assessment. In the following

[1]IBM Research, IBM Research Australia Lab, 3006 Melbourne, VIC, Australia
[2]Massachusetts Institute of Technology, Media Lab, 02139 Cambridge, MA, USA
[3]IBM Research, IBM T.J. Watson Research Center, 10598 Yorktown Heights, NY, USA

*Correspondence: sharrer@au.ibm.com (S. Harrer).

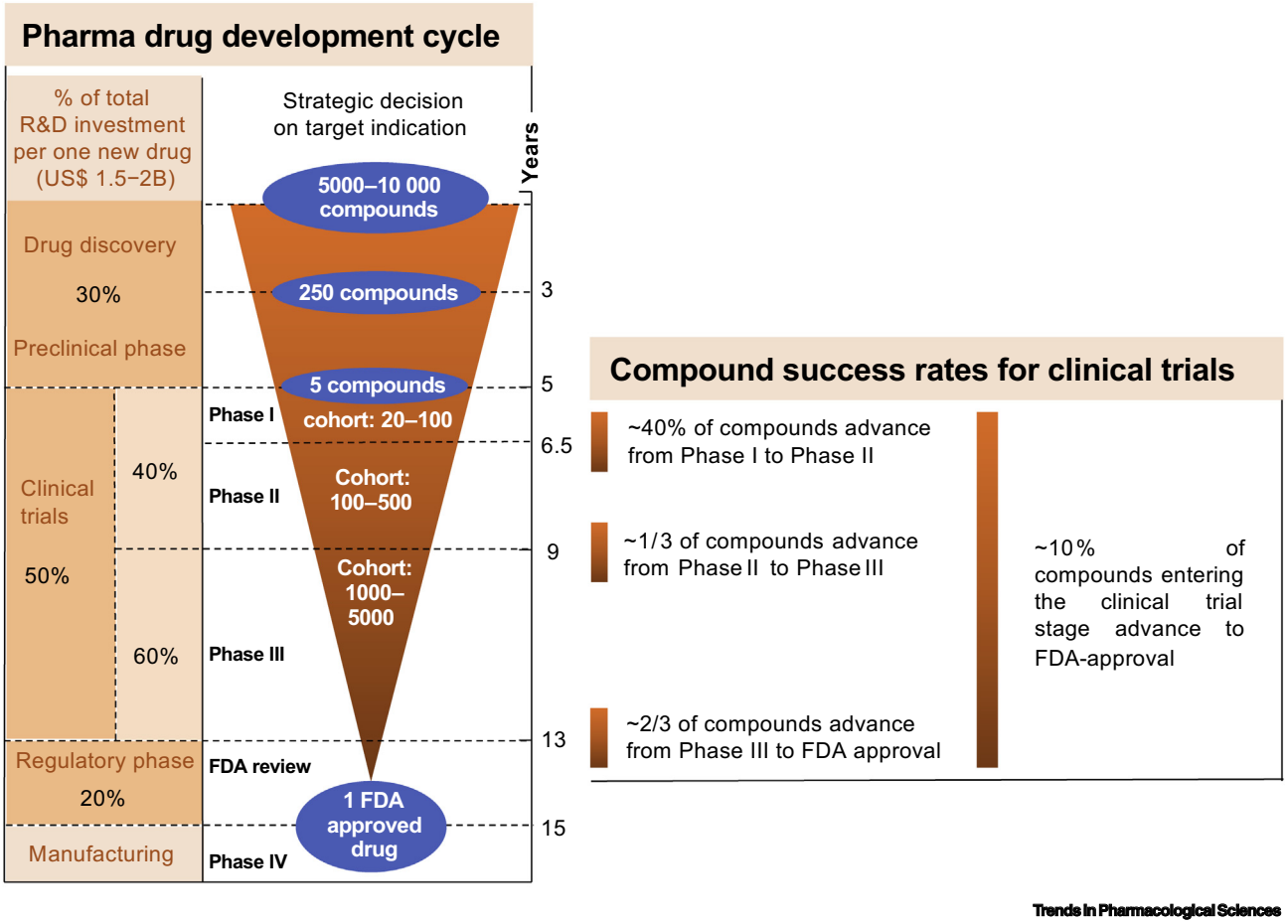


Figure 1. The Pharma Drug Development Cycle. It takes up to 15 years and an average total R&D expenditure of 1.5–2 billion (B) USD to bring a single new drug to market. About half of this investment is spent on clinical trials, with Phase III trials being the most complex and most expensive. Probabilities of success for compounds to proceed through the clinical trial stages vary from phase to phase, and lead to a situation where only one of 10 compounds entering clinical trials advances to FDA approval. High clinical trial failure rates are one major cause for the prevailing inefficiency of the drug development cycle.

sections we highlight aspects of clinical trial design with immediate potential entry points for AI, and explain specific AI techniques of interest and how their application will improve trial performance (Figure 2, Key Figure).

Patient Selection

Every clinical trial poses individual requirements on participating patients with regards to eligibility, suitability, motivation, and empowerment to enrol. The medical history of a specific patient might render them ineligible. An eligible patient might not be at the stage of the disease, or belong to a specific sub-phenotype, that is targeted by the drug to be tested, thus making that patient unsuitable. Eligible and suitable patients might not be properly incentivized to participate, and, even if they are, they might not be aware of a matching trial or find the recruitment process too complex and cumbersome to navigate. Moving enough patients through these bottlenecks under tight recruitment timelines constitutes a major challenge and is in fact the number one cause for trial delays: 86% of all trials do not meet enrolment timelines, and close to one third of all Phase III trials fail owing to enrolment problems[iii]. Patient recruitment takes up one third of the overall trial duration[iv]. For example, Phase III trials carry 60% of the total costs for moving a drug through all trial phases

Box 1. The Evolution of AI

The use of AI in medicine dates back to the early 1970s when expert systems such as MYCIN were first introduced to provide diagnostic decision support [48]. However, early medical AI systems relied heavily on medical domain experts to train computers by encoding clinical knowledge as logic rules for specific clinical scenarios. Such systems suffered from the limitation that they were labor-intensive and time-consuming to construct, and once built they were rigid and difficult to update [49]. More advanced ML systems that are capable of training themselves to learn these rules by identifying and weighing relevant features from data such as unstructured text, medical images, and EHRs emerged in the 90s and 2000s, but were relatively slow to be adopted by the medical field, largely because of the lack of widely available data and the fact that the early methods required intense feature-engineering efforts involving serious commitments from medical domain experts [50].

This situation has changed dramatically recently because of two factors. First, the field of AI itself went through transformational advances, particularly in DL and related ML methods, enabled by hardware improvements and very large training datasets [21,51]. Second, medical data became increasingly available in digital form thanks to new technology advances as well as public policy efforts such as the Electronic Records Meaningful Use Programs in the USA[xxxix]. Recent years have witnessed a surge in efforts as well as early proof-of-concept successes of AI in medicine, starting from medical imaging for detecting diabetic retinopathy [52] and skin cancer [53], to the use of EHR data to predict important clinical parameters ranging from disease onset to mortality [54]. The field of biomedicine has also benefited from this surge in AI methods at many levels, from sophisticated natural language processing (NLP) searches of the biomedical literature [55], to cancer sub-phenotyping using DL [56], to predictions of gene targets of microRNAs [57], drug–target interactions, and drug repositioning hypotheses [58].

because they require the largest patient cohorts. A 32% failure rate because of patient recruitment problems[v] in Phase III trials illustrates one of the most severe shortcomings of state-of-the-art clinical trial design: those trials with the highest patient demand suffer most from inefficient

Box 2. Different Methods Used in AI

Artificial Intelligence (AI): machine simulation of human intelligence processes including learning, reasoning, and self-correction[xl]. The ultimate goal of AI is to build machines that can perceive the world and make decisions in the same way as humans do.

Association rule mining: ML algorithms for discovering interesting relations between variables in large databases to help a machine to mimic the extraction and abstract association capabilities of the human brain from new uncategorized data.

Brain–machine interface (BMI): a direct communication pathway between an enhanced or wired brain and an external device. Also referred to as a brain–computer interface (BCI), a mind–machine interface (MMI), or a direct neural interface (DNI).

Deep learning (DL): a class of ML methods based on artificial neural networks, inspired by information processing and distributed communication nodes in biological systems, that use multiple layers to progressively extract higher level features from raw input[xli]. The 'deep' in 'deep learning' refers to the number of layers through which the data is transformed.

Deep reinforcement learning (DRL): reinforcement learning (RL) is an area of ML that is concerned with building software agents that can take actions in an environment so as to maximize some notion of cumulative reward. DRL combines DL and RL principles to create efficient algorithms to achieve this task.

Human–machine interface (HMI): a direct communication pathway between a human and a device. For example, an artificial system capable of automatically understanding and responding to spoken or written human language constitutes a human–machine Interface.

Machine learning (ML): the scientific study of algorithms that build a mathematical model of sample data to make predictions or decisions without being explicitly programmed to perform the task[xlii]. ML is often considered to be a branch of AI.

Natural language processing (NLP): a subfield of AI concerned with the interactions between computers and human (natural) languages, in particular how to program computers to process and analyze large amounts of natural language data. NLP draws from many disciplines including computer science and computational linguistics.

Optical character recognition (OCR): a field of research in AI, pattern recognition, and computational vision aimed at the electronic conversion of images of typed, handwritten, or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo, or from subtitle text superimposed on an image.

Glossary

Absence seizure: an epileptic seizure episode that the affected patient is not aware of.
Blockbuster drug: a drug that creates in excess of $1B in annual sales.
Blockchain technology: a blockchain is a digital list – often referred to as digital ledger – of permanent data records which have been uploaded sequentially in form of 'blocks'. Each time a new block is added to this 'chain' a cryptographic signature is created which connects the newly uploaded block with the latest existing block in the ledger. This process ensures that every change to the blockchain is self-validated and that a record of the change is permanently stored in the blockchain, without a need for validation through a third party. The blockchain itself is decentralized and self-governing, allowing anyone with valid access rights to use it.
Clinical trial endpoint: a clinical trial attempts to assess the potential impact of a medical intervention on the occurrence of a disease, as for example assessed by a specific symptom. The appearance of such a symptom in a patient during the course of the trial marks the clinical endpoint for that patient.
Explainability of AI: the ability to explain the inner workings of AI algorithms and the outputs they produce.
Exposome: the impact on an organism of all environmental factors to which it has been exposed to during its lifetime.
Internet of things (IoT): a state in which real-world devices are interconnected such that information and data can flow between them.
Moore's law: in 1965, Gordon Moore postulated that the power of computing would increase while its relative cost would decrease at an exponential pace. This trend held for decades and became known as 'Moore's Law'.
Overfitting of ML models: overfitting describes the state of a machine learning (ML) model that has overly precisely learned the distinct features of the training dataset such that it can no longer generalize well to datasets to which it has never been exposed to before.
Reasoning: a reasoning system automatically draws conclusions from data and knowledge.
Time-series data: data that capture the value of a single or multiple

patient recruitment techniques. AI- and ML-driven systems can help to improve patient cohort composition and provide assistance with patient recruitment (Figure 2).

Cohort Composition

Clinical trials are usually not designed to demonstrate the effectiveness of a treatment in a random sample of the general population, but instead aim to prospectively select a subset of the population in which the effect of the drug, if there is one, can more readily be demonstrated, a strategy referred to as 'clinical trial enrichment' [vi]. If a patient is *a priori* not part of the suitable subset, then their participation in the trial will automatically decrease the observed efficacy of the drug being tested. Suitability may not be confused with the degree of treatment success or absence thereof during the trial: it denotes a condition that does not render it outright impossible or highly unlikely for participating patients to respond to the tested drug. Recruiting a high number of suitable patients does not guarantee success of a trial, but enrolling unsuitable patients increases the likelihood of its failure[vii].

In an ideal world the assessment of suitability would use patient-specific diagnostic genome-to-**exposome** profiling [5] to determine whether biomarkers which the drug targets are sufficiently strongly represented in the patient profile or not. Although trials which could benefit from such an approach form a relatively small subset of all trials, they also tend to be the most expensive trials – especially when medical imaging techniques are used. Hence, although in practice there may not be a comprehensive 'omic profile', and effective biomarkers may need to be identified for the majority of therapies under clinical development, biomarker testing should still be considered whenever applicable. Sophisticated analytics methods are necessary to combine omic data with electronic medical record (EMR) and other patient data, scattered among different locations, owners, and formats – from handwritten paper copies to digital medical imagery – to surface biomarkers that lead to endpoints that can be more efficiently measured, and thereby identify and characterize appropriate patient subpopulations. This presents a unique opportunity for NLP and computer vision algorithms such as optical character recognition (OCR) (Box 2) to automate the reading and compiling of this evidence. Moreover, treating data from different sources and formats as a single coherent dataset for the purpose of its comprehensive analysis is especially challenging in the case of EMR data owing to their volume, velocity, veracity, and variety. The data source-agnostic nature of AI models makes them a unique tool for EMR data harmonization which is key to designing tools for clinical trial enrichment and biomarker discovery. However, care must be taken to reduce **overfitting of ML models** as a result of class-imbalance in the training data.

Preclinical compound discovery, compound-target testing, and defining lead compounds for clinical trials can be assisted by using generative and prediction-based AI, ML, and **reasoning** techniques [6–8]. For example, a broader and more efficient search for correlations between indications and biomarkers than conventional discovery techniques has been reported [8]. This may allow lead candidates to be chosen that have a higher chance of success during clinical trials, and the elimination of those with a higher likelihood of failing before they enter the clinical phase.

AI models and methods can also be used to enhance patient cohort selection through one or more of the following means identified by the Food and Drug administration (FDA): (i) by reducing population heterogeneity, (ii) by choosing patients who are more likely to have a measurable clinical endpoint, also called 'prognostic enrichment', and (iii) by identifying a population more capable of responding to a treatment, also termed 'predictive enrichment' [vi] (Figure 2). Electronic phenotyping is a well-established discipline within health informatics that focuses on reducing population heterogeneity, namely the process of identifying patients with specific characteristics

measured parameters over time at defined points in time.
Wearable: a device that can execute a measurement or data-processing task, and that is fully functional while attached to the human body directly or indirectly through clothing, and that has no hardwired connection to any other non-wearable device.

of interest. The characteristics can be as simple as patients with type 2 diabetes, or as complex as patients with stage II prostate cancer and urinary urgency without evidence of urinary tract infection. The task of electronic phenotyping is far more challenging than a simple code search, and requires sophisticated methods to account for heterogeneity among patient records, across multiple data types, and to leverage complex representations of clinical domain knowledge. Although early methods relying on hand-crafted rules were effective for simple cases, they proved to be insufficient for more complex and more nuanced cases. In recent years there have been increasing efforts to design a diverse range of ML methods, ranging from NLP to association rule mining to DL (Box 2), that have shown great progress towards being able to handle complex real-world situations [9].

