AI and machine learning for clinical pharmacology

David K Ryan[1], Rory H Maclean[1,3], Alfred Balston[2], Andrew Scourfield[1], Anoop D Shah[1,3], Jack Ross[1]

1. Department of Clinical Pharmacology, University College London Hospitals NHS Foundation Trust, London, UK.
2. Department of Clinical Pharmacology, Guy's and St Thomas' NHS Foundation Trust, London, UK.
3. Institute of Health Informatics, University College London, 222 Euston Road, London, NW1 2DA, UK.

Keywords:
Artificial intelligence; machine learning; clinical pharmacology; clinical trials; real-world data

Word count:
5, 589 words

Abstract

Artificial intelligence (AI) will impact many aspects of clinical pharmacology including drug discovery and development, clinical trials, personalised medicine, pharmacogenomics, pharmacovigilance and clinical toxicology. The rapid progress of AI in healthcare means clinical pharmacologists should have an understanding of AI and its implementation into clinical practice. As with any new therapy or health technology, it is imperative that AI tools are subject to robust and stringent evaluation to ensure that they enhance clinical practice in a safe and equitable manner. This review serves as an introduction to AI for the clinical pharmacologist, highlighting current applications, aspects of model development and issues surrounding evaluation and deployment. The aim of this article is to empower clinical pharmacologists to embrace and lead on the safe and effective use of AI within healthcare.

Introduction

Artificial intelligence (AI) is the use of software to perform tasks that were traditionally thought to require human intelligence, and there is growing interest of AI in healthcare.[1] Improvements in computing power, data availability, and developments in machine learning (ML) algorithms have led to rapid progress in the field,[2] as evident by the exponential growth in the number of AI publications.[1]

However, there have been a number of “false dawns” for AI in healthcare, with concerns that so far AI has over-promised and under-delivered.[3] It is important to critically evaluate the benefits and potential negative effects with the same stringency we would for any new therapy or

intervention.[4,5] AI has specific strengths and weaknesses that need to be considered. Indeed, despite approval by regulators, there are growing concerns about the clinical applicability of AI tools in medicine.[6] These concerns relate to the robustness of model development, accuracy, ability to generalise to wider populations and how models sit within existing pathways and systems among other factors.[6]

Whilst many areas of healthcare can be positively impacted by innovative AI technologies, clinical pharmacologists will be well placed to lead the development, evaluation and implementation of AI technologies in healthcare systems. For example, criteria for evaluating AI have been developed in line with those used in the regulation and evaluation of medicines.[7,8] The specialty's expertise and experience in trials, safety monitoring boards and regulatory systems will all be transferable to the development and roll-out of AI tools in clinical practice. The regulation of AI tools is a major challenge and governments worldwide are beginning to design legislation and safeguards to ensure the safe implementation of AI in all fields. It is important that developers of any AI tool, not only medical AI, are aware of existing regulatory landscapes and forthcoming regulations such as the European Commission AI Act.

Despite the growing relevance of medical AI, it is rarely mentioned in medical curricula and many doctors feel that they have not had sufficient training to understand how to evaluate it.[9] Educating clinicians in AI has been identified as a priority by the World Medical Association[10], the UK National Institute for Health Research[11] and the UK National Health Service.[12] We highlight key areas of model development, evaluation and appraisal through the use of an exemplar machine learning model, with publicly available data and code allowing readers to delve deeper into aspects of model development. In addition, we will emphasise, through

examples, the potential for AI in clinical pharmacology and discuss concerns relating to the implementation of AI into clinical environments. The emphasis of this article is on the clinical aspects of AI; detailed discussions of the computer science behind different machine learning algorithms are beyond the scope of this article.[2,13]

What is medical AI?

AI relies on algorithms, which can be defined as a set of instructions for completing a specific task. Historically most AI algorithms used rules-based systems written by humans. Recent advances in AI have been driven by machine learning, where algorithms are trained through the use of examples rather than explicit instruction, iteratively processing data and adjusting the model's internal parameters to improve performance.[2,13]

There are a number of different machine learning models, which can be broadly divided into unsupervised and supervised learning.[2,13] Unsupervised learning models work with unlabelled data to identify patterns. These include clustering models, which could be used to group patients with similar prescription patterns, and dimensionality reduction models, which are useful for analysing complex datasets such as genetic markers.[2,13] Supervised learning models use labelled data, learning from examples to map inputs to known outputs. These include regression models, which could use demographic and clinical data to predict the optimal drug doses for patients, or classification models, which could be used to identify toxidromes to diagnose poisonings.[2,13] Other supervised learning models commonly used in healthcare include decision trees and neural networks.

Decision tree models can be used to make predictions from labelled data - for example, a decision tree model could be applied to a database of patients to predict whether a person developed an outcome, such as stroke, following an exposure to an anti-hypertensive medication. Decision tree models make predictions by separating cases from non-cases, or determining the value of an output, based on features in the data. The model structure is said to resemble a tree, where each feature acts as a branch – splitting the cases from the non-cases based on the values of certain features.[2,13] The model learns to predict based on a training dataset consisting of the features in the model and the labelled outcome. These models have the advantage of being relatively easy to understand and interpret, but they are vulnerable to overfitting, where they learn from noise in the training examples and so do not perform as well on unseen data. Random forest models are a common decision tree-based model and aim to address this by using multiple different decision trees, each with a random subset of features, to arrive at a combined prediction or classification.[2,13]

Neural networks models have a structure inspired by the biological networks of neurons in a brain, though using mathematical functions rather than electrochemical processes to transmit information at synapses.[14] Their architecture can be complicated, with deep learning models having many layers of connections, and it can be hard for humans to interpret how they reach their outputs. Nevertheless, they have useful clinical applications, and various neural networks are better at different tasks. Convolutional neural networks are well suited to detect features in images and can be used for image recognition and segmentation tasks such as detecting cancer or measuring response to chemotherapy.[2,14] Recurrent neural networks have a design that allows them to process sequential data and could be used to monitor drug effects over time.[2] Generative

adversarial networks can create realistic synthetic text and images, which could be used to design cases for medical education or generate synthetic images to train further classification models.[15]

Transformer models are a type of model that focus on identifying words in context and can be used for natural language processing to analyse large datasets like the scientific literature or electronic health records.[16] There has been particular interest in using these to make large language models, such as GPT-4, LLaMa and PaLM, which can be used to create chatbots to support both patients and clinicians.

More recently, there has been a rise in code-light and code-free models where the algorithm makes automatic decisions regarding feature selection and engineering, model selection and parameter settings, making them more approachable for clinicians without programming experience. These include proprietary automated machine learning systems[17] and nnU-net for biomedical image segmentation.[18]

How can AI be used in clinical pharmacology?

AI has the potential to revolutionise many core areas of clinical pharmacology in addition to medical diagnostics, prognostication and biomarker development. AI will also likely transform areas of specialist interest for many clinical pharmacologists such as: drug discovery and development,[19,20] clinical trials,[21] medicines management,[22] personalised medicine and precision dosing,[23,24] pharmacogenomics,[25] pharmacovigilance and adverse drug reactions,[26] and clinical toxicology[27] (figure 1). The use of AI in pre-clinical pharmacology is extensively covered elsewhere.[28,29]

Diagnosis and prognostication

The most significant advancements of AI in healthcare have been made in diagnostic support and imaging analytics. This is evident by the high proportion of diagnostic and imaging analytics tools among the 178 approved AI algorithms by the Food and Drug Administration.[30] Examples of approved technologies include arrhythmia detection using smart watch data, image pre-processing in radiology and real-time detection of colonic mucosal lesions during endoscopy.[30] The relative proliferation of AI imaging tools likely reflects the advanced nature of computer vision algorithms, particularly convolutional neural networks, and the availability of extensive databases of anonymised clinical images like chest x-ray, retinal photographs and skin lesions.

Prognostication and risk-stratification tools can be developed using machine learning, which allows for flexible modelling of complex relationships between covariates in underlying datasets. For example, a decision tree-based algorithm was used to stratify risk of patients who presented with symptoms of acute coronary syndrome using clinical factors and high-sensitivity troponin, using a dataset of 10,038 patients.[31] This model was reported to have excellent discrimination for myocardial infarction and increased the identification of low-risk myocardial infarctions compared to fixed troponin thresholds in an external validation set.[31] Other machine learning risk stratification tools have been developed for many conditions including COVID-19,[32,33] heart failure,[34] and various cancers.[35] However, there remain concerns about the validation and generalisability of these tools, and the impact of AI risk scores on clinical outcomes in practice has not been established.

There are numerous studies that utilise AI models to predict clinical events, particularly acute deterioration. Many of these studies are based on longitudinal and high-dimensional critical care datasets, though can expand to other medical specialties as data availability and remote monitoring becomes more widespread. The HAVEN model was developed in the UK to predict cardiac arrest or unplanned intensive care admissions among 230,000 patients admitted to hospital.[36] This was a real-time, tree-based machine learning model based on retrospective data that used clinically-determined static (time invariant – e.g. demographics) and dynamic variables (time-varying – e.g. vital signs, laboratory results). The model was able to accurately predict deterioration at 12 hours, with slightly worse performance at 48 hours. This model was reported to outperform standard early warning systems, identifying more than 40% of cardiac arrests or unplanned intensive care admissions within the preceding 48 hours and provide as much as 12-hour notice for more than 25%.[36] In addition, machine learning can be used to model disease trajectories and detect signals for deteriorations in chronic diseases such as cystic fibrosis (CF).[37,38] A study among 147 participants with CF showed that data from home monitoring (weight, lung function, oximetry) and wearable technologies (pulse rate, activity) could be utilised to predict acute pulmonary exacerbations, with clustering analysis showing distinct profiles for different types of exacerbations.[37,38] This highlights the role of AI in predicting disease trajectories, anticipating future healthcare needs and allowing bespoke personalised interventions at an earlier stage.

Novel biomarker development

There are a huge range of applications for AI within the diagnostic pathway, particularly for identifying novel digital biomarkers. Retinal images have been studied extensively and deep neural networks have shown potential to detect not only retinal pathology such as diabetic

retinopathy and other retinal diseases,[39] but also wider systemic illness including schizophrenia,[40] anaemia,[41] cardiovascular disease[42] among others. The implications of such digital biomarkers are yet to be fully explored but with the widening access of retinal imaging technologies, it is possible that these tools will become used much more broadly across medical specialties. In addition, voice recognition models have been suggested to predict heart failure exacerbations, although the exact signals mediating this are yet to be fully elucidated.[43] There is also a growing body of work where models combine data across different modalities such as text, image, sensor and tabular data, so-called 'multimodal' sources, to make predictions.

In imaging, machine learning can extract large amounts of quantitative data such as intensity, gradient and texture and use this to predict outcomes, a technique known as radiomics when used in radiology[44] and pathomics when applied to pathology.[45] This has shown potential in breast cancer, using whole slide histopathology images to predict response to neoadjuvant chemotherapy,[46] as well as in ovarian cancer, using CT images to predict response to immunotherapy.[47]

Drug discovery and development

The rising costs of drug development present a significant challenge for the pharmaceutical industry. There is a growing investment in AI tools to improve target identification, target validation, molecule selection and optimisation, with an anticipation that AI supported research and development could lead to significant time and cost savings, and improve the rate of translational successes.[48] Several AI tools are already in use in drug discovery, particularly using deep neural network models. These include AlphaFold to predict 3D protein structures,[49]

DeepAffinity to predict drug-protein interactions,[50] and DeepTox to predict toxicity of candidate molecules.[51] This is an expansive topic, and has been covered in recent review articles.[52]

AI tools can be used to simplify the growing availability of -omics data to identify previously unknown disease pathways and drug targets.[48] One such approach is to incorporate this data in the form of a knowledge graph, which describes the relationships between data related to a gene, protein, signalling pathway, drug, symptom or disease among other entities.[53,54] These knowledge graphs can be analysed by machine learning models to help repurpose existing drugs for new indications or identify novel drug targets.[53,54] This method was used during the COVID-19 pandemic to identify baricitinib, a janus kinase inhibitor, as a potential therapeutic, leading to its inclusion in the RECOVERY trial.[55]

Clinical trials

AI has been proposed to help improve efficiency and generate better evidence for treatments throughout clinical trial design, conduct, and analysis. A registry based audit of ClinicalTrials.gov showed AI studies were being registered across a range of specialties, most commonly in oncology, followed by cardiology, ophthalmology and psychiatry.[56] AI can be used to optimise trial design and selection of patients. AI has been applied to trial registries to identify the impact of clinical trial eligibility criteria on the probability of trial termination.[57] It was reported that features such as number of countries in the study and exclusion of inflammatory conditions were important in determining trial termination.[57] In addition, AI can help identify populations with higher event rates in order to reduce the total sample size required to detect a treatment effect. A machine learning model was used to predict which individuals with acute respiratory distress syndrome would have longer hospital admissions in order to focus

recruitment.[58] However, the use of an AI tool to enrich event-rates in a clinical trial may alter the interpretation of a trial's conclusion. Any treatment effect would now be conditional on having a high probability of an event, which may not generalise to real-world practice. This may also influence regulatory decisions – where the event-rate detection model may be required to select eligible people for the treatment in real world deployment.

AI can also be used to screen patients for inclusion into studies.[59,60] For example, the computer system IBM Watson demonstrated that an AI model could identify suitable patients for four breast cancer trials, with 86.7% accuracy and high agreement with manual patient screening.[51] In a summary meta-analysis of 19 cancer clinical trial datasets, AI was found to match patients to clinical trials well, with a summary sensitivity of 91% and specificity of 99%.[60] Furthermore, machine learning can also be used to gauge a patient's interest in trial participation.[61]

There are several applications for AI in the conduct and analysis of trials. A reinforcement learning model was used to aid with decisions relating to the timing and order of chemotherapy regimens in a non-small cell lung cancer trial.[62] An AI platform on a mobile phone was used to monitor participants adherence to anticoagulation treatment for stroke, visually identifying the patient, medication and confirming ingestion.[63] This was associated with a 50% improvement based on plasma drug concentration levels. AI may be helpful in the analysis of clinical trials, particularly in relation to imputing missing trial data, estimating personalised dosing regimens and calculating treatment effects. For example, meta-learners can be used to estimate treatment effects within individuals and heterogeneous treatment effects. These methods have been developed and tested in simulation studies, but have yet been applied to real-world data.[64]

Furthermore, random forest models and other methods have been developed to predict missing values, reducing data attrition and optimising analysis of trial data.[65]

Medicines management

AI has the potential to improve the safe and efficacious prescribing of medicines. Multiple models have been developed using electronic health record data to identify erroneous medicine prescribing, drug omissions and drug-drug interaction.[66] For example, one model applied to retrospective electronic health record data detected 10,668 potential prescribing errors over 4 years.[67] It was estimated that approximately 80% of such alerts were determined to be clinically significant.[67] No prospective trial has been conducted to test the real-world effects of such a system.

There is also a growing use-case for AI to identify the most optimal therapeutic strategy for a specific patient and improve how we integrate observational data into clinical decision making. AI, coupled with the growing availability of real-world clinical data, could lead to greater data-driven, personalised treatment decisions. One of the major concerns with the current paradigm of evidence-based medicine is that it does not readily apply to the growing number of patients with comorbidities.[68] For example, patients with multimorbidity and polypharmacy are frequently excluded from clinical trials, a matter of increasing concern within clinical pharmacology. As a result, clinical trials often do not generalise to the population that will be eligible for treatment, thus creating an evidence gap between RCTs and real-world clinical practice.[68,69]

The use of real-world data, mostly from electronic health records, is made challenging by the non-random allocation of treatments. As a result, there is a risk of making biased causal

inferences due to factors such as confounding by indication - where the outcomes is caused by a factor that precipitated the drug being prescribed in the first instance, rather than the drug itself.[70] Traditional statistical techniques to account for such confounding including matching studies, adjustment using regression models and propensity scoring. More advanced machine learning methods have been developed for selecting features to adjust in models,[71] determine individual treatment effects and treatment heterogeneity.[64]

One example of this was a study in the Swedish Heart Failure Registry which used a random forest model initially to identify features that were associated with 1-year survival.[72] A clustering algorithm was then applied to identify 4 distinct clusters within this cohort. An interaction effect was found between the clusters and various heart failure treatments such as beta-blockers, ACE inhibitors and diuretics - with differential response to these drug classes among different heart failure clusters.[72] Whilst this study was limited by the small number of features available for clustering, it demonstrates the potential role for AI in identifying clinically meaningful treatment heterogeneity that clinical trials may not be sufficiently powered to study. It is likely that with the increasing completeness of clinical data and development of sophisticated machine learning methods to analyse such data, real-world evidence will promote the efficacious use of medicines.

The use of AI to derive insights from real-world data has a potential to offer real-time, real-world insights about treatment effects, which could be incorporated into existing evidence-based guidelines to offer 'live', adaptive guidelines. The ability for informatics to contribute directly to clinical decision making has been termed *'the informatics consult'* and with the growing availability of clinical data and sophisticated state-of-the-art methods to analyse such data, it is likely that such consults could grow in scope and utility.[73]

Precision dosing

AI methods can support pharmacokinetic and pharmacodynamic analyses to guide drug dosing.[23,24] A decision tree based model was used to personalise vancomycin dosing regimens using data from electronic health records, and similar techniques have been suggested for other drugs with a narrow therapeutic index.[74] In a large multi-centre study, a reinforcement learning tool was developed to suggest the dose of erythropoietin for patients receiving haemodialysis, aiming to improve haemoglobin level while minimising the total doses of erythropoietin given.[75]

A systematic review identified eight studies reporting machine learning models to support personalised dosing of intravenous unfractionated heparin.[24] However, of these studies only one included an external validation, and no studies evaluated the model impacts on clinical practice, leading the authors to conclude more rigorous evaluation of machine learning methods were needed before these models should be used clinically.[24] Furthermore, tree-based machine learning models have been successfully utilised to inform immunosuppression dosing using either mycophenolic acid or tacrolimus in patients following a transplant. In these models, two or three blood drug-level concentrations were combined with clinical details (type of transplantation, dose, age, time from transplantation, other immunosuppressant agents) to predict area-under-the curve for immunosuppressive agents. This model performed better than the existing standard Bayesian estimation in four independent PK datasets and is accessible online. Similar approaches have shown improvements in isavuconazole exposure prediction.[76] Therapeutic drug management and model-informed precision dosing have been covered elsewhere.[77]

Pharmacogenomics

Pharmacogenomic tests are increasingly used in clinical practice.[78] The cytochrome P450 enzyme family is responsible for the metabolism of many drugs, and CYP2D6 in particular metabolises around one-fifth of licensed drugs.[78] Two approaches using machine learning have demonstrated improved identification of genetic variants that might require an actionable change in prescribing for an individual patient. A neural network model achieved more accurate prediction of CYP2D6 mediated tamoxifen metabolism in breast cancer patients compared to conventional methods.[79] A convolutional neural network trained on a combination of predominantly synthetic and limited real data was able to accurately predict CYP2D6 haplotype function.[80]

Pharmacovigilance

AI natural language processing techniques can be used to extract coded diagnoses or clinical events from free text. These techniques have been used to identify adverse drug reactions from hospital electronic health records,[26] as well as other data sources including X (formerly Twitter) and social media posts.[81] AI methods have also been used to identify signals for adverse drug reactions by regulators including as part of the UK spontaneous adverse event reporting (yellow card) scheme.[82] A machine learning model has been developed to predict the likelihood of a causal association of an observed drug-reaction combination based on individual case study reports.[83]

Clinical Toxicology

There are few studies of AI in clinical toxicology published so far. One group designed a rules-based model, which is a type of AI system explicitly designed by humans rather than

machine learning, to identify poisoned patients and differentiate between the 6 toxidromes: anticholinergic, cholinergic, opioid, sedative-hypnotic, serotonin toxicity and sympathomimetic. When tested against synthetic cases, the rules-based model outperformed a decision tree model but did not reach the level of human physicians.[27]

How is an AI model developed and evaluated?

With the growing proliferation of medical AI tools, it is important that clinicians have an awareness of AI model development and evaluation. This is essential not only to ensure the development of clinically valid and applicable models but also to guarantee their accurate and appropriate use in real-world deployment. Clinician input is vital for all stages of model development, including defining use-cases, informing decisions about data used to train models, feature selection and evaluation metrics. Furthermore, a broad understanding of medical AI is vital for any clinician using an AI tool in the future. To make a safe and effective translation into clinical use, clinicians will need to understand the strengths and weaknesses of the AI models that they are using, as well as the potential pitfalls in the development cycle of AI tools. This is particularly important from an accountability standpoint as clinical decision makers will remain responsible for the decisions made, even if they are informed by AI.[84]

In this section, we will highlight critical aspects of model development relevant to a clinical pharmacologist. An example model, with an annotated coding notebook, has been designed to emphasise key aspects of model development (table 1). This example model is a binary classification model built using publicly available data from the US National Health and

Nutritional Survey[85] to predict dysglycaemia. This model was built for educational purposes and is not intended for clinical practice. The focus is on clinical decisions involved in the development of AI and potential pitfalls which extend beyond this binary classification model. The data and code are publicly available for readers Further details available at: https://github.com/dkdryan/ai_cpt/

Data and bias

A model is only as good as the data it is trained on.[86] If a model is trained on incomplete, inaccurate or unrepresentative data, it is likely to perform poorly when deployed. These administrative or societal biases can then be introduced into algorithms, which can have detrimental effects including differential performances across groups.[87] For example, previously developed algorithms for predicting population health needs were found to underperform on black patients.[88] AI tools should be monitored to ensure they do not exacerbate health inequalities or perpetuate societal biases.

Clinical trials and the external validity of AI tools

A key concept in model evaluation is that an AI model should be evaluated in a dataset that the model was not trained on. A major concern in model development and evaluation is overfitting, where machine learning models "over-learns" from example data, focusing on the noise in the data and learning spurious correlations.[2] For example, a model trained to automate the classification of skin lesions learned to interpret images with rulers as more likely to be associated with malignancy.[89] As a result of this overfitting, the model does not work when tested in an external dataset.

The gold-standard to assess medical AI tools are randomised controlled trials. Reporting guidelines such as the Consolidated Standards of Reporting Trials - Artificial Intelligence (CONSORT-AI) and Standard Protocol Items: Recommendations for Interventional Trials - Artificial Intelligence (SPIRIT-AI) extensions have been designed for reporting trials with an AI component.[7,8] These guidelines advocate for transparency in model development, external validation of models and consideration of how the tool sits within existing clinical pathways. However, given the wide range of potential AI applications in healthcare, there is debate as to whether all such applications require an RCT prior to deployment. For example, tools intended to aid clinicians rather than directly alter patient management (such as clinical decision aids, risk prediction scores or new imaging techniques), were historically not subjected to an RCT prior to deployment in routine clinical practice. In a similar light, it could be argued that AI tools in these areas may not require RCT evidence of benefit and safety. There is a growing recognition that clinical decision tools should be validated prior to use, and this may be more pronounced when novel AI technologies are used in model development. Indeed, a recent example of a multi-centre RCT comparing an e-alert system for acute kidney injury versus standard care demonstrated that the e-alert system was associated with increased mortality in some centres.[90] This highlights that interventions, even those suspected to have low risk of harm such as electronic health record pop-ups, should still have an evaluation within clinical environments to ensure no unanticipated consequences. We believe that this is particularly relevant to clinical AI tools.

Human-computer interaction

There has been limited consideration of how to incorporate AI tools into a healthcare setting, with few studies of how AI models may affect a clinical practice and patients.[7,8,91] Automation bias and “alert fatigue” may reduce the potential benefits of an AI intervention. For example, an

RCT studying the effects of 'pop-up' boxes for acute kidney injury within the electronic health record failed to improve outcomes for patients.[90] These e-alerts were criticised for driving "alert fatigue", wherein clinicians are desensitised to, and so ignore alerts because of an overwhelming quantity being presented to them.[90] Further studies show that 90% of predicted drug interaction alerts shown to clinicians using e-prescribing software are overridden,[92] again emphasising the need for developers to consider how any AI intervention interacts with humans in a complex healthcare system.

Interpretability

A key barrier to the acceptance of clinical AI by practising clinicians is the limited transparency about how a model arrived at a decision. Such models are referred to as 'black box' models. Some models such as logistic regression and decision trees are more interpretable, and clinicians can measure how much each piece of input data impacted the output, much like classical clinical risk scores. Other models such as neural networks can have very complicated architecture making their decisions less explainable, and so evaluation requires assessing their outcomes, such as did they correctly identify a tumour on a scan or predict a diagnosis. There are emerging techniques being used to help explain AI predictions. For example, for imaging analytic tasks, such as diagnosis of a tumour from a chest X-ray, techniques such as salience mapping can be used to highlight which part of the image was important to determine a prediction, thus allowing some level of interpretability.[94]

Regulation and market authorisation

Given the rapid advancements in AI technology, the regulatory frameworks governing AI are struggling to keep pace. There is now growing societal and governmental interest in developing regulations to ensure the safe use of AI, not only for medical AI, but all AI innovations. Approval for medical AI falls under the remit of medical device regulators, for example the Medicines and Healthcare Regulation Agency (MHRA) within the UK[95] and Food and Drug Administration (FDA) within the USA.[96] Software as a medical device is regulated under the Medical Device Regulation 2002 (as amended) act within the UK - requiring evidence that new devices are safe under normal conditions of use and perform as intended.[97] The European Union is developing an overarching "AI Act" which will seek to regulate new AI technologies.[98] The field of AI regulation is rapidly evolving and adapting to ever increasing developments in technology and is likely that an international consensus will be needed to develop global AI standards.[99] Developers of AI need to be aware of such regulatory landscapes when building, testing and deploying such tools. Legal liability in AI is a contentious area, and there is uncertainty on the legal implications for errors or patient-related harm arising from medical AI tools.[100,101]

AI lifecycle and 'AI-vigilance'

Once an AI model has achieved satisfactory performance in a clinical setting, it must continue to be evaluated to monitor for adverse effects and ensure its performance does not degrade, akin to the post-marketing surveillance that occurs with newly licensed drugs. A model's performance can be eroded through data "drift", where the population changes over time away from the dataset it was trained on.[102] For example, a model predicting the diagnosis of fever in a returning traveller may need to be retrained if a vaccine for malaria enters clinical practice, or a lung cancer recognition model if a new CT scanner is used. AI models must also be evaluated for

hidden stratification, where a model's overall performance is good but is persistently poor in certain subgroups as they give too much weight to spurious correlations or have little training examples for rarer subgroups of a disease state.[103] Any errors or adverse events relating to an approved AI technology should be reported via systems such as the UK spontaneous adverse event reporting (yellow card) scheme. Furthermore, AI innovations in healthcare should be monitored to ensure that they are benefiting all of society in an equitable manner. There are also concerns that AI innovations in healthcare could lead to intervention generated inequality, disproportionately benefitting people with high levels of self-advocacy and health literacy.[104]

Limitations of medical AI

AI in clinical pharmacology is at the beginning of a wave, and it is not yet widely used in clinical practice. Before medical AI tools can transition from code to clinic, there needs to be a robust and transparent process to develop, evaluate and authorise models for clinical use. Many of these issues are familiar from drug development and RCTs, however the adoption of trial methodologies and regulatory frameworks has yet to be fully extended to clinical AI. Developers will need to overcome barriers, both technical and clinical, before being deployed into live clinical environments (table 2). Indeed, like the annual influenza vaccine, AI models may need to be continuously iterated to reflect changes in the world that might influence predictive capacity and ensure that it is optimised for real-world use. In addition, current AI models are limited in terms of their scope - particularly in that they are mostly developed to address a single, highly specific task. Medical AI is likely to remain "human in the loop" - meaning AI will augment, but

not replace, traditional clinical decision making. As such, it is unlikely that clinical pharmacologists will be replaced by AI. Rather, clinical pharmacologists will need to harness the potential of AI and apply it to clinical practice in a safe, efficient and equitable manner.

Conclusions

Recent advances in AI and machine learning have the potential to revolutionise many aspects of clinical pharmacology, with the ability to support and automate work traditionally done by healthcare professionals. Current use of AI in clinical settings is limited by lack of clinician confidence and understanding of AI, concerns around safety and applicability to clinical practice, and a lack of pathways for how these technologies will undergo regulation and evaluation. Understanding the strengths and limitations of AI models is critical prior to using them in a patient setting, and clinicians should be involved in the whole cycle of AI development and appraisal to ensure the products are safe and effective. Clinical pharmacologists, given their skills in drug development, evaluation and regulation, are ideally placed to help support the deployment of these technologies within the wider healthcare system.

Conflict of interest statement:
None to declare.

Funding information:
DKR, RHM and AB are all funded by a NIHR Academic Clinical Fellowship.

References:

1. Ross, J., Webb, C. & Rahman, F. *Artificial Intelligence in healthcare*. https://www.aomrc.org.uk/reports-guidance/artificial-intelligence-in-healthcare/.
2. Badillo, S. *et al.* An Introduction to Machine Learning. *Clin. Pharmacol. Ther.* **107**, 871–885 (2020).
3. Ge, W., Lueck, C., Suominen, H. & Apthorp, D. Has machine learning over-promised in healthcare? *Artif. Intell. Med.* **139**, (2023).
4. Meskó, B. & Görög, M. A short guide for medical professionals in the era of artificial intelligence. *Npj Digit. Med.* **3**, 1–8 (2020).
5. Corrigan, B. W. Artificial Intelligence and Machine Learning: Will Clinical Pharmacologists Be Needed in the Next Decade? The John Henry Question. *Clin. Pharmacol. Ther.* **107**, 697–699 (2020).
6. Ebrahimian, S. *et al.* FDA-regulated AI Algorithms: Trends, Strengths, and Gaps of Validation Studies. *Acad. Radiol.* **29**, 559–566 (2022).
7. Cruz Rivera, S., Liu, X., Chan, A.-W., Denniston, A. K. & Calvert, M. J. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat. Med.* **26**, 1351–1363 (2020).
8. Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J. & Denniston, A. K. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat. Med.* **26**, 1364–1374 (2020).
9. Banerjee, M. *et al.* The impact of artificial intelligence on clinical education: perceptions of postgraduate trainee doctors in London (UK) and recommendations for trainers. *BMC Med. Educ.* **21**, 429 (2021).
10. WMA Statement on Augmented Intelligence in Medical Care – WMA – The World Medical Association. https://www.wma.net/policies-post/wma-statement-on-augmented-intelligence-in-medical-care/.
11. Roy, A. Artificial intelligence for healthcare: 10 promising interventions. *NIHR Evidence* https://evidence.nihr.ac.uk/collection/artificial-intelligence-10-promising-interventions-for-healthcare/ (2023) doi:10.3310/nihrevidence_56221.
12. NHS AI Lab roadmap. *NHS Transformation Directorate* https://transform.england.nhs.uk/ai-lab/nhs-ai-lab-roadmap/.
13. Mahesh, B. *Machine Learning Algorithms -A Review*. (2019). doi:10.21275/ART20203995.
14. Choi, R. Y., Coyner, A. S., Kalpathy-Cramer, J., Chiang, M. F. & Campbell, J. P. Introduction to Machine Learning, Neural Networks, and Deep Learning. *Transl. Vis. Sci. Technol.* **9**, 14 (2020).
15. Yoon, J., Jordon, J. & Schaar, M. GAIN: Missing Data Imputation using Generative Adversarial Nets. in *Proceedings of the 35th International Conference on Machine Learning* 5689–5698 (PMLR, 2018).
16. Vaswani, A. *et al.* Attention is All you Need. in *Advances in Neural Information Processing Systems* vol. 30 (Curran Associates, Inc., 2017).

17. Karmaker (“Santu”), S. K. *et al.* AutoML to Date and Beyond: Challenges and Opportunities. *ACM Comput. Surv.* **54**, 175:1-175:36 (2021).
18. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J. & Maier-Hein, K. H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, 203–211 (2021).
19. Zhavoronkov, A., Vanhaelen, Q. & Oprea, T. I. Will Artificial Intelligence for Drug Discovery Impact Clinical Pharmacology? *Clin. Pharmacol. Ther.* **107**, 780–785 (2020).
20. Liu, Q. *et al.* Landscape Analysis of the Application of Artificial Intelligence and Machine Learning in Regulatory Submissions for Drug Development From 2016 to 2021. *Clin. Pharmacol. Ther.* **113**, 771–774 (2023).
21. Inan, O. T. *et al.* Digitizing clinical trials. *Npj Digit. Med.* **3**, 1–7 (2020).
22. Eggerth, A., Hayn, D. & Schreier, G. Medication management needs information and communications technology-based approaches, including telehealth and artificial intelligence. *Br. J. Clin. Pharmacol.* **86**, 2000–2007 (2020).
23. Ribba, B., Dudal, S., Lavé, T. & Peck, R. W. Model-Informed Artificial Intelligence: Reinforcement Learning for Precision Dosing. *Clin. Pharmacol. Ther.* **107**, 853–857 (2020).
24. Falconer, N. *et al.* Systematic review of machine learning models for personalised dosing of heparin. *Br. J. Clin. Pharmacol.* **87**, 4124–4139 (2021).
25. Turner, R. M., Magavern, E. F. & Pirmohamed, M. Pharmacogenomics: Relevance and opportunities for clinical pharmacology. *Br. J. Clin. Pharmacol.* **88**, 3943–3946 (2022).
26. Wasylewicz, A. *et al.* Identifying adverse drug reactions from free-text electronic hospital health record notes. *Br. J. Clin. Pharmacol.* **88**, 1235–1245 (2022).
27. Chary, M., Boyer, E. W. & Burns, M. M. Diagnosis of Acute Poisoning using explainable artificial intelligence. *Comput. Biol. Med.* **134**, 104469 (2021).
28. Kumar, M. *et al.* Opportunities and challenges in application of artificial intelligence in pharmacology. *Pharmacol. Rep.* **75**, 3–18 (2023).
29. van der Lee, M. & Swen, J. J. Artificial intelligence in pharmacology research and practice. *Clin. Transl. Sci.* **16**, 31–36 (2023).
30. Health, C. for D. and R. Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices. *FDA* (2022).
31. Doudesis, D. *et al.* Machine learning for diagnosis of myocardial infarction using cardiac troponin concentrations. *Nat. Med.* **29**, 1201–1210 (2023).
32. Wynants, L. *et al.* Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* **369**, m1328 (2020).
33. Shamout, F. E. *et al.* An artificial intelligence system for predicting the deterioration of COVID-19 patients in the emergency department. *Npj Digit. Med.* **4**, 1–11 (2021).
34. Guo, A., Pasque, M., Loh, F., Mann, D. L. & Payne, P. R. O. Heart Failure Diagnosis, Readmission, and Mortality Prediction Using Machine Learning and Artificial Intelligence Models. *Curr. Epidemiol. Rep.* **7**, 212–219 (2020).

35. Kumar, Y., Gupta, S., Singla, R. & Hu, Y.-C. A Systematic Review of Artificial Intelligence Techniques in Cancer Prediction and Diagnosis. *Arch. Comput. Methods Eng.* **29**, 2043–2070 (2022).
36. Pimentel, M. A. F. *et al.* Detecting Deteriorating Patients in the Hospital: Development and Validation of a Novel Scoring System. *Am. J. Respir. Crit. Care Med.* **204**, 44–52 (2021).
37. Sutcliffe, D. *et al.* Machine Learning Predicts Acute Pulmonary Exacerbations in Cystic Fibrosis. SSRN Scholarly Paper at https://doi.org/10.2139/ssrn.3750683 (2021).
38. Abroshan, M., Alaa, A. M., Rayner, O. & Schaar, M. van der. Opportunities for machine learning to transform care for people with cystic fibrosis. *J. Cyst. Fibros.* **19**, 6–8 (2020).
39. De Fauw, J. *et al.* Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat. Med.* **24**, 1342–1350 (2018).
40. Association Between Retinal Features From Multimodal Imaging and Schizophrenia | Ophthalmic Imaging | JAMA Psychiatry | JAMA Network. https://jamanetwork.com/journals/jamapsychiatry/article-abstract/2802551?casa_token=w3lB6TzkIu0AAAAA:Hc6QrHnArvICEqCbOXW_t3T1GKMgLR9_91pL_b_ZixDLBGtY0fmQvyFxilQ_kaoqDFBBY8sgjw.
41. Mitani, A. *et al.* Detection of anaemia from retinal fundus images via deep learning. *Nat. Biomed. Eng.* **4**, 18–27 (2020).
42. Poplin, R. *et al.* Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat. Biomed. Eng.* **2**, 158–164 (2018).
43. Nahar, J. K. & Lopez-Jimenez, F. Utilizing Conversational Artificial Intelligence, Voice, and Phonocardiography Analytics in Heart Failure Care. *Heart Fail. Clin.* **18**, 311–323 (2022).
44. Radiomics: Images Are More than Pictures, They Are Data | Radiology. https://pubs.rsna.org/doi/full/10.1148/radiol.2015151169.
45. Gupta, R., Kurc, T., Sharma, A., Almeida, J. S. & Saltz, J. The Emergence of Pathomics. *Curr. Pathobiol. Rep.* **7**, 73–84 (2019).
46. Saednia, K. *et al.* Quantitative digital histopathology and machine learning to predict pathological complete response to chemotherapy in breast cancer patients using pre-treatment tumor biopsies. *Sci. Rep.* **12**, 9690 (2022).
47. Lu, H. *et al.* A mathematical-descriptor of tumor-mesoscopic-structure from computed-tomography images annotates prognostic- and molecular-phenotypes of epithelial ovarian cancer. *Nat. Commun.* **10**, 764 (2019).
48. Unlocking the potential of AI in drug discovery | Reports | Wellcome. https://wellcome.org/reports/unlocking-potential-ai-drug-discovery.
49. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
50. Karimi, M., Wu, D., Wang, Z. & Shen, Y. DeepAffinity: interpretable deep learning of compound–protein affinity through unified recurrent and convolutional neural networks. *Bioinformatics* **35**, 3329–3338 (2019).

51. Mayr, A., Klambauer, G., Unterthiner, T. & Hochreiter, S. DeepTox: Toxicity Prediction using Deep Learning. *Front. Environ. Sci.* **3**, (2016).
52. Paul, D. *et al.* Artificial intelligence in drug discovery and development. *Drug Discov. Today* **26**, 80–93 (2021).
53. MacLean, F. Knowledge graphs and their applications in drug discovery. *Expert Opin. Drug Discov.* **16**, 1057–1069 (2021).
54. Himmelstein, D. S. *et al.* Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *eLife* **6**, e26726 (2017).
55. Richardson, P. *et al.* Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. *The Lancet* **395**, e30–e31 (2020).
56. Karekar, S. R. & Vazifdar, A. K. Current status of clinical research using artificial intelligence techniques: A registry-based audit. *Perspect. Clin. Res.* **12**, 48–52 (2021).
57. Kavalci, E. & Hartshorn, A. Improving clinical trial design using interpretable machine learning based prediction of early trial termination. *Sci. Rep.* **13**, 121 (2023).
58. Schwager, E. *et al.* Utilizing machine learning to improve clinical trial design for acute respiratory distress syndrome. *NPJ Digit. Med.* **4**, 133 (2021).
59. Haddad, T. *et al.* Accuracy of an Artificial Intelligence System for Cancer Clinical Trial Eligibility Screening: Retrospective Pilot Study. *JMIR Med. Inform.* **9**, e27767 (2021).
60. Chow, R. *et al.* Use of artificial intelligence for cancer clinical trial enrollment: a systematic review and meta-analysis. *J. Natl. Cancer Inst.* **115**, 365–374 (2023).
61. Vazquez, J. *et al.* Using supervised machine learning classifiers to estimate likelihood of participating in clinical trials of a de-identified version of ResearchMatch. *J. Clin. Transl. Sci.* **5**, e42 (2020).
62. Zhao, Y., Zeng, D., Socinski, M. A. & Kosorok, M. R. Reinforcement Learning Strategies for Clinical Trials in Non-small Cell Lung Cancer. *Biometrics* **67**, 1422–1433 (2011).
63. Labovitz, D. L., Shafner, L., Reyes Gil, M., Virmani, D. & Hanina, A. Using Artificial Intelligence to Reduce the Risk of Nonadherence in Patients on Anticoagulation Therapy. *Stroke* **48**, 1416–1419 (2017).
64. Curth, A. & Schaar, M. van der. Nonparametric Estimation of Heterogeneous Treatment Effects: From Theory to Learning Algorithms. in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics* 1810–1818 (PMLR, 2021).
65. Shah, A. D., Bartlett, J. W., Carpenter, J., Nicholas, O. & Hemingway, H. Comparison of random forest and parametric imputation models for imputing missing data using MICE: a CALIBER study. *Am. J. Epidemiol.* **179**, 764–774 (2014).
66. Damiani, G. *et al.* Potentiality of algorithms and artificial intelligence adoption to improve medication management in primary care: a systematic review. *BMJ Open* **13**, e065301 (2023).
67. Rozenblum, R. *et al.* Using a Machine Learning System to Identify and Prevent Medication Prescribing Errors: A Clinical and Cost Analysis Evaluation. *Jt. Comm. J. Qual. Patient Saf.* **46**, 3–10 (2020).

68. Liu, F. & Panagiotakos, D. Real-world data: a brief review of the methods, applications, challenges and opportunities. *BMC Med. Res. Methodol.* **22**, 287 (2022).
69. MHRA guidance on the use of real-world data in clinical studies to support regulatory decisions. *GOV.UK* https://www.gov.uk/government/publications/mhra-guidance-on-the-use-of-real-world-data-in-clinical-studies-to-support-regulatory-decisions.
70. Sendor, R. & Stürmer, T. Core Concepts in Pharmacoepidemiology: Confounding by Indication and the Role of Active Comparators. *Pharmacoepidemiol. Drug Saf.* **31**, 261–269 (2022).
71. Lee, B. K., Lessler, J. & Stuart, E. A. Improving propensity score weighting using machine learning. *Stat. Med.* **29**, 337–346 (2010).
72. Ahmad, T. *et al.* Machine Learning Methods Improve Prognostication, Identify Clinically Distinct Phenotypes, and Detect Heterogeneity in Response to Therapy in a Large Cohort of Heart Failure Patients. *J. Am. Heart Assoc.* **7**, e008081.
73. Lai, A. G. *et al.* An informatics consult approach for generating clinical evidence for treatment decisions. *BMC Med. Inform. Decis. Mak.* **21**, 281 (2021).
74. Wang, Z., Ong, C. L. J. & Fu, Z. AI Models to Assist Vancomycin Dosage Titration. *Front. Pharmacol.* **13**, (2022).
75. Barbieri, C. *et al.* An international observational study suggests that artificial intelligence for clinical decision support optimizes anemia management in hemodialysis patients. *Kidney Int.* **90**, 422–429 (2016).
76. Destere, A., Marquet, P., Labriffe, M., Drici, M.-D. & Woillard, J.-B. A Hybrid Algorithm Combining Population Pharmacokinetic and Machine Learning for Isavuconazole Exposure Prediction. *Pharm. Res.* **40**, 951–959 (2023).
77. Poweleit, E. A., Vinks, A. A. & Mizuno, T. Artificial Intelligence and Machine Learning Approaches to Facilitate Therapeutic Drug Management and Model-Informed Precision Dosing. *Ther. Drug Monit.* **45**, 143–150 (2023).
78. Sanghvi, S. *et al.* How to assess pharmacogenomic tests for implementation in the NHS in England. *Br. J. Clin. Pharmacol.* **n/a**,.
79. A unifying model to predict variable drug response for personalised medicine | bioRxiv. https://www.biorxiv.org/content/10.1101/2020.03.02.967554v2.abstract.
80. McInnes, G. *et al.* Transfer learning enables prediction of CYP2D6 haplotype function. *PLoS Comput. Biol.* **16**, e1008399 (2020).
81. DeepADEMiner: A Deep Learning Pharmacovigilance Pipeline for Extraction and Normalization of Adverse Drug Effect Mentions on Twitter | medRxiv. https://www.medrxiv.org/content/10.1101/2020.12.15.20248229v1.full.
82. Freedom of Information request on the use of artificial intelligence (AI) in our Yellow Card system, when it went live and the impact it has had on reporting and concomitant investigations. (FOI 22/494). *GOV.UK* https://www.gov.uk/government/publications/freedom-of-information-responses-from-the-mhra-

week-commencing-14-march-2022/freedom-of-information-request-on-the-use-of-artificial-intelligence-ai-in-our-yellow-card-system-when-it-went-live-and-the-impact-it-has-had-on-re.
83. Cherkas, Y., Ide, J. & van Stekelenborg, J. Leveraging Machine Learning to Facilitate Individual Case Causality Assessment of Adverse Drug Reactions. *Drug Saf.* **45**, 571–582 (2022).
84. European Parliament. Directorate General for Parliamentary Research Services. *Artificial intelligence in healthcare: applications, risks, and ethical and societal impacts.* (Publications Office, 2022).
85. NHANES Questionnaires, Datasets, and Related Documentation. https://wwwn.cdc.gov/nchs/nhanes/.
86. Teno, J. M. Garbage in, Garbage out—Words of Caution on Big Data and Machine Learning in Medical Practice. *JAMA Health Forum* **4**, e230397 (2023).
87. Panch, T., Mattie, H. & Atun, R. Artificial intelligence and algorithmic bias: implications for health systems. *J. Glob. Health* **9**, 020318.
88. Vokinger, K. N., Feuerriegel, S. & Kesselheim, A. S. Mitigating bias in machine learning for medicine. *Commun. Med.* **1**, 1–3 (2021).
89. Narla, A., Kuprel, B., Sarin, K., Novoa, R. & Ko, J. Automated Classification of Skin Lesions: From Pixels to Practice. *J. Invest. Dermatol.* **138**, 2108–2110 (2018).
90. Wilson, F. P. *et al.* Electronic health record alerts for acute kidney injury: multicenter, randomized clinical trial. *BMJ* **372**, m4786 (2021).
91. Patel, V. L. & Kannampallil, T. G. Cognitive informatics in biomedicine and healthcare. *J. Biomed. Inform.* **53**, 3–14 (2015).
92. Bryant, A. D., Fletcher, G. S. & Payne, T. H. Drug interaction alert override rates in the Meaningful Use era: no evidence of progress. *Appl. Clin. Inform.* **5**, 802–813 (2014).
93. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G. & King, D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* **17**, 195 (2019).
94. Mundhenk, T. N., Chen, B. Y. & Friedland, G. Efficient Saliency Maps for Explainable AI. Preprint at https://doi.org/10.48550/arXiv.1911.11293 (2020).
95. Software and AI as a Medical Device Change Programme - Roadmap. *GOV.UK* https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme-roadmap.
96. Health, C. for D. and R. Artificial Intelligence and Machine Learning in Software as a Medical Device. *FDA* (2022).
97. The Medical Devices Regulations 2002. https://www.legislation.gov.uk/uksi/2002/618/contents/made.
98. *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS*. (2021).

99. PM urges tech leaders to grasp generational opportunities and challenges of AI. *GOV.UK* https://www.gov.uk/government/news/pm-urges-tech-leaders-to-grasp-generational-opportunities-and-challenges-of-ai.
100. Locating Liability for Medical AI by W. Nicholson Price II, I. Glenn Cohen :: SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4517740.
101. Scott, I. A., Carter, S. M. & Coiera, E. Exploring stakeholder attitudes towards AI in clinical practice. *BMJ Health Care Inform.* **28**, e100450 (2021).
102. Finlayson, S. G. *et al.* The Clinician and Dataset Shift in Artificial Intelligence. *N. Engl. J. Med.* **385**, 283–286 (2021).
103. Oakden-Rayner, L., Dunnmon, J., Carneiro, G. & Ré, C. Hidden Stratification Causes Clinically Meaningful Failures in Machine Learning for Medical Imaging. Preprint at https://doi.org/10.48550/arXiv.1909.12475 (2019).
104. Veinot, T. C., Mitchell, H. & Ancker, J. S. Good intentions are not enough: how informatics interventions can worsen inequality. *J. Am. Med. Inform. Assoc.* **25**, 1080–1088 (2018).

Figure 1: Potential for AI in Clinical Pharmacology and Therapeutics

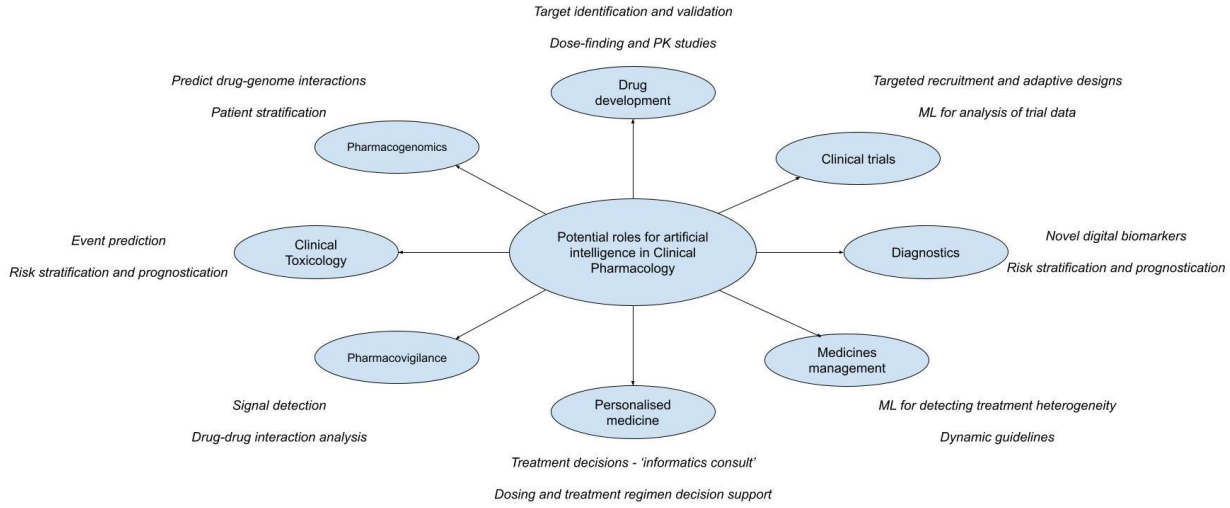

Figure 1: *Potential roles for artificial intelligence and machine learning in Clinical Pharmacology. ML: machine learning, PK: pharmacokinetics.*

Tables:

Table 1: Model Development: A Clinician's perspective

A supervised learning binary classification model was built using data from 35,000 participants in the cross-sectional US National Health and Nutrition Examination Survey (NHANES 2007 – 2016). This model was trained to classify participants based on measures of dysglycaemia (composite of either abnormal glucose tolerance test or elevated glycated haemoglobin A1c ≥ 48 mmol/mol). The purpose of this model development is to highlight key considerations in AI development and is not a clinical decision aid.
1. **Carefully define the use-case:** An AI model should have an ethical and well-defined purpose, for example, aid clinical decisions, promote patient safety or reduce healthcare inefficiencies. In this example, one could envisage such a model being useful for a medical admissions unit, where prediction of dysglycemia might trigger a formal test for diabetes or other health intervention. Decisions to develop such a model need to be cognisant of how the model would interact with wider systems (for example, local diabetes services), resource availability and patient preferences.
2. **Data quality:** Clinical AI models should utilise standard features that would be readily available in real-world settings. Developers need to be aware of what data is available in real-time, any biases in the underlying data and consider how missing data will be handled. In this example, the model was developed using publicly available data from a cross-sectional survey of the US population, with little missing data. As a result, this model may not apply to other contexts (acutely unwell patients) or populations.
3. **Data organisation:** The available data is commonly divided at random into a training and test data set prior to model development. The training data allows the algorithm to learn the relationship between the input and the output. Once the final model is determined, it is deployed to the test data to allow the assessment of model performance on unseen data. Other strategies such as "cross-validation" involve repeatedly partitioning data into different training and test sets and averaging the results for each partition.
4. **Feature selection**: Decisions relating to what features to include in the model can be driven by domain expertise, data availability or alternatively can be data-driven, relating to how important certain variables are with the outcome variable. In this example, clinical knowledge was used to select features relating to commonly recorded demographic, routine laboratory tests, measurements such as blood pressure, BMI and prescribed medication.
5. **Model development:** Various binary classification models were tested but our focus was on tree-based methods including random forests and XGBoost models. Hyperparameters (characteristics of the model) were optimised during training by randomly searching for best settings in an iterative manner. Models were compared using several evaluation metrics on the test set. Methods used included: confusion matrix, sensitivity, specificity, positive predictive values and area-under-receiver operating curve for each model. Other factors such as speed of prediction may be important in other medical contexts.
6. **External validation:** One common error in AI model development is the concept of overfitting, where the model in development "rote-learns" the training data and fails to learn broader patterns in the data. As a result, it does not generalise to new data or data originating from other contexts.In addition, models can be sensitive to minor changes in data acquisition, demographics or changes in clinical guidelines or practices. It is essential to evaluate the model performance in different populations, contexts and over-time. There may be a need to update model parameters for a new context, in a process known as transfer learning. In this example, the model would need to be tested in clinical settings before an accurate evaluation of the model could be made.
7. **Deployment and clinical evaluation:** Randomised clinical trials represent the gold standard to ensure that an AI tool performs correctly in the real-world and has no unintended or adverse consequences. Reporting guidelines have been adapted for AI based tools, including the CONSORT-AI[21] and SPIRIT-AI[20] statements. Any adverse or potentially adverse events should be reported via the UK Yellow Card Scheme.

Table 1: *A clinician's guide to AI model development. Data and coding notebooks are publicly available at: https://github.com/dkdryan/ai_cpt*

Table 2: Anatomy of machine learning models

Regression-based models:
Linear regression: model that finds the straight line that best represents the relationship between a continuous outcome variable and one or more predictor variables by minimising the overall distance between the observed and predicted values. Use case: studying the effect of variables on a continuous outcome, commonly used for prediction purposes.

Logistic regression: model used to estimate the probability of a binary outcome based on predictor variables. It relates various predictors to the probability of the outcome and constrains the probability to be between 0 and 1.

Both linear and logistic regression have strengths in terms of their interpretability and familiarity to most clinicians. However, weaknesses relate to the requirement for the developer to correctly specify how the variables enter the model (e.g. should age be entered as a squared term?) and ensure assumptions hold for the model.

Decision trees models:
Models that classify based on separating individuals according to different values of a feature. The model seeks to optimise the splitting of individuals to maximise information gain at each level of the tree.

Random forests: an ensemble model which combines multiple weak decision trees that are used to classify or make predictions.

Key strengths of tree-based methods relate to their interpretability and intuition. In addition, model developers do not need to define how variables enter the model and the model can handle complex, non-linear associations between variables. Disadvantages can relate to the model's tendency to over-fit - "rote-learn" the training data and fail to learn underlying patterns in data, resulting in the model not generalising to unseen data.

K-nearest neighbours (KNN):
Data is organised in a two-dimensional (or higher) space. Model makes a classification or prediction for a new data point based on the value or label of the data in the vicinity. The value or label of a data point's 'nearest neighbours' informs prediction. KNN tends to work best in the context of low-dimensionality data.

Neural networks:
Models that learn patterns and relationships in data, inspired by design of the human neuron. Data is transformed through different layers. The number of layers as well as the relative importance and data transformation of each layer is determined based on training data and updated to reduce the error in the system. In this way, the neural network learns how to map the input to the output. Deep learning refers to a neural network with many hidden layers.

Generative AI model: is a model that can create new content, such as images or text, based on learning patterns from millions of examples. A common format is a generative adversarial network (GAN) where a generator model synthesises new content and tries to beat a classification model that aims to detect if it is human or computer generated. The models compete against each other, with the generator model learning over time how to outwit the classification model.

Large language models: is a specific form of generative AI models which generates text based on learning patterns from a large corpus (bank) of text. A notable example is ChatGPT - chat generative pre-trained transformer model.

Neural network models have shown remarkable performance in recent years. However, key disadvantages can relate to lack of interpretability, requirement for large training datasets to train models and significant computational expense. In addition, there are growing concerns around the environmental cost of training such large models.

Table 2: *Description of common AI models, detailing model structure, advantages and disadvantages*

Table 3: Potential pitfalls in the development of medical AI interventions

Potential pitfalls in the development of medical AI interventions
Model development ● Biased datasets fail to represent real-world patient cohorts ● Missing data introduces bias into the model ● Data leakage - where the model learns spurious correlations ● Poorly labelled data **Model evaluation** ● Lack of external validation ● Lack of interpretability ● Hidden stratification - model performs poorly in certain subgroups ● Lack of prospective RCT evidence **Model deployment** ● Model does not fit within clinical pathways or has limited real world value ● Limited consideration of how model interacts with human processes ● Drift - model performance can change over time relating to other changes in healthcare environment ● Cost effectiveness assessment ● AI-vigilance - analysis of model errors

Table 3: *potential pitfalls in the lifecycle of an AI model, from model development to final deployment.*