

Aberystwyth University

Current state and prospects of artificial intelligence in allergy

van Breugel, Merlijn; Fehrmann, Rudolf S. N.; Bügel, Marnix; Rezwan, Faisal I.; Holloway, John W.; Nawijn, Martijn C.; Fontanella, Sara; Custovic, Adnan; Koppelman, Gerard H.

Published in:
Allergy

DOI:
[10.1111/all.15849](https://doi.org/10.1111/all.15849)

Publication date:
2023

Citation for published version (APA):

van Breugel, M., Fehrmann, R. S. N., Bügel, M., Rezwan, F. I., Holloway, J. W., Nawijn, M. C., Fontanella, S., Custovic, A., & Koppelman, G. H. (2023). Current state and prospects of artificial intelligence in allergy. *Allergy*, 78(10), 2623-2643. <https://doi.org/10.1111/all.15849>

Document License CC BY-NC

General rights

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal




Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

REVIEW

Current state and prospects of artificial intelligence in allergy

Merlijn van Breugel^{1,2,3}  | Rudolf S. N. Fehrmann⁴  | Marnix Bügel³ | Faisal I. Rezwan^{5,6} | John W. Holloway^{5,7}  | Martijn C. Nawijn^{2,8} | Sara Fontanella^{9,10} | Adnan Custovic^{9,10} | Gerard H. Koppelman^{1,2}

¹Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children's Hospital, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

²Groningen Research Institute for Asthma and COPD (GRIAC), University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

³MIcompany, Amsterdam, the Netherlands

⁴Department of Medical Oncology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

⁵Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK

⁶Department of Computer Science, Aberystwyth University, Aberystwyth, UK

⁷National Institute for Health and Care Research Southampton Biomedical Research Centre, University Hospitals Southampton NHS Foundation Trust, Southampton, UK

⁸Department of Pathology and Medical Biology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

⁹National Heart and Lung Institute, Imperial College London, London, UK

¹⁰National Institute for Health and Care Research Imperial Biomedical Research Centre (BRC), London, UK

Correspondence

Gerard H. Koppelman, Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children's Hospital, University Medical Center Groningen, PO Box 30.001, Groningen, 9700 RB, the Netherlands.

Email: g.h.koppelman@umcg.nl

Abstract

The field of medicine is witnessing an exponential growth of interest in artificial intelligence (AI), which enables new research questions and the analysis of larger and new types of data. Nevertheless, applications that go beyond proof of concepts and deliver clinical value remain rare, especially in the field of allergy. This narrative review provides a fundamental understanding of the core concepts of AI and critically discusses its limitations and open challenges, such as data availability and bias, along with potential directions to surmount them. We provide a conceptual framework to structure AI applications within this field and discuss forefront case examples. Most of these applications of AI and machine learning in allergy concern supervised learning and unsupervised clustering, with a strong emphasis on diagnosis and subtyping. A perspective is shared on guidelines for good AI practice to guide readers in applying it effectively and safely, along with prospects of field advancement and initiatives to increase clinical impact. We anticipate that AI can further deepen our knowledge of disease mechanisms and contribute to precision medicine in allergy.

KEYWORDS

artificial intelligence, deep learning, diagnosis, machine learning, precision medicine

Abbreviations: AI, artificial intelligence; CE, Conformité Européenne; EHR, electronic health records; FDA, US Food and Drug Administration; GIGO, garbage in garbage out; ML, machine learning; RCT, randomized control trial.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *Allergy* published by European Academy of Allergy and Clinical Immunology and John Wiley & Sons Ltd.

1 | INTRODUCTION

As of February 2023, the US Food & Drug Administration (FDA) has approved 521 medical applications that utilize artificial intelligence (AI) and machine learning (ML).¹ Most of these (75%) are in radiology, followed by cardiology, hematology, and neurology. Similar trends are observed in Conformité Européenne (CE)-marked medical devices incorporating AI within the European Union.² Currently, no registered AI and ML-based applications are being utilized in the field of allergy. One can therefore question if this field is missing out on new research opportunities and clinical applications either because of insufficient access to AI applications or a lack of awareness of potential applications. However, given the rapid pace of technological advances, it can be anticipated that AI and ML algorithms will be increasingly employed in allergy research and applications soon.

Over the past decade, medicine has witnessed an exponential growth of interest in AI and the yearly number of scientific articles on AI has increased tenfold since 2012.³ This trend is fueled by the explosion of (bio)medical data, including multi-omics, image data, and digital electronic health records (EHRs), along with advancements in computing power. These developments have paved the way for advanced analytical approaches to address new research questions on large-scale datasets. Traditional analytical techniques are no longer adequate to handle such data complexity, volume, and structure. The introduction of accessible software and methodological advancements in AI have further promoted the use of AI in the (bio-)medical field. Most importantly, ML and AI can identify complex patterns in vast amounts of data, such as images, text, or audio and deliver superior predictive power, often surpassing traditional statistical methods.⁴⁻⁶

This review provides a fundamental understanding of ML and AI's core concepts. A framework is presented to structure the broad umbrella term AI, and an overview of several state-of-the-art applications of AI in medicine and allergy, specifically, is provided. The focus is on applications that preferably adhere to any, and ideally multiple, of the following conditions: (1) are externally or prospectively validated, (2) demonstrate a positive effect on clinically relevant patient outcomes, (3) FDA and/or CE approval, (4) outperform traditional methods, and (5) answer research questions where traditional analytical techniques fail. Additionally, we critically discuss the limitations and open challenges of AI applications and share an outlook on good practices of AI and ML in allergy.

2 | THE FUNDAMENTALS AND TERMINOLOGY OF ARTIFICIAL INTELLIGENCE

2.1 | Machine learning and deep learning

AI is the discipline in computer science that develops computer systems that can simulate human intelligence and perform tasks that generally require humans (Table 1). This discipline can be further

narrowed down into ML and its subdiscipline *deep learning* (Figure 1). ML can be described as an algorithm that learns from data by automatically mapping input data to an output prediction. While this draws a parallel to traditional regression methods such as ordinary least squares, most machine learning techniques have the advantage of inherently modeling complex patterns, including non-linearity and interaction effects. ML concentrates primarily on prediction and finding patterns in vast amounts of data without making prior assumptions about the distribution of these data.⁷ The predictive performance generally improves when more (high-quality) data are fed to the algorithm.