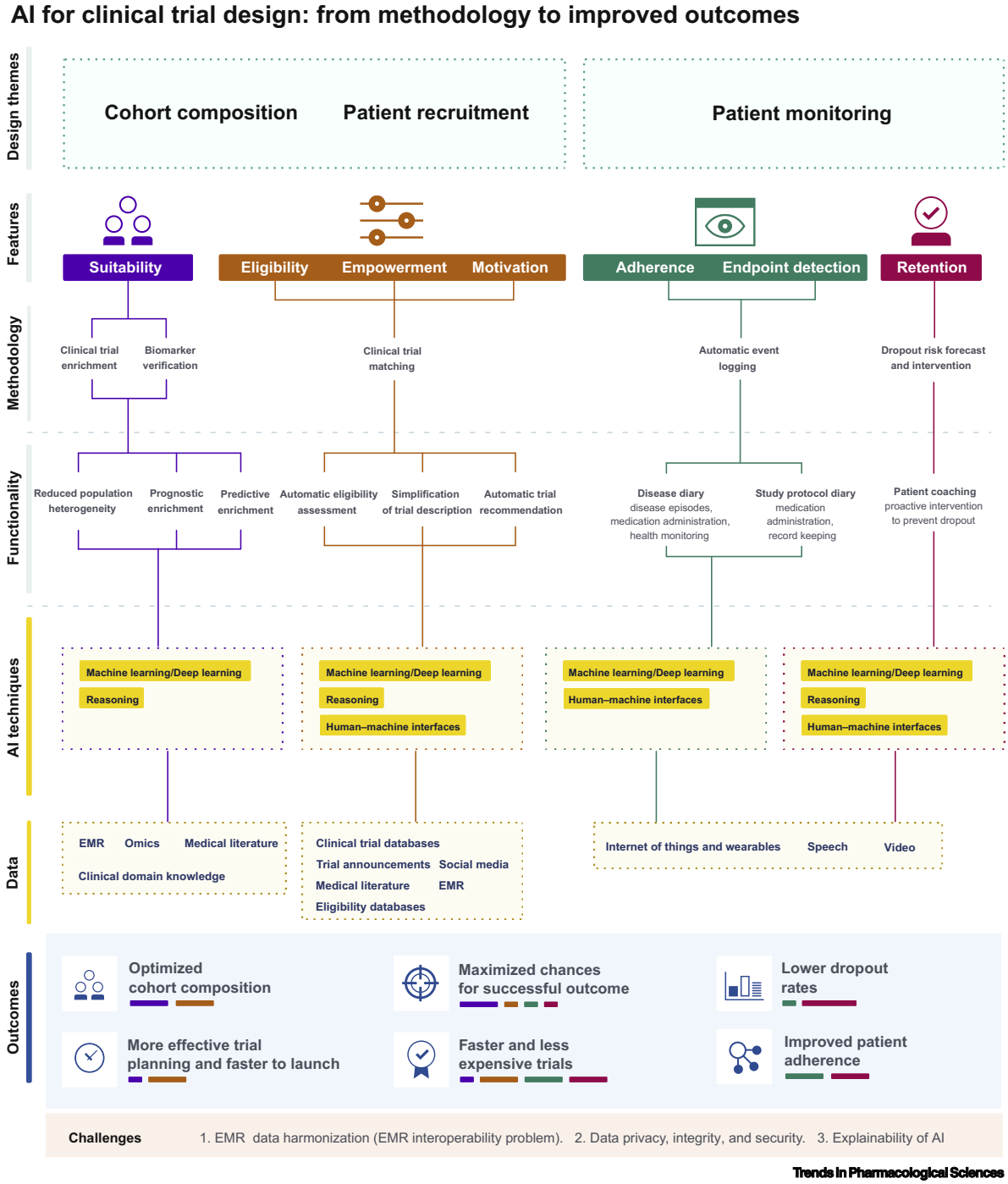
Although electronic phenotyping can be leveraged to reduce patient population heterogeneity, it is not designed to achieve prognostic or predictive enrichment. ML methods are increasingly being deployed for prognostic enrichment for neurological diseases where key biomarkers, which are typically expensive or invasive to measure, are approximated by non-linear combinations of multiple cheap and non-invasive measures which provide similar prognostic information [10,11]. Predictive enrichment requires more complex models that are necessary to characterize and assess disease progression. The Coalition Against Major Diseases (CAMD) recently led a process that successfully advanced a clinical trial simulation (CTS) tool for Alzheimer's disease (AD) through the formal regulatory review process at the FDA and the European Medicines Agency. The CTS tool includes computational components for modeling drug, disease, and progression of mild cognitive impairment (MCI) and early AD that can be used for model-based clinical trial design [12]. Expanding on this effort, ML methods for disease progression modeling are being developed to provide increasingly accurate and nuanced understanding and characterization of complexity and heterogeneity of many diseases, particularly those such as AD where disease-modifying drugs are not yet available [13–17].

Assistance in Recruitment

The complexity of trial eligibility criteria in terms of number and medical jargon generally makes it challenging for a patient to comprehend and assess their own eligibility. Manually extracting meaningful information from this large and unstructured data-source is a significant task that imposes a heavy processing burden on doctors and patients alike. Nonetheless, it is this step that largely defines whether a patient is deemed suitable and eligible to participate in a study, and also whether the recruiting site and the patient become aware of each other. Several AI techniques can offer viable assistance with automatically finding the needles in the EMR haystack: NLP [18] can be used to comprehend written and spoken language from a variety of structured and unstructured data types. A detailed summary of NLP techniques applicable to clinical trial design is provided in a recent review by Fogel [19]. Reasoning [20] techniques allow content to be digested into actionable recommendations for the human decision-maker. ML [21] and in particular deep reinforcement learning (Box 2) empowers systems to learn and integrate feedback on the quality of their analytic output into adapted underlying algorithms. Assistive systems using these AI techniques or subsets thereof can be used to automatically analyze EMR and clinical trial eligibility databases, find matches between specific patients and recruiting trials, and recommend these matches to doctors and patients (Figure 2). Such AI-based clinical trial matching systems have successfully been demonstrated and have proved their value in real-life use cases [22][viii]. Because of the AI nature of these systems, any added future functionality and improved performance predominantly will depend on the quality and amount of data which are accessible for analytical model development and pilot-study field validation work.

Key Figure

Artificial Intelligence (AI) for Clinical Trial Design



(See figure legend at the bottom of the next page.)

AI and ML techniques such as NLP and OCR have also been proposed to proactively mine publicly available web content such as, for example, digital trial databases, trial announcements, and social media to automatically identify potential matches between trials of relevance and specific patients. By assisting patients in their conventional manual web search, such a system could make patients aware of trials of interest much faster and allow them to proactively engage with clinicians for further assessment of eligibility and suitability[ix]. Indeed, the first enrolment plans employing a social media component have successfully been demonstrated[x]. We expect the integration of AI will improve the reach, efficiency, and thus the impact of such digital enrolment plans substantially in the future.

Challenges

The digitalization and accessibility of EMR data that are used extensively by AI methods are not trivial. Both tasks are challenging for contrary reasons: on the one hand a lack of regulatory frameworks on data collection causes EMR formats to differ widely, to be incompatible with each other or not digital at all, and to reside in a decentralized ecosystem without established data exchange or access gateways. On the other hand, a strongly regulated legal environment strictly limits third-party access to patient data and even makes it difficult for patients themselves to access their own data. This so-called 'EMR interoperability dilemma' is being recognized as major hurdle to making healthcare systems more efficient, and substantial investments are being made by governments and medical institutions towards overcoming this hurdle [23]. In parallel, legal frameworks such as, for example, the US Health Insurance Portability and Accountability Act (HIPAA) and the EU General Data Protection Regulation (GDPR) continue to evolve as governing and protecting sensitive health data becomes an increasingly complex endeavor in the growing network of devices, data owners, and service providers [24,25]. Further, exactly as with EMR mining, for clinical trial matching the legal aspects of data privacy and security as well as a sufficient degree of **explainability of AI** models need to be addressed to ensure that AI-based systems are operable and gain regulatory approval.

Patient Monitoring

Recruiting the right patients into a clinical trial is a massive investment of both time and funding. The return on this investment can only be realized through successful completion of the trial. Hence, it is imperative that patients stay in the trial, adhere to trial procedures and rules throughout the trial, and that all data-points for monitoring the impact of the tested drug are collected efficiently and reliably. Only 15% of clinical trials do not experience patient dropout, and the average dropout rate across clinical trials is 30%[iii]. Dropouts caused by a lack of adherence to trial protocols require additional recruiting efforts, which lead to trial delays and substantial additional costs. A linear increase of the non-adherence rate in a trial leads to an exponential increase in additional patients required to maintain the statistical power of the outcomes. For example, a study in which half of the patients are non-adherent means an additional 200% of patients need to be recruited to keep the statistical power of the results stable[xi]. Improved patient monitoring and coaching methods during ongoing trials can be used to lower the adherence burden, make endpoint detection more efficient, and thus reduce dropout and non-adherence rates[xii]. AI techniques

Figure 2. The schematic visualizes the major ways to infuse AI into the clinical trial design pipeline. The three core design themes – cohort composition, patient recruitment, and patient monitoring (top row) – are based on patient features regarding suitability, eligibility, enrolment empowerment, and motivation, as well as trial features including endpoint detection, adherence control, and patient retention (second row). A variety of design methodologies (third row) are used to implement target functionalities (fourth row). These functionalities are enabled through individual combinations of the three main AI technologies: machine/deep learning, reasoning, and human–machine interfaces (fifth row) which each analyzes a specific set of patient- and functionality-specific data sources (sixth row). The relative improvement brought about by such implementations on the study outcome is indicated by the length of the horizontal lines in the color bar code underneath the main outcome aspects (seventh row). Every AI-based study design application is directly dependent on the quality and amount of data it can tap into, and hence faces the same fundamental challenges (bottom row). Abbreviation: EMR, electronic medical record.

in combination with **wearable** technology offer new approaches to developing such power-efficient, mobile, real-time, and personalized patient monitoring systems. We review some examples in the sections below.

Patient Adherence Control, Endpoint Detection, and Retention

To comply with adherence criteria, patients are required to keep detailed records of their medication intake and of a variety of other data-points related to their bodily functions, response to medication, and daily protocols. This can be an overwhelming and cumbersome task, leading to on average 40% of patients becoming non-adherent after 150 days into a clinical trial [26]. Wearable sensors and video monitoring can be used to automatically and continuously collect patient data, thereby relieving the patient of this task. ML and particularly DL models can then be used to analyze such data in real-time for detecting and logging events of relevance (Figure 2). This approach allows disease diaries to be generated which – because the underlying analytical DL models are periodically retrained with updated measurement data – evolve to be patient-specific and adaptive to any changes in disease expression and patient behavior. Such disease diaries may serve as evidence for adherence or lack thereof and – as minimal or no manual patient input is required – will also collect data-points for endpoint detection more reliably and efficiently than current patient-driven self-monitoring methods. AI also has an important role to play in image-based endpoint detection – a task that is currently addressed manually at reading centers. ML technologies have been proposed[xiii] [27] – and recently approved [28,29] – for screening applications for the rapid detection of diseases from medical images. Complementing this with algorithms that quantify pathological conditions [30–32] will reduce the cost associated with image-based studies by circumventing manual processing.

AI and ML methods may also be used to dynamically predict the risk of dropout for a specific patient, in other words to detect the onset of patient behavior that suggests the patient might be experiencing issues with adhering to the study protocol (Figure 2). One such example described the use of deep reinforcement learning algorithms by Yauney and Shah [33] to determine the fewest, smallest doses that could still shrink brain tumors, while reducing toxicity associated with chemotherapy dosing regimens. Powered by a 'self-learning' ML technique, the system looks at treatment regimens currently in use, and iteratively adjusts the doses. Eventually, it finds an optimal treatment plan, with the lowest possible potency and frequency of doses that should still reduce tumor sizes to a degree comparable to that of traditional regimens. In simulated trials of 50 patients, the ML model designed treatment cycles that reduced the potency to a quarter or half of nearly all the doses while maintaining the same tumor-shrinking potential, and thus promises improvements in patient adherence and reductions in dropouts and censuring. Picking up early warning signs for non-adherence allows proactive engagement with individual patients and permits the root causes of problematic behavior to be addressed: for example, severe side effects or incompatibility of study and personal routines could be detected and remedied before they lead to dropout. The choice of sensors and analytical models is highly disease-specific and will need to be part of the clinical study design.

Using DL for object recognition in images and video, as well as for analyzing **time-series data** from wearable sensors, first studies for testing and exploring AI-assisted patient monitoring systems have recently been started[xiv] or completed successfully[xv] [34]. The advent of commercially available wearable devices with medical-grade health-sensing capabilities[xvi], as well as complementary software ecosystems for running advanced DL models[xvii–xix] on such mobile platforms, will allow more diversified sensor combinations to be investigated for a variety of diseases. In a previous study, Shah *et al.* evaluated the significance and efficacy of clinical evidence generated from advanced technology-enabled non-invasive diagnostic screening (TES) using low-cost

smartphones and other point-of-care medical sensors versus conventional vital signs examination. They report that, although routine health screening continues to be important, the emerging techniques of TES can play an important synergistic role in stratifying populations and providing personalized screening and care in support of clinical trial designs and observational studies to generate innovative, new treatment approaches [35]. We expect to see more pilot studies benchmarking the impact of such technologies on trial efficiency alongside ongoing clinical trials in the near future; to illustrate this, in the following we give a detailed view into neurology.

Neurological diseases occupy a special role regarding drug development: neurology trials are among the four lowest-performing trial types alongside cardiovascular, psychiatry, and oncology trials [2]. It is exceptionally challenging to monitor patients in neurological trials for adherence control and endpoint detection. Often the nature of acute episodes of neurological disorders makes it impossible for patients to self-monitor, to control their behavior, or to keep an event log. For example, an epileptic patient experiencing an **absence seizure** will simply not be able to self-report the incident. A patient going through prolonged depressive episodes might decide to not take medication and also to not report such a deviation from the trial protocol. Even third-party monitoring by an experienced medical practitioner or independent observer does not allow reliable event-logging in most neurological diseases.

The reason for this complication is predominantly that neurological diseases usually tend to be highly individualized, in other words they look different in individual patients, and even for a specific patient their manifestations tend to change over time. This makes the diagnosis and treatment of neurological disorders a particularly challenging endeavor: for the same disease a diagnostic profile for one patient may not apply to another patient, and even for the same patient diagnostic patterns might shift over time. This prevents following a one-size-fits-all, rule-based diagnostic and treatment path. A technology platform that can be trained to continuously analyze and interpret patient data for an individual patient as they accumulate over time, and automatically adjusts to changes in disease expression and treatment response, will be necessary to allow learning from these heterogeneous data. Recent AI hardware and software developments demonstrate that AI and wearable sensors can be combined to implement such an automatic, real-time, patient monitoring and analytics technology.

DL algorithms are uniquely positioned to cater to these requirements and to thus bring precision medicine to neurology. Recent advances in custom-developing mobile processors and coding environments allow DL models to be run close to or at the point of sensing. This transforms wearables from pure information storage and transmission devices into information digestion and analytics devices – a novel concept which we call 'cognitive sensing'. Wearables measure biometric parameters through mobile systems attached to the human body, and either store the collected data on the device or send it to the cloud for offline analysis. As the wearable revolution unfolds, a rapidly increasing number of parameter types can be monitored simultaneously, making storage and transmission of unfiltered sensor data impossible. Algorithms for analyzing, in other words continuously correlating, contextualizing, and filtering raw data in real-time directly at the point of sensing, will be necessary to extract actionable information before the need for data storage or transmission arises. DL models in combination with on-sensor data preprocessing and curation systems allow this task to be accomplished. The architecture of such wearable, autonomously operating, always-on, cognitive sensors (Figure 3) consists of the following system components: (i) minimum-footprint biosensors feeding into (ii) low-power mobile processors capable of locally running DL models with (iii) closed-loop interfaces to (iv) an event diary which instantly and proactively logs information on specific disease episodes and interacts with wearer or caregiver for patient support, guidance, and intervention. The event diary can thus utilize a local

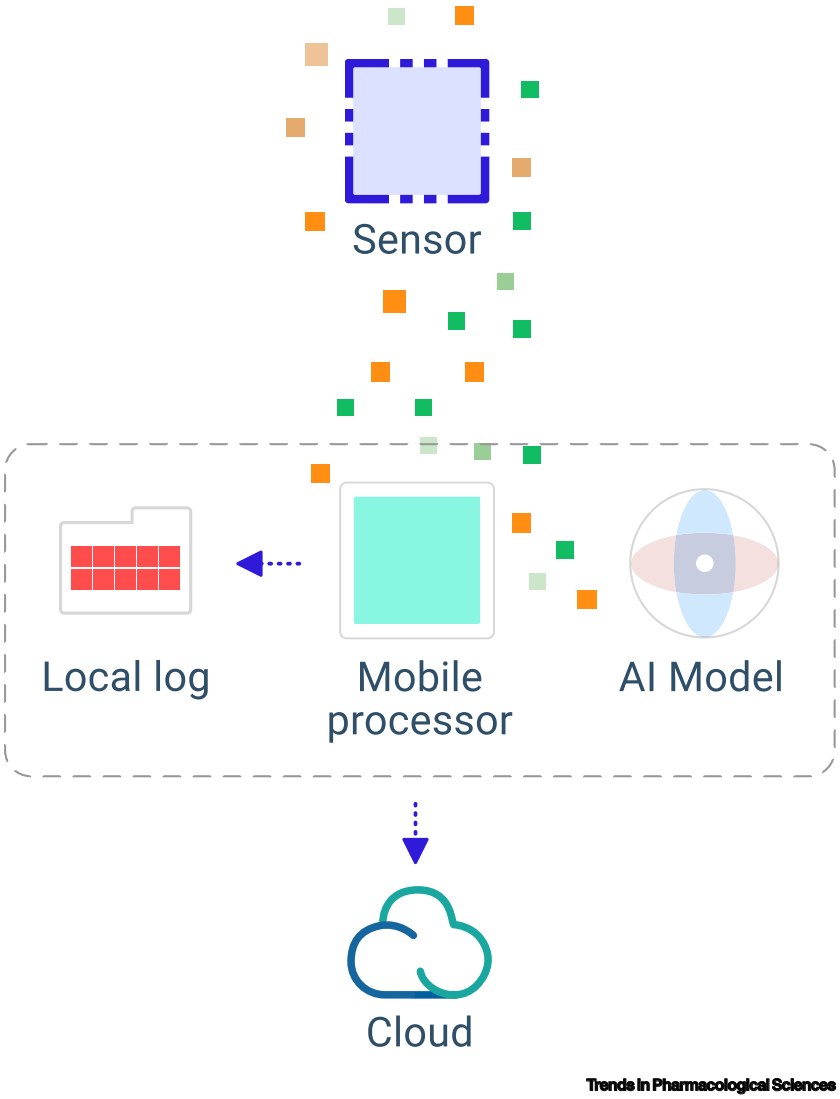


Figure 3. Cognitive Sensors. Data are measured by a wearable sensor and analyzed in real-time at the point of sensing by a mobile processor that runs an artificial intelligence (AI) model. Analysis results are then stored on a local log, in the cloud, or through a combination of both.

memory unit, a remote cloud repository, or a hybrid version of both. Various wearable biosensor and actuator platforms in different stages of technical maturity have been demonstrated or are currently under development [36] (Table 1).

The predominant type of data which most of these sensor types measure is time-series data. Although DL has traditionally focused on analyzing imagery data using deep convolutional neural networks, recent work has demonstrated that custom-designed neural network models are also uniquely suitable to analyze complex time-series streams [37–39]. To run DL algorithms continuously in real-time at the point of sensing, ultra-low-power consumption mobile processors are needed. Advances in developing both novel AI hardware and AI software techniques over the past 3 years have led to several versions of such AI-tailored mobile processing solutions now being available for real-life use. These solutions can be categorized into three general types:

Table 1. Possible Candidates for Incorporation into Cognitive Sensors

Type of sensor	Components[a]	Application	Refs
Neural implants	Retinal stimulation electrodes	Bionic Eye	[59]
	EEG and ECoG electrodes	Brain activity monitoring, deep brain-stimulation, controlling prostheses with thought	[xlii,xliii]
	Artificial skin sensors	Tactile prostheses	[60]
	Electroceuticals	Nerve- and brain- stimulation	[xliv]
Tattoo sensors	Smart contact lenses	Biomarker detection	[xlv]
	Electrochemical tattoo batteries	Multimodal data measurement	[61]
	Always-on EEG electrode tattoos		
	Low-cost integrated circuit patches		[62]
Molecular sensors	Nano- and Microfluidic sensors, portable DNA sequencers	DNA sequencing	[63]
	Smart pills, nanobiosensors, functionalized nanoparticles	Biomarker detection	[64]

[a]Abbreviations: ECoG, electrocorticography; EEG, electroencephalography.

(i) custom-developed hardware requiring custom developed AI coding environments, such as IBM's TrueNorth chip [40], (ii) custom-developed hardware compatible with standard AI programming tools, such as Qualcomm's Snapdragon chip series[xx] and Intel's Movidius processor[xxi], and (iii) conventional mobile processors which can be programmed using standard AI coding platforms, such as the Apple Watch[xvi], Apple's XS iPhone series carrying the A12 Bionic chip[xxii], and also a variety of other smartphones[xxiii].

Based on some of these techniques, the first cognitive-sensing applications have emerged in the field of applied neuroscience for monitoring and interpreting brain activity, diagnostics, and predictive prevention in epilepsy and mental disorders, deep-brain stimulation, brain–machine interfacing, and bionics: several groups have demonstrated the feasibility of using mobile AI for real-time epileptic seizure detection [41][xxiv] and prediction [42,43][xxv]. The same DL algorithm can be deployed across multiple patient cohorts, and automatically adapts to the individual disease patterns of each patient as they evolve and change over time. The demonstrated monitoring systems remain operational over extended periods of time without the need for any third-party input, and will thus become the personal seizure detector and predictor for each patient. In other demonstrations, wearable devices and ML models have been used to automatically detect cognitive and emotional states[xxvi], to monitor patients in Parkinson's disease trials[xiv], and to assess quality of sleep (among other parameters) in neurology trials[xxvii].

As pointed out previously, interoperability and standardization of data and methodology are key challenges for integration of AI into clinical trial design. The same is true for wearable AI technology and devices. Regulatory bodies, in collaboration with academic, medical, and pharma institutions, have started to produce standardization frameworks and best practice recommendations for incorporating wearable technology into clinical trial design[xxviii–xxx].

Ongoing research at the intersection of AI, **Internet of things** (IoT), and healthcare will produce more medical-grade devices with advanced analytics capabilities for continuous real-time monitoring of patients and disease progression [44][xvi]. If an equally strong focus on standardization and interoperability is maintained, these devices might make cognitive sensing an effective tool for improving the performance of neurology trials. It is important to note, however, that data

integrity and safety occupy a central role in the conception, implementation, and exploitation of digital disease diaries: patients, doctors, and regulatory bodies will rely on the integrity and safety of sensitive patient data and of analytical insights derived from it. While HIPAA-compliant environments constitute the data security baseline, advanced generations of AI-based monitoring and data-housing platforms will employ **blockchain technology** for ensuring trusted and traceable multiparty communication and exchange of monitoring data[xxxi–xxxiii].

Conclusions and Future Perspectives

The status quo of the development process for new drugs has put big pharma and other sponsors of clinical development in a dilemma [45][i] where the era of blockbuster drugs is coming to an end but the R&D process for adding new drugs to the portfolio is too slow and too expensive to compensate for this change. A fundamental transformation of the underlying business and innovation model of the entire industry is needed for a paradigm shift to a new sustainable trajectory of growth and progress.

Over the past 5 years modern AI techniques have advanced to a level of maturity that allows them to be employed under real-life conditions to assist human decision-makers in computer vision, navigation, and in some cases of medical and healthcare environments [46]. At the same time, pharma and healthcare are still among the most highly regulated and risk-averse industries. Infusing innovation that changes established processes is a difficult task that needs to be approached and implemented in a stepwise manner. Although AI has the potential to impact numerous steps of clinical trial design from preparation to execution [47], any AI pitch that aims to tackle all aspects at once is predestined for failure. Instead, data scientists and medical scientists should jointly define achievable use cases where the application of well-understood AI tools to a specific subtask of clinical trial design promises the greatest improvement of overall trial performance (Figure 2). Such AI technology first needs to be tested alongside the existing technology it aims to complement or replace, and the added value must be demonstrated and benchmarked in an explainable, ethical, repeatable, and scalable way – not only to users but also to regulatory bodies. Following this approach AI may be adopted into the clinical trial ecosystem step-by-step, making trials faster, while at the same time hopefully lowering failure rates and R&D costs. Several pharma and AI companies have started to jointly explore this avenue [47][xxxiv–xxxvii]. Regulators have put in place and continue to expand frameworks for assessing AI-based technologies in healthcare[xxxviii].

Further, completed trials have amassed a corpus of data which carries a wealth of information on correlations between trial design features and trial performance. This includes data from failed clinical trials. These large and unstructured datasets are predestined to be analyzed by AI technologies. Insights could be used to educate future improved trial designs and also to investigate the potential relevance of already trialed drugs against comorbidities for drug repurposing [8]. Nevertheless, failed trial data in particular tend to be a neglected asset that has remained largely untouched on the shelves.

It is important to note that the measurable impact of any such steps on the efficiency of the pharma R&D pipeline – even if implemented successfully now[xxxvii] – will not show up in the statistics until after a 5–8 year delay. Moreover, there will be additional R&D costs on top of the ongoing costs; in other words, from a required investment perspective, things will get worse before they will get better.

The AI techniques described in this review offer real-life practicability; however, particularly with respect to explainability, these techniques must mature to allow their broader inclusion in healthcare and life sciences applications. Although these developments are in full swing, we

Outstanding Questions

How do we collate and mine large sets of genomic data, past clinical studies, journal articles, and related real-world data, potentially distributed over multiple institutions and geographies, to improve patient selection, protocol adherence, patient eligibility computations, patient visit management, site performance, retention statistics, and adverse event detection?

Can an open repository where researchers can upload their protocols and practices, share them in public or private groups, and receive credit for them, be deployed to make these data available for training AI and ML models?

Can we create collaborative ecosystems which incentivize owners of proprietary datasets to allow their data to be collectively used to train AI and ML models, while at the same time preserving the value and honoring access and usage restrictions of these datasets?

Can AI and ML algorithms be encrypted to preserve patient identity and facilitate sharing of trained models with various stakeholders?

Will AI and ML toolboxes, developed predominantly for non-medical fields, be successfully adapted or ported over for learning, classifying, and predicting from heterogeneous medical datasets?

Is a new clinical development process for more comprehensive data collection feasible, where a network of trained healthcare professionals visit patients in their homes or places of work to collect essential data for them to participate in clinical trials?

need to acknowledge that the opportunity to transform the drug development cycle through AI comes with a great responsibility for all the disciplines involved and the mandate to qualify the value and reliability of any innovation through rigorous R&D work. This exploratory research pilot phase may not be bypassed for any reason because any breach of research protocol or premature setting of unreasonable expectations will inevitably undermine trust and ultimately the success of AI in the clinical sector.

In the same way as a change of clinical trial design alone will not turn efficiency of the pharma R&D cycle from decay to growth, AI is not a magic bullet that will make the success rates of clinical trials skyrocket overnight (see Outstanding Questions). Both reshaping clinical trial design and using AI techniques for doing so are important building blocks of a much-needed overhaul of the drug development cycle.

Acknowledgments

The authors would like to thank Josh Andres for help with designing and preparing Figure 2, Benjamin Goudey for discussions on prognostic trial enrichment, and Subhrajit Roy and Isabell Kiral-Kornek for discussions on endpoint detection in epilepsy trials. We would also like to thank Dr Ajay Royyuru from IBM Research for sharing his knowledge about clinical trial design.

Resources

[i] https://endpts.com/pharmas-broken-business-model-an-industry-on-the-brink-of-terminal-decline/
[ii] www.clinicalleader.com/doc/the-high-price-of-failed-clinical-trials-time-to-rethink-the-model-0001
[iii] www.clinpal.com/clinpal-blog/recruitment-infographic/
[iv] https://trialfacts.com/patient-recruitment-timelines-why-they-should-be-important-to-you/
[v] www.cognizant.com/whitepapers/patients-recruitment-forecast-in-clinical-trials-codex1382.pdf
[vi] www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM332181.pdf
[vii] www.wsj.com/articles/bristol-myers-opdivo-failed-to-meet-endpoint-in-key-lung-cancer-study-1470400926
[viii] www.techemergence.com/ai-machine-learning-clinical-trials-examining-x-current-applications/
[ix] www.appliedclinicaltrialsonline.com/clinical-trial-recruitment-digital-era-some-smart-ideas
[x] http://social.eyeforpharma.com/clinical/5-tips-how-facilitate-clinical-trial-recruitment
[xi] www.appliedclinicaltrialsonline.com/non-adherence-direct-influence-clinical-trial-duration-and-cost
[xii] www.appliedclinicaltrialsonline.com/mhealth-perspective-patient-centricity
[xiii] www.ibm.com/blogs/research/2018/10/glaucoma-detection
[xiv] www.ibm.com/blogs/research/2017/04/monitoring-parkinsons-disease/
[xv] https://britehealth.co/
[xvi] www.forbes.com/sites/davidphelan/2018/09/27/apple-watch-series-4-when-will-the-ecg-feature-arrive-in-the-states-and-worldwide/#6c4061193a09
[xvii] www.tensorflow.org/lite/
[xviii] https://code.fb.com/ml-applications/delivering-real-time-ai-in-the-palm-of-your-hand/
[xix] https://developer.apple.com/documentation/coreml
[xx] www.qualcomm.com/snapdragon
[xxi] https://blogs.windows.com/buildingapps/2018/03/07/ai-platform-windows-developers/
[xxii] https://venturebeat.com/2018/09/12/apple-unveils-7-nanometer-a12-bionic-chip-for-new-iphones/
[xxiii] https://venturebeat.com/2018/11/21/the-best-smartphones-for-the-ai-enthusiast/
[xxiv] www.wired.co.uk/article/empatica-embrace-epilepsy-wearable-medical-device
[xxv] www.ibm.com/blogs/research/2017/12/ai-epileptic-seizure-prediction/
[xxvi] www.affectiva.com/how/deep-learning-at-affectiva/
[xxvii] www.mobihealthnews.com/content/shimmer-unveils-wearable-sensor-platform-clinical-trials
[xxviii] https://healthpolicy.duke.edu/sites/default/files/atoms/files/duke-margolis_mhealth_action_plan.pdf
[xxix] www.mobihealthnews.com/content/fda-looking-real-world-evidence-device-evaluation

[xxx]www.mobihealthnews.com/content/rising-use-consumer-devices-clinical-trials-has-pharma-hungry-data-standards

[xxxi]https://medium.com/@kabagak/blockchain-in-pharma-clinical-trial-data-management-556ee88b46b9

[xxxii]www.iconplc.com/insights/blog/2018/06/21/harnessing-blockchain-technology/

[xxxiii]https://pharmaphorum.com/views-and-analysis/how-blockchain-will-revolutionise-clinical-trials-clinical-trials/

[xxxiv]https://m.nasdaq.com/article/3-emerging-technologies-set-to-transform-the-pharma-world-cm944011

[xxxv]https://health.economictimes.indiatimes.com/news/industry/analytical-tools-and-the-ai-will-play-a-pivotal-role-in-the-pharma-industry-sumit-peer/65777024

[xxxvi]www.mobihealthnews.com/content/mayo-clinic-finds-ibm-watson-increases-enrollment-clinical-trials

[xxxvii]www.techemergence.com/ai-in-pharma-and-biomedicine/

[xxxviii]www.fda.gov/newsevents/newsroom/pressannouncements/ucm604357.htm

[xxxix]www.cdc.gov/ehrmeaningfuluse/introduction.html

[xl]www.technologyreview.com/s/612404/is-this-ai-we-drew-you-a-flowchart-to-work-it-out/

[xli]www.technologyreview.com/s/612437/what-is-machine-learning-we-drew-you-another-flowchart/

[xlii]https://about.unimelb.edu.au/newsroom/news/2018/december/stentrode-developed-for-brain-treatments-without-major-surgery

[xliii]https://thenextweb.com/artificial-intelligence/2018/08/17/developing-bionics-how-ibm-is-adapting-mind-control-for-accessibility/

[xliv]www.scientificamerican.com/article/electroceuticals/

[xlv]www.zdnet.com/article/how-smart-contact-lenses-will-help-keep-an-eye-on-your-health/

References

1. Scannell, J.W. *et al.* (2012) Diagnosing the decline in pharmaceutical R&D efficiency. *Nat. Rev. Drug Discov.* 11, 191–200
2. Thomas, D.W. *et al.* (2016) *Clinical Development Success Rates 2006–2015*, BIO, Biomedtracker, and Amplion
3. Hay, M. *et al.* (2014) Clinical development success rates for investigational drugs. *Nat. Biotechnol.* 32, 40–51
4. Wong, C.H. *et al.* (2019) Estimation of clinical trial success rates and related parameters. *Biostatistics* 20, 273–286
5. Harrer, S. (2015) Measuring life: sensors and analytics for precision medicine. In *Bio-MEMS and Medical Microdevices II* (van den Driesche, S., ed.), pp. 51802-1–951802-5, SPIE
6. Bakkar, N. *et al.* (2018) Artificial intelligence in neurodegenerative disease research: use of IBM Watson to identify additional RNA-binding proteins altered in amyotrophic lateral sclerosis. *Acta Neuropathol.* 135, 227–247
7. Spangler, S. *et al.* (2014) Automated hypothesis generation based on mining scientific literature. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1877–1886, ACM
8. Chen, Y. *et al.* (2016) IBM Watson: how cognitive computing can be applied to big data challenges in life sciences research. *Clin. Ther.* 38, 688–701
9. Fanda, J.M. *et al.* (2018) Advances in electronic phenotyping: from rule-based definitions to machine learning models. *Annu. Rev. Biomed. Data Sci.* 1, 53–68
10. Goudey, B. *et al.* (2019) A blood-based signature of cerebrospinal fluid $A\beta_{1-42}$ status. *Sci. Rep.* 9, 4163
11. Palmqvist, S. *et al.* (2019) Alzheimer's Disease Neuroimaging Initiative, 2018. Accurate risk estimation of β-amyloid positivity to identify prodromal Alzheimer's disease: cross-validation study of practical algorithms. *Alzheimers Dement.* 15, 194–204
12. Romero, K. *et al.* (2015) The future is now: model-based clinical trial design for Alzheimer's disease. *Clin. Pharmacol. Ther.* 97, 210–214
13. Ghosh, S. *et al.* (2017) An exploration of latent structure in observational Huntington's disease studies. *AMIA Jt. Summits Transl. Sci. Proc.* 2017, 92–102
14. Sun, Z. *et al.* (2017) A data driven method for generating robust symptom onset indicators in disease registry data. *AMIA Annu. Symp. Proc.* 2017, 1635–1644
15. Sun, Z. *et al.* (2019) A probabilistic disease progression modelling approach and its application to integrated Hunting's disease observational data. *JAMIA Open* 2, 123–130
16. Che, C. *et al.* (2017) An RNN architecture with dynamic temporal matching for personalized predictions of Parkinson's disease. In *Proceedings of the 2017 SIAM International Conference on Data Mining*, pp. 198–206, SIAM
17. Schobel, S.A. *et al.* (2017) Motor, cognitive, and functional declines contribute to a single progressive factor in early HD. *Neurology* 89, 2495–2502
18. Young, T. *et al.* (2018) Recent trends in deep learning based natural language processing. *IEEE Comput. Intell. Mag.* 13, 55–75
19. Fogel, D.B. (2018) Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: a review. *Contemp. Clin. Trials Commun.* 11, 156–164
20. LeCun, Y. (2018) The Power and Limits of Deep Learning: in his IRI Medal address, Yann LeCun maps the development of machine learning techniques and suggests what the future may hold. *Res. Technol. Manag.* 61, 22–27
21. LeCun, Y. *et al.* (2015) Deep learning. *Nature* 521, 436–444
22. Helgeson, J. *et al.* (2018) Clinical performance pilot using cognitive computing for clinical trial matching at Mayo Clinic. *J. Clin. Oncol.* 36, e18598
23. Reisman, M. (2017) EHRs: the challenge of making electronic data usable and interoperable. *Pharm. Ther.* 42, 572–575
24. Orlando, A.W. and Rosoff, A.J. (2019) The new privacy crisis: what's health got to do with it? *Am. J. Med.* 132, 127–128
25. Topol, E.J. (2015) *The Patient Will See You Now: The Future of Medicine Is in Your Hands,* Basic Books
26. Blaschke, T.F. *et al.* (2012) Adherence to medications: insights arising from studies on the unreliable link between prescribed and actual drug dosing histories. *Annu. Rev. Pharmacol. Toxicol.* 52, 275–301
27. Pallab, R. *et al.* (2017) A novel hybrid approach for severity assessment of diabetic retinopathy in colour fundus images. Biomedical Imaging (ISBI 2017), pp. 1078–1082
28. Abràmoff, M.D. *et al.* (2018) Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *npj Digit. Med.* 1, 39
29. Rodriguez-Ruiz, A. *et al.* (2019) Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. *J. Natl. Cancer Inst.* Published online March 5, 2019. https://dx.doi.org/10.1093/jnci/djy222
30. Suman, S. *et al.* (2018) Joint segmentation and uncertainty visualization of retinal layers in optical coherence tomography images using Bayesian deep learning. In *Computational*

Pathology and Ophthalmic Medical Image Analysis (Stoyanov, D., *et al.*, eds), pp. 219–227, Springer
31. Mahapatra, D. *et al.* (2018) Deformable medical image registration using generative adversarial networks. Biomedical Imaging (ISBI 2018), pp. 1449–1453
32. De Fauw, J. *et al.* (2018) Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat. Med.* 24, 1342
33. Yauney, G. and Shah, P. (2018) Reinforcement learning with action-derived rewards for chemotherapy and clinical trial dosing regimen selection. *PMLR* 85, 161–226
34. Labovitz, D.L. *et al.* (2017) Using artificial intelligence to reduce the risk of nonadherence in patients on anticoagulation therapy. *Stroke* 48, 1416–1419
35. Shah, P. *et al.* (2018) Technology-enabled examinations of cardiac rhythm, optic nerve, oral health, tympanic membrane, gait and coordination evaluated jointly with routine health screenings: an observational study at the 2015 Kumbh Mela in India. *BMJ Open* 8, e018774
36. Yetisen, A.K. *et al.* (2018) Wearables in medicine. *Adv. Mater.* 30 1706910 1-26
37. Roy, S. *et al.* (2018) ChronoNet: a deep recurrent neural network for abnormal EEG identification. *arXiv* Published online January 30, 2018. https://arxiv.org/abs/1802.00308
38. Asif, U. *et al.* (2019) SeizureNet: a deep convolutional neural network for accurate seizure type classification and seizure detection. *arXiv* Published online March 8, 2019. https://arxiv.org/abs/1903.03232
39. Roy, S. *et al.* (2019) Machine learning for seizure type classification: setting the benchmark. *arXiv* Published online February 4, 2019. https://arxiv.org/abs/1902.01012
40. Merolla, P.A. *et al.* (2014) A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science* 8, 668–673
41. Ulate-Campos, A. *et al.* (2016) Automated seizure detection systems and their effectiveness for each type of seizure. *Seizure* 40, 88–101
42. Kiral-Kornek, I. *et al.* (2018) Epileptic seizure prediction using big data and deep learning: toward a mobile system. *EBioMedicine* 27, 103–111
43. Stacey, W. (2018) Seizure prediction is possible – now let's make it practical. *EBioMedicine* 27, 3–4
44. Gabrani, M. *et al.* (2018) *When data meets disease head-on: new trends in treating and managing epilepsy*. https://researcher.watson.ibm.com/researcher/files/au1-sharrer/EpilepsyWP_Nov2018.pdf
45. Weiss, D. *et al.* (2009) The 'big pharma' dilemma: develop new drugs or promote existing ones? *Nat. Rev. Drug Discov.* 8, 533–534
46. Walczak, S. (2018) The role of artificial intelligence in clinical decision support systems and a classification framework. *Int. J. Comput. Clin. Pract.* 3, 31–47
47. Lee, J. and Varadharajan, D. (2018) *The Future of Clinical Trials: How AI and Big Tech Could Make Drug Development Cheaper, Faster, and More Effective,* CBInsights
48. Clancey, W.J. and Shortliffe, E.H. (1984) *Readings in Medical Artificial Intelligence: The First Decade,* Addison-Wesley Longman
49. McCauley, N. and Ala, M. (1992) The use of expert systems in the healthcare industry. *Inf. Manag.* 22, 227–235
50. Niu, F. *et al.* (2011) HOGWILD!: a lock-free approach to parallelizing stochastic gradient descent. *arXiv* Published online June 28, 2011. https://arxiv.org/abs/1106.5730
51. Wang, F. *et al.* (2018) Deep learning in medicine – promise, progress and challenges. *JAMA Intern. Med.* 179, 293–294
52. Gulshan, V. *et al.* (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 316, 2402–2410
53. Esteva, A. *et al.* (2017) Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542, 115–118
54. Topol, E. (2019) High-performance medicine: the convergence of human and artificial intelligence. *Nat. Med.* 25, 44–56
55. Mohan, S. *et al.* (2017) Deep learning for biomedical information retrieval: learning textual relevance from click logs. In *Proc. BioNLP 2017 Workshop*, pp. 222–231, Association for Computational Linguistics
56. Tan, J. *et al.* (2014) Unsupervised feature construction and knowledge extraction from genome-wide assays of breast cancer with denoising autoencoders. *Pac. Symp. Biocomput.* 20, 132–143
57. Zurada, J. (1994) End effector target position learning using feedforward with error back-propagation and recurrent neural networks. In *Proc. 1994 IEEE International Conference on Neural Networks (ICNN'94)*, pp. 2633–2638, IEEE
58. Wang, Y. and Zeng, J. (2013) Predicting drug–target interactions using restricted Boltzmann machines. *Bioinformatics* 29, i126–i134
59. Hadjinicolaou, A.E. *et al.* (2015) Prosthetic vision: devices, patient outcomes and retinal research. *Clin. Exp. Optom.* 98, 395–410
60. Ahmed, A. *et al.* (2019) An ultra-shapable smart sensing platform based on a multimodal ferrofluid-infused surface. *Adv. Mater.* 31, 1807201
61. Norton, J.J.S. *et al.* (2015) Soft, curved electrode systems capable of integration on the auricle as a persistant brain-computer interface. *Proc. Natl. Acad. Sci. U. S. A.* 112, 3920–3925
62. Yang, S. *et al.* (2015) Cut-and-paste manufacture of multiparametric epidermal sensor systems. *Adv. Mater.* 27, 6423–6430
63. Harrer, S. *et al.* (2015) Label-free screening of biomolecules through resistive pulse-sensing technology for precision medicine applications. *Nanotechnology* 26, 182502
64. Shandilya, R. *et al.* (2019) Nanobiosensors: point-of-care approaches for cancer diagnostics. *Biosens. Bioelectron.* 130, 147–165