Over the past years, the ML subfield of *deep learning* has gained tremendous popularity, as it has yielded superior results in analyzing *unstructured data* such as medical images,⁸ text data, and audio data. This technique is based on large artificial neural networks (ANNs). ANNs form networks inspired by the biological animal brain, consisting of multiple layers of processing units called neurons. Deep learning methods can detect complex data relationships by automatically compressing data and distilling relevant features in various levels of abstraction.⁹ This makes it different from statistical approaches such as regression methods, which require explicitly defining independent variables and making assumptions about their relationship to the outcome variable. Another advantage of deep learning is its ability to continue learning and improve performance with larger datasets.¹⁰ Besides applications in computer vision, which is the ability to interpret image data by an AI system, deep learning has also propelled natural language processing (NLP) forward, which is the capacity of a computer to understand written and spoken human language.¹¹ We refer to recent, extensive reviews that cover the subfields of deep learning and its applications in medicine.¹²⁻¹⁴ A state-of-the-art example is the detection of diabetic retinopathy from retinal images,¹⁵ for which the IDx-DR deep-learning-based software has been FDA-approved¹⁶ and validated in a clinical setting.¹⁷ Relatedly, deep learning approaches have outperformed trained physicians in breast cancer detection using imaging data,^{18,19} with currently nine applications FDA-approved.²⁰ Some of the latest AI breakthroughs that ignited the general public's interest are based on deep learning approaches. These involve generative models that are trained to create new data. Examples include deep fakes, DALL-E²¹ (an OpenAI application that creates figures and art based on written descriptions), and most recently, ChatGPT,²² an AI tool that generates highly realistic written text based on user prompts. Figure S1 displays two AI-generated illustrations of this review's topic.

2.2 | Learning strategies

A useful categorization of AI is made on the learning strategy, which defines how an algorithm learns from data.²³ Three different approaches are distinguished: supervised, unsupervised, and reinforcement learning. We provide a conceptual framework to structure AI applications based on learning strategy, learning goal, data modality, and medical domain in Figure 2.

TABLE 1 Glossary of key terms in the AI field. This table contains important and popular terms and methods, with a concise definition. The selection covers those terms most relevant in the context of this article; a general or exhaustive table is outside the scope of this work.

| Term | Definition |
|--------------------------------------|--|
| Artificial neural network (ANN) | Technique that is build up of a network of interconnected nodes (neurons) that process signals over weighted arcs. Central to deep learning with a wide range of complex variations |
| Chatbot | Software designed to mimic human-like conversation and enable interactive communication with a user via text or voice. Newer chatbots leverage AI-driven NLP, whereas earlier version were often predefined rule-based systems |
| Convolutional neural network (CNN) | Class of neural networks that is characterized by convolution filters that slide over input data to extract relevant patterns. Frequently used in in deep learning-based image analysis |
| Decision tree | Among the most popular ML algorithms that learns to split data on certain conditions of variables to classify or predict an outcome variable, creating an hierarchical tree shape with predictions in leaf nodes |
| Generative adversarial network (GAN) | Class of deep neural networks for the generation of new data samples. GANs have formed a rapidly advancing field since the 2016 introduction, used in applications such as deepfakes |
| Gradient boosting | Machine learning model type that uses an ensemble of weak prediction models (often decision trees), optimized over a differentiable loss function. XGBoost and LightGBM are popular algorithms in this family |
| k-means | Unsupervised clustering method that aims to partition observations into k clusters, where each observation is mapped to the closest cluster centroid |
| LightGBM | Popular ML algorithm of relatively recent origin (2016), similar performance to XGBoost but with more efficient training due to improved decision tree splitting strategy |
| Natural language processing (NLP) | The discipline in AI involved in the understanding of written and spoken human language |
| Overfitting | A model that captures the training data too closely, hereby hindering generalization and prediction on future data |
| Principal component analysis (PCA) | Dimensionality reduction technique that uses linear transformation to map data to a lower dimension than the initial data |
| Random forest | Popular ML algorithm that builds an ensemble of decision trees, improving on the performance and generalizability of single decision trees |
| Support vector machine (SVM) | Supervised model that aims to find the optimal hyperplane that best separates different categories of observations |
| Tabular data | Data that are organized in a table with rows and column |
| Transfer learning | Technique to improve model learning by leveraging knowledge gained on a related problem. Often used for recalibrating large-scale pre-trained deep learning models |
| Unstructured data | Data that have an internal structure but one that is not represented in a row-column table, such as image, text, and audio |
| XGBoost | Popular ML algorithm that uses gradient boosting and builds decision trees iteratively, often delivering best-of-class performance and fast model training |

2.2.1 | Supervised machine learning

Most machine learning applications concern *supervised learning*, where a model is trained to predict a known outcome, called the target variable, label or dependent variable. Supervised learning often requires manual labeling of the target variable. Supervised learning can be applied in almost all medical domains, such as disease diagnosis, treatment outcome prediction, or classifications in medical imaging. One such example is the FDA-approved *Koios DS for Breast* application.²⁴ This tool supports clinicians in breast cancer diagnosis by classifying ultrasound images into benign, probably benign, suspicious, and probably malignant. It has been shown to improve assessment performance compared to the clinician's assessment in a retrospective study.²⁵ Supervised machine learning has also surged in screening, predicting, contact and tracing, and drug

development.^{26,27} For example, during the COVID-19 pandemic, supervised ML was used to predict which potential drug compounds could be effective against SARS-CoV-2 targets by developing prediction models for the drug-likeness of candidate compounds from chemical libraries based on chemical descriptors.^{28,29} Within supervised ML, gradient-boosted decision tree methods have been among the most popular and performant, with the leading algorithms being Random Forest, XGBoost, and LightGBM (Glossary).³⁰⁻³²

2.2.2 | Unsupervised machine learning

In *unsupervised learning*, the aim is to learn groupings in data or reduce their dimensionality. Contrary to its supervised counterpart, there are no known labels to predict. Unsupervised learning is often

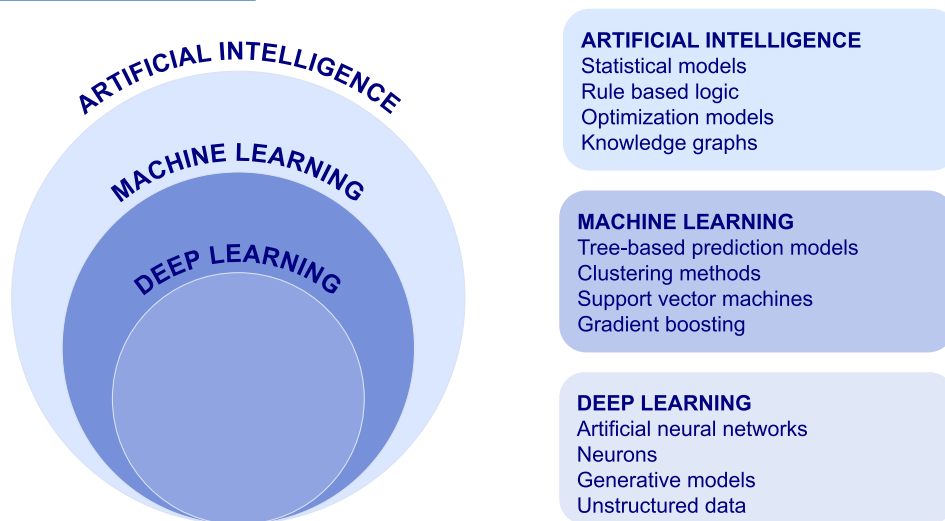


FIGURE 1 Hierarchy of disciplines of artificial intelligence. The discipline of AI is often categorized as part of computer science, although it also builds upon other fields, such as mathematics and cognitive sciences. ML is a subdiscipline of AI, whereas deep learning is a further specialization within ML, generally characterized by large-scale artificial neural networks consisting of many layers, hence the term *deep*. The right part of the figure displays typical terms that one encounters within the (sub)discipline.

used for clustering analysis. Here, the algorithm aims to describe the data in a limited number of clusters or groups, where goodness-of-fit tests determine the most parsimonious model. An example³⁰ is the discovery of asthma phenotypes based on longitudinal wheezing patterns or clinical variables.^{33–35} Techniques for unsupervised learning are latent class analysis (LCA), k-means clustering, principal component analysis (PCA), and Multidimensional Scaling (MDS).^{36–39} Recently, also *semi-supervised learning* has grown in popularity, which aims to overcome the lack of sufficiently large, labeled datasets⁴⁰ and the tedious task of manual labeling. It leverages a dataset of yet unlabeled data to improve the performance of a model that is initially trained on labeled data.

2.2.3 | Reinforcement learning

Reinforcement learning has recently delivered breakthroughs in the biomedical field.⁴¹ This strategy is characterized by an iterative process that aims to take actions that deliver maximum reward based on a defined objective function. This is comparable to nature, where animals have learned to interpret signs such as hunger as negative, whereas satiety after food intake is seen as positive reinforcement. When animals learn how to behave to gain optimal positive reinforcement, they show reinforcement learning. Applications within medicine have indicated its potential.⁴¹ For example, an “AI Clinician” algorithm has been developed to improve the treatment of sepsis by suggesting the personalized treatment of intravenous fluids and vasopressors.⁴² While still requiring prospective validation, an independent validation cohort was used to assess this algorithm, showing that mortality was lowest when clinicians' treatment policy was close to AI recommended policy and higher when deviating from it.

3 | CHALLENGES AND PITFALLS FOR AI APPLICATION IN MEDICINE

Artificial Intelligence is not without pitfalls, and serious challenges must be overcome to deliver its full potential. The most critical challenges are described below, with potential directions to surmount them. For a more in-depth discussion of AI's current most pressing issues, the reader is referred to several excellent reviews.^{8,43–45}

3.1 | Data

AI systems and models are as good as the data they learn from. This relates to the data's (1) quality and quantity, (2) suitability, and (3) availability. The first challenge refers to data quality and quantity. Low quality of input data, leading to biased outcomes, is often referred to as the GIGO (“garbage in garbage out”) principle. Data quantity also remains challenging, since AI models are extremely “data hungry,” especially for deep learning methods.⁴⁶ The availability and quality of data labels are critical, as label inaccuracies directly impair model reliability.⁴⁷ In particular for images, manual labeling of images is time-consuming. Combining and harmonizing multiple datasets is increasingly used to overcome these data limitations. The use of synthetic data may also help, where additional data are generated by simulating from a known data distribution, which has been shown to improve model performance.⁴⁸ Similarly, in image analysis, data augmentation is often used to (fictively) increase the data sample size by applying data transformations on existing (non-synthetic data points).⁴⁹ Another strategy to improve model reliability on relatively small datasets is *transfer learning*, especially popular in NLP and image analysis. This technique enables researchers to train a complex model on relatively small datasets by recalibrating existing parameters of known models.⁵⁰

| Learning strategy | Learning goal | Examples | Allergy references |
|--------------------------|--|---|--|
| Supervised learning | Classification | Prostate cancer classification in biopsy WSIs Pathologists improved their performance by AI assistance that classified biopsies and pinpointed suspicious areas | Cardiovascular disease classification on heart sound recordings The CardioXNet algorithm classifies phonocardiograms into five classes of cardiac auscultation, with high accuracy on challenge datasets |
| | Regression | Joint space width measure in osteoarthritis knee radiographs The FDA-approved clinical support tool uses deep learning to determine the minimum joint space width, while also detecting osteoarthritis | Chemotherapy dose recommendation with CURATE-AI The AI platform predicts the optimal dosing in a combination regimen, leading to increased efficacy and patient safety. |
| | Clustering | Identification of wheezing phenotypes in asthmatic children Latent class analysis of birth cohort data identified wheezing types, consistent with another cohort's study, while also detecting a new type | Dietary pattern analysis using principal component analysis PCA identified five dietary patterns based on questionnaire and health data, some with significantly increased risk for Crohn's disease and ulcerative colitis |
| | Dimensionality reduction | Multi-omics dimension reduction for ovarian cancer analysis DL is used to integrate and compress genomics, transcriptomics and epigenomics into latent representations for improved downstream analysis | Text analysis on electronic health records data Using DL to create low-dimensional latent vectors from large-scale and diverse text data, which can be used for further clustering to find disease subtypes |
| | Generation | Accelerated CS-MRI reconstruction using generative AI RefineGAN uses generative adversarial networks for faster and more accurate reconstruction of CS-MRI, outperforming current benchmarks | Histological cancer image generation for data size increase Synthetic images were visually indistinguishable from real images, assessed by trained pathologists, while also improving a cancer classification model |
| Semi-supervised learning | Ventricular hypertrophy detection based on partly labelled ECGs A semi-supervised GAN architecture uses mostly unlabelled and a small portion of labelled training data, delivering still 92% detection accuracy | Semi-supervised cancer detection on pathological images Semi-supervised learning achieves similar diagnostic performance to fully labeled datasets on multiple cancer types, reducing costly annotations | |
| Reinforcement learning | Robot-assisted guidewire navigation for coronary intervention The autonomous guidewire navigation can reach all target locations in arteries, reaching over 98% accuracy in 2D and 3D setting | Clinical support for adaptive radiotherapy Use individual patient dose response to improve clinical decision-making, showing 10% improvement potential compared to unaided clinical practice. | |

FIGURE 2 A conceptual framework for AI applications in the biomedical domain. The framework is structured by learning strategy, learning goal and data modality. The included studies are selected as illustrations of how AI is used within the medical field and how its applications can be conceptually categorized. They are not necessarily selected based on the inclusion criteria stated in the introduction. References for the Example studies are included in the Supplementary References. The "Allergy references" column maps the cited studies in the *Current state of AI in the allergy research field* section to their primary learning goal.

Data modality

 Tabular
  Time series
  Image
  Audio
  Text

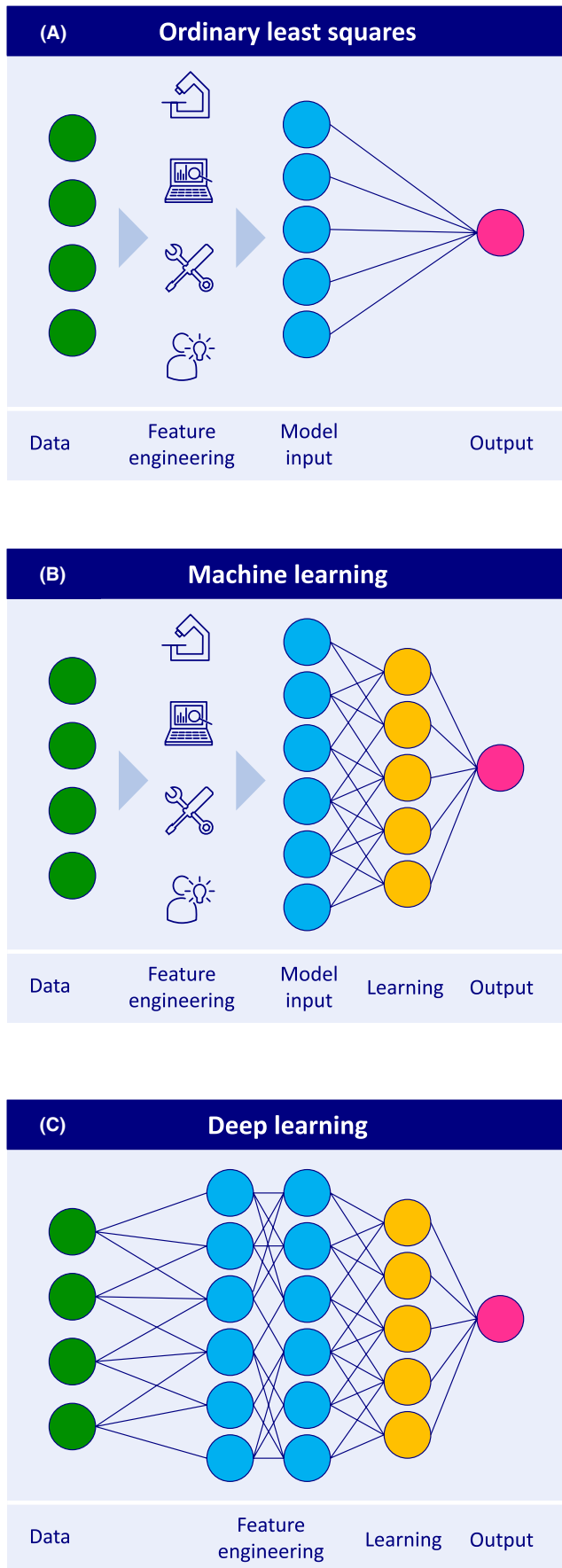


FIGURE 3 Difference between ordinary least squares (OLS), machine learning (ML), and deep learning (DL) methodology. (A) In OLS, features (or predictors) are modeled manually, and their relationship is assumed linear to the output variable unless specified differently. Interpretation of the model and learned patterns (inference) is straightforward. (B) A similar procedure is followed with ML, but the algorithm can learn more complex patterns from the provided features. Nevertheless, thorough feature engineering by the practitioner is a critical step for delivering a performant model. (C) With DL, especially when applied to unstructured data, feature engineering is an inherent behavior of the interconnected neural network layers. The relationship between input features (tabular data fields, image pixels, text snippets, etc.) to the predicted output is more opaque and harder to interpret.

Data suitability poses a second challenge. Akin to traditional analytical methods, AI approaches need adequate study designs to yield reliable outcomes, from data collection to the appropriate analytical strategy. Training algorithms based on unsuitable data may lead to biased outcomes.⁴⁴ For example, it is increasingly clear that AI and ML algorithms can engrain ethnic bias when models are trained using data sets with imbalanced genetic ancestry.

Data availability may pose a third challenge, as data are often siloed within individual institutions, and curated, publicly accessible clinical datasets remain rare. The reason for this includes patient privacy, lack of data-sharing infrastructure, and competition among institutions. In immunology research, efforts are being made to break open silos and democratize datasets.⁵¹ Examples include the National Institutes of Health (NIH)-curated resources on open-access COVID-19 data,⁵² or the European Health Data Space for the safe exchange and reuse of health data.⁵³ These developments are aided by novel data sharing and integration approaches, such as federated learning, where a model is centrally trained while the data are kept locally.⁵⁴ Recently Swarm Learning⁵⁵ was introduced, a decentralized machine learning approach that does not require central coordination. The researchers demonstrate that the model outperforms individual sites in disease classification while retaining complete confidentiality.

3.2 | Explainability

The lack of explainability of AI algorithms hampers clinical implementation.⁵⁶ Unlike statistical methods such as regression, which are inherently explainable, the learned patterns of AI models are more complex, and their estimated parameters are not directly interpretable (Figure 3).

Deep learning models suffer from this due to their hidden learning behavior and up to billions of parameters. One recent study made this vividly clear. When researchers trained a model to distinguish COVID-19 patients with pneumonia from those with other respiratory diseases based on chest radiographs, the algorithm

based its prediction on the printed dates on the radiological images; it found a shortcut and classified all patients dated since 2020 as COVID-19 cases.⁵⁷ Thus, there is a growing demand for “white box” approaches, referring to methods and models that are easy to explain and interpret. This need is further amplified when the aim is to bring applications to clinical practice, which has many technical,^{56,58} medical, legal, and ethical dimensions.⁵⁹⁻⁶¹ The urgent need for explainability has accelerated methodological innovations to “open the black box.” Relevant examples are SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), and CAM (Class Activation Maps).⁶²⁻⁶⁴ For example, SHAP was recently implemented to describe the contribution of features selected for inclusion in asthma prediction models.⁶⁵ These analytical methods calculate how each input feature contributes to each prediction, providing detailed insights into the learning patterns of the AI model.

3.3 | Validation and generalizability

A structured modeling process is essential in developing an ML prediction model to create a reliable model and establish confidence in its outcomes. There are many ML algorithms, and it is difficult to tell which will perform best beforehand. This is called the *no-free lunch* theorem,⁶⁶ which emphasizes the need to develop and evaluate ML models iteratively. Thus, multiple ML methods should be applied to the data and their performance compared. Figure 4 depicts the steps to build a supervised learning prediction model for disease risk.⁶⁷⁻⁶⁹ The steps needed for unsupervised learning overlap to a large extent. Skipping or mismanaging these steps poses a risk to model reliability, for example, by not properly separating the training and validation data, which may lead to overfitting of the prediction model.

One of AI's significant benefits is its ability to scale intelligence at an unprecedented pace. In a time in which a clinician could diagnose a single patient, an AI system could analyze unlimited number of patients, at least in theory. However, the same scalability holds for mistakes and faulty diagnoses, and validation is of the utmost importance to prevent the lack of generalizability of AI models.⁷⁰ AI models tend to “overfit” on the training data, which results in a model that works seemingly well on the training population but poorly predicts future or other patient outcomes, especially in high-dimensional models.⁷¹ One example was from IBM's Watson, which recommended unsafe cancer treatments,⁷² because it was trained on a sample size too limited for its dimensionality. For models to be more broadly applicable and generalizable to other populations, diligent validation and replication (in external datasets) are paramount. Unfortunately, this is often insufficient or altogether missing, respectively. Even FDA-approved AI applications fall short in this domain: Only 11/118 FDA applications (up until 2021) reported a validation set of more than 1000 samples, and only 19/118 reported a multi-reader, multicenter validation study.⁷³ Site-specific recalibration or retraining on multiple datasets are solutions to adapt a model

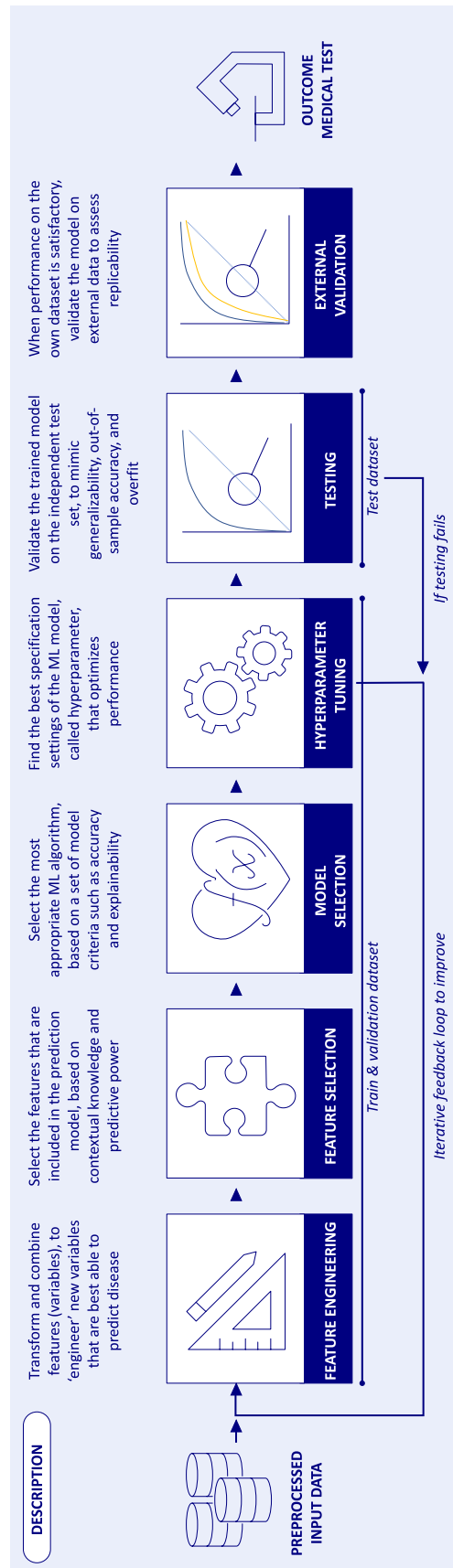


FIGURE 4 Workflow of developing a machine learning model to predict disease risk. The best practice in machine learning modeling is using distinct training, validation (or tuning), and test datasets. The modeling steps till testing are generally executed in sequential order, where it is common to iterate multiple times based on validation results that inform model improvements. It is discouraged to assess and improve test performance iteratively, as this can lead to overfitting. The steps preceding model development are excluded, primarily consisting of problem definition and study design, data collection, and preprocessing.

to another context, although caution is required to avoid spurious learning patterns.⁴⁵

Randomized control trials (RCTs) or prospective validations are scarce in medical AI.^{74,75} Most applications are only tested on retrospective data and have not passed prospective validation in an independent dataset. New guidelines have been emerging for reporting and evaluating RCTs with an AI intervention component in the past 2 years, such as the CONSORT-AI standard^{76,77} and SPIRIT-AI.⁷⁷ A systematic review from 2022 reported that none of the 41 assessed RCTs adhered to this standard and suggested that AI applications with FDA approval do not always prove efficacy.⁷⁸ Thus, the clinical utility and safety remain uncertain, providing a clear direction for future research to confidently implement AI in clinical medicine.

Lastly, the process of validation must not be considered as a one-time event that ceases upon the system's implementation. There is a persistent risk of model drift,⁷⁹ which is a serious concern in the deployment of AI systems. This refers to the degradation of model performance over time, typically caused by alterations in the distribution of input data. It is important to institute a quantitative monitoring system to promptly identify any decline in accuracy, or even better, flag changes in input data before predictions are made. This approach also facilitates the establishment of a feedback loop that can support ongoing performance enhancement by learning from real-live system behavior.⁷⁴

3.4 | Ethical considerations

AI systems often rely on and are trained on confidential personal data, such as health records, imaging, or genomic data. The more voluminous these data become, especially with integrating multiple data sources and unlocking new data sources, the more critical privacy becomes.⁷⁵ The EU's General Data Protection Regulation (GDPR) already provides a "right to explanation" when decisions are based on "automated processing" such as AI.⁸⁰ There is a complicated relationship between privacy and trust. If the mechanisms of algorithms remain hidden for privacy reasons, this could also impede trust in the solution and slow down adoption by patients and clinicians. Furthermore, being over-protective of privacy in data collection, usage, and sharing can also hinder potential patient benefits from using these data to drive AI solutions for novel diagnostic or therapeutic options.^{75,81} Novel approaches are emerging that preserve privacy without slowing down innovation, such as the generation of synthetic data.⁸² Rather than (pseudo)-anonymizing samples, AI-generated synthetic data samples can be used for safe data sharing or even new model development.^{83,84}

While AI systems are not moral agents, their decisions can have ethical consequences. Especially bias and fairness are two key concepts in this context, and various cases of embedded biases exist in developed models.⁸⁵ The 2021 AI action plan from the FDA, warns that biases in healthcare systems, such as ethnic or gender biases, can be inadvertently introduced to algorithms.⁸⁶ This will lead to research conclusions and applications biased toward specific populations while overlooking others. If they are not corrected, this could

further reinforce biases and exacerbate health inequalities experienced by certain underrepresented populations by excluding them from AI-driven medical innovations.⁸⁷⁻⁸⁹ Therefore, researchers need to ensure that the training sample is diverse and represents any future population to which the AI model will be applied.

While the above risks are important, it is essential to realize that humans are not free from implicit biases. For instance, cardiologists are trained to recognize symptoms of coronary artery disease more frequently in men,⁹⁰ resulting in underdiagnosis in women.⁹¹ The advantage of data and algorithms is that biases may be detected, corrected, or prevented. From the study's outset, during the data collection phase, investigators should strive for a representative training dataset that resembles the data distribution the algorithm would encounter once deployed.⁹² Before model development, guidelines have been defined to assess the risk of algorithmic bias, such as the PROBAST tool.⁹³ Likewise, new techniques for the modeling phase are emerging that can help to mitigate bias,⁹⁴ such as adversarial debiasing.⁹⁵ Lastly, dedicated tools have been developed to evaluate the fairness of algorithms along a variety of fairness definitions,⁹⁶ like the open-source Python library AI Fairness 360.⁹⁷

When the above considerations are not managed adequately, an AI system may make mistakes. This raises the intricate question of (moral) accountability,⁹⁸ which becomes increasingly pressing with more clinical applications in place. However, the traditional notion of accountability is problematic in the context of an AI system.^{90,99} It is questionable whether a clinician can be held responsible for such a system's decisions. Furthermore, the system's complexity can make it infeasible for the clinician, and sometimes even the designer, to understand precisely why certain decisions are made.¹⁰⁰ Therefore, we anticipate that the introduction of AI in clinical medicine will first be limited to decision support systems, with the final clinical decision to be made by the caring physician.

3.5 | Clinical implementation

Despite exciting showcases, AI has been criticized for underdelivering tangible clinical impact. Translating solid AI models to effective action remains an open challenge¹⁰¹ and actual clinical use is still nascent.^{102,103} Recently, even with the surge of COVID-19-related AI research, the clinical value of AI applications remained limited.¹⁰⁴ Important challenges for clinical implementation include questionable clinical advantages, inadequate reporting, and adoption and integration in clinical practices.

Developers of algorithms are also urged to be transparent and complete in their reporting to provide a fair view of improving patient care. RISE criteria (Regulatory aspects, Interpretability, Interoperability, Structured Data, and Evidence) can support overcoming major pitfalls in developing AI applications for clinical practice.¹⁰⁵ Recently, the DECIDE-AI guideline has been introduced as a reporting checklist of AI-based (early-stage) clinical evaluation of decision support systems.¹⁰⁶ In addition, clinicians and patients must adapt to working with and trusting new AI systems, and such

behavioral change is notoriously hard.¹⁰¹ There is a need for (better) AI education for clinicians that will need to adapt to new roles and tools to support them in their decision-making. To smoothen this transition, integration into the medical education system has been proposed.^{8,88} The recent American Academy of Asthma, Allergy and Immunology workgroup has underscored a knowledge and an educational gap in the allergy and immunology field.¹⁰² Furthermore, interoperability of AI systems is vital to ensure that they can be integrated with existing clinical and technical workflows across sites and health systems.¹⁰⁷

4 | CURRENT STATE OF AI IN THE ALLERGY RESEARCH FIELD

Artificial Intelligence applications within the allergy field can be broadly categorized into three domains: clinical research, fundamental research, and drug and therapy development (Figure 5). The studies cited in this section are mapped to their respective learning goals in the framework of Figure 2.

Many studies leveraging AI have been published in the research setting. Virtually, all applications concern supervised learning and unsupervised clustering, whereas semi-supervised learning and reinforcement learning are mainly absent. Overviews of the use of ML in asthma and eczema research over the last seven decades have recently been published.^{108,109}

4.1 | Clinical research

4.1.1 | Diagnosis of allergic diseases

The diagnosis or classification of allergic disease has been the area in which AI has been applied most, an exemplary case of supervised learning. ML has used a wide range of data sources to improve allergy or asthma diagnosis: text data from electronic health records (EHRs),

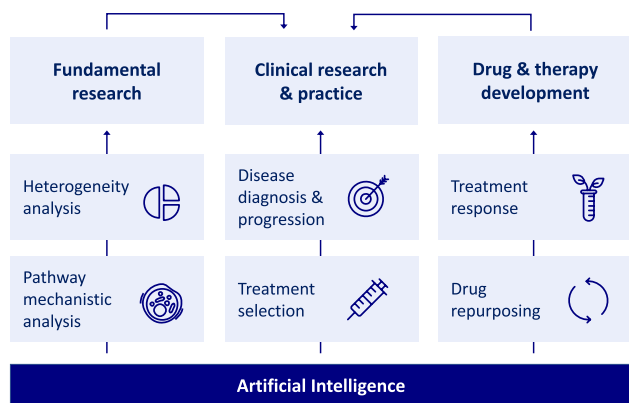


FIGURE 5 Application domains of artificial intelligence within the allergy field. Domains identified as currently most active and discussed in the main text. Other areas, such as clinical trial optimization, have been excluded due to the limited number of impactful applications.

sound data of wheezes, image data from lung CT scans, or large-scale multi-omics data.¹¹⁰⁻¹¹² The extraction of relevant clinical features from EHRs using NLP has successfully diagnosed (childhood) asthma in discovery and replication cohorts.^{111,113,114} In a study of a US birth cohort study, Seol et al.¹¹¹ applied an AI algorithm to define asthma using established predictive and diagnostic criteria in 8196 children. Of all patients that met those criteria, 30% did not have a physician diagnosis of asthma, signifying the potential for early disease identification and population management with EHRs.

Additionally, several studies have investigated the potential of omics data for diagnosis.^{112,115,116} One study developed an ML model that diagnosed IgE-sensitized allergic disease in 16-year-old children based on nasal cell DNA methylation of only three CpG sites.¹¹² External validation in an independent cohort indicated the prospect of reproducible epigenetic tests for diagnosis. Alag¹¹⁶ pursued a similar approach to diagnosing food allergy, where neural networks were trained on blood epigenetic markers. The predictive markers were subsequently associated with a 13-gene profile linked to immune response. This study highlights the potential of novel diagnostic approaches to food allergy.

Machine learning-based modeling of the component-resolved diagnostic multiplex array data has shown that component-specific IgE responses to multiple allergenic proteins are functionally coordinated and co-regulated, and that the networks of interactions are associated with asthma diagnosis and severity.^{117,118} Machine learning has also been used to predict disease risk or persistence. In a prospective study of 704 children aged 2–13 months, unsupervised clustering on 16S rRNA data was used to identify profiles of longitudinal changes in nasal airway microbiota that were significantly associated with asthma risk at age seven.¹¹⁹ These results affirm that the microbiome plays a vital role in the early development of asthma and show promise for early identification and prevention strategies.¹²⁰ In another study, a supervised machine learning model was able to predict asthma persistence in almost 10,000 patients diagnosed before age 5 for persistence by age 10.¹²¹ The XGBoost algorithm delivered the most robust performance (AUC=0.86), using clinically relevant features such as the number of (non) asthma-related visits before age five and noninvasive pulse oximetry data. The study was not independently replicated, which is essential in pursuing clinical support tools. Kothalawala et al.⁶⁵ used data from birth cohorts to train and validate two predictive models, CAPE and CAPP, to predict the likelihood of asthma at school-age using predictors from 0–2 to 0–4 years of age, respectively. Predictive performance was externally validated in the Manchester Asthma and Allergy Study (MAAS) cohort.¹²² Support Vector Machine (SVM) algorithms provided the best performance for both the CAPE (AUC=0.71) and CAPP (AUC=0.82) models, and both demonstrated good generalizability in the replication cohort, performing better than previous regression-based models.

AI guided image analysis has been performed to diagnose eczema.¹²³⁻¹²⁶ One study developed a classifier of atopic dermatitis in multiphoton tomography images, reaching over 97% accuracy¹²⁷ through transfer learning. Highlighted areas of interest in the images could support clinicians in faster diagnosis.

The significant heterogeneity of allergic diseases, combined with the substantial influence of environmental factors, can undermine the generalizability of diagnostic and prognostic models. Notable differences in sensitization patterns across geographical regions enhance this challenge.¹²⁸ This predicament can further amplify the enduring issue of model bias and inequality in medical research—an issue from which the field of allergy is not exempt,¹²⁹ as evident in models demonstrating lesser accuracy in underprivileged populations.¹³⁰ To alleviate this risk, three strategies are recommended. First, ideally, comprehensive data about a patient's environment, inclusive of historical allergen exposure, should be procured and included in the model training procedure. Second, meticulous validation of model performance across diverse subpopulations or disease phenotypes can help identify biases and discrepancies. Lastly, an innovative approach entails the application of continual learning, a concept where operational models are persistently updated and refined based on new observations.¹³¹

4.1.2 | Prediction of asthma exacerbations and hospitalizations

Asthma exacerbations are related to increased morbidity, mortality, and healthcare use, yet these are challenging to predict. Several studies have applied ML to predict exacerbations.^{132–134} In a large study involving EHR data from 60,000 patients, researchers used different ML techniques in a supervised setup to predict three exacerbation outcomes: oral glucocorticoid bursts, ED visits, and hospitalization.¹³⁵ The study achieved a ROC AUC of 0.88 on the latter outcome, which is significantly higher than the results of previous studies (AUC of 0.77); this was replicated in an independent cohort.¹³⁶ Important predictors for hospitalization included oral glucocorticoid burst, inhaled corticosteroid, and blood creatinine. Another study used self-reported daily home monitoring data of asthma symptoms and peak expiratory flow, which were reduced in dimensionality using PCA and then fed to various supervised ML methods.¹³⁷ The best model achieved a sensitivity of 90% and specificity of 83%, predicting severe asthma exacerbations on the same day or up to 3 days in the future. A more extensively validated example is Asthma-Guidance and Prediction System (a-GPS), an AI tool to optimize asthma management.¹³⁸ A-GPS uses NLP on open text from EHRs to provide clinicians with the most relevant clinical information. In a randomized control trial, the tool significantly reduced the time for reviewing EHRs (11.3–3.5 min),¹³⁹ but no significant change in clinical outcome (i.e., exacerbations) was observed. Sensor data from an electronic multi-dose dry powder inhaler (eMDPI), such as inhalation volume and duration, have also been utilized to predict exacerbations with a ROC AUC of 0.83.¹⁴⁰

4.1.3 | Disease management

Medication non-adherence in allergic diseases is common in clinical practice and can negatively impact disease control.^{141,142} To address this issue, researchers explored ML approaches for disease management

and medication adherence.¹⁴³ One such approach involves using ML to provide early warnings for loss of control in the Asthma Mobile Health Study data of 5875 patients,¹⁴⁴ containing over 75,000 daily surveys on symptoms and medicine use, medical history, demographics, location, and EuroQol 5D questionnaire. The supervised classifier obtained an AUC of 0.87, but peak flow readings did not further enhance its performance. External or prospective validation is strongly needed.

In addition to early warning systems, chatbots have also been proposed to support disease management by providing personalized advice to patients¹⁴⁵ and tracking medication compliance. One example is KBot, an early prototype of a chatbot for asthma that utilizes contextual information (such as high pollen triggers) and NLP for dialogue processing.¹⁴⁶ AI can also leverage the capabilities of wearables and mHealth technologies to monitor disease outside clinical contexts.¹⁴⁷ A recent study tested a prototype application for real-time counting of coughs using a deep learning model on ambient sound recorded by mobile phone.¹⁴⁸ This yielded accurate and real-time cough count with a specificity of 92% and a specificity of 98%. Another study applied ML to analyze the sounds of asthma inhalers to predict adequate usage and drug actuations.¹⁴⁹ Recorded sound on mobile devices has also been proposed to monitor lung function in asthmatics.¹⁵⁰ While requiring further validation, these techniques could be used to develop future telehealth solutions including smartphone-based applications, which have the potential to aid decision-making and self-monitoring in asthma.

AI has also attracted attention for the monitoring of airborne pollen and allergens. Traditionally, pollen monitoring has relied on manual labeling methods, that, while accurate, can be labor-intensive and time-consuming.¹⁵¹ In contrast, public weather websites now incorporate real-time pollen counts sourced from research institutes that provide AI-based pollen counts.¹⁵² Recently, the BAA500 monitoring system was introduced, which automates the counting using microscopy and a supervised image recognition algorithm,¹⁵³ reaching a 90% detection rate within a 3-h sampling time. Another example is the *beenose* software, which uses a classification model based on optical signatures of pollen.¹⁵¹ Besides detection, ML has also been used to forecast future pollen concentrations.^{154,155} Cordero et al.¹⁵⁵ combined various models to predict the *Olea* pollen concentrations in the Community of Madrid. By incorporating sources such as meteorological and historical pollen data, they generated a model that could explain over 75% of variance.

4.2 | Fundamental research

AI can provide insights into disease classification, pathophysiology, and the underlying biological mechanisms, by clustering large numbers of data points into interpretable patterns.

4.2.1 | Heterogeneity and endotype discovery

There is an increasing awareness that allergic diseases (asthma, eczema, rhinitis, food allergy) are umbrella terms of subtypes

characterized by distinct disease mechanisms (endotypes).¹⁵⁶ Developments in ML techniques provide new ways to capture the heterogeneity in longitudinal patterns of the development of distinct symptoms of allergic diseases in individual patients. For example, childhood wheezing illness has been extensively investigated using ML approaches to derive more homogenous groups for genetic, mechanistic, and therapeutic studies.^{33,157–163} Most studies modeled repeated measurements of wheeze through the life-course to derive classes. These different symptom patterns may indicate distinct biological mechanisms,^{164,165} and their discovery may facilitate stratified treatment,¹⁶⁶ but this is not certain (i.e., the classes may not directly translate to endotypes).¹⁰⁹ However, recent studies from the US CREW and UK UNICORN consortia demonstrated that LCA using binary information on wheezing might classify individuals imprecisely,^{161,167,168} and children with identical wheezing patterns can be assigned to different phenotypes.¹⁶⁷ Recently, a novel data-driven method suggested a potential way to improve assignment to wheeze “phenotypes”.¹⁶⁹ Repeated observations of current wheezing were transformed to derive multidimensional indicators of wheezing spells (reflecting duration, temporal sequencing, and the extent of persistence/recurrence). Clustering these indicators resulted in a structure that was much more robust to data imputation, and with a remarkably high agreement between cluster assignment of individual children when using complete or imputed data.^{170,171}

Similarly, over the past 5 years, longitudinal data on eczema were clustered using data-driven approaches.^{172–179} There were notable differences in the estimated prevalence of each phenotype, and inconsistent associations with the filaggrin (*FLG*) genotype.

Bayesian machine learning approach has been used to model the development of eczema, wheeze, and rhinitis from birth to school-age.¹⁸⁰ The developmental profiles were heterogeneous, and the progression of the symptoms fitting the atopic march profile was rare among those with atopic comorbidities.¹⁸⁰ The findings revealed eight latent profiles of symptom development, each with different temporal patterns of their co-manifestation,¹⁸⁰ and distinct genetic associates.¹⁸¹ Further studies indicated that atopic march, as initially described, occurs rarely, that most 2-disease combinations occur by chance, but that there is a very important cluster of multimorbidity (affecting ~8% of the population that have a high disease burden).¹⁶⁹

Numerous studies have applied ML clustering to identify asthma subtypes too.^{182–184} Different endotypes may have a specific response to treatment, making this differentiation potentially clinically significant.^{185,186} Using k-means clustering, researchers identified four distinct clusters of asthma patients in the Severe Asthma Research Program with different responses to corticosteroids (CS).¹⁸⁷ One cluster involves patients, that despite severe baseline airflow limitations, have the lowest response to CS with almost no improvement in lung function, suggesting that this group would benefit from alternative treatment options. The authors also show that the variables that characterize the clusters robustly predict cluster assignment in an independent test set.

While there are fewer examples, ML clustering has also advanced phenotyping in rhinitis, a condition that particularly is

characterized by significant clinical heterogeneity. Hierarchical unsupervised clustering has been applied to a Chinese population suffering from chronic rhinitis. This methodology identified six distinct rhinitis phenotypes, each demonstrating unique inflammatory patterns—characterized by elevated levels of local and serum IgE—and comorbidities of allergic rhinitis alongside asthma.¹⁸⁸ In a concurrent study, data from nasal cytology were integrated with sociodemographic information and medical history, employing latent class analysis to discern new phenotypes. The research team identified two clusters, each demonstrating distinct cellular patterns indicative of varying inflammation levels. These disparities were based on the diverse frequencies of neutrophils, mast cells, and eosinophil counts.¹⁸⁹

A hypothesis-generating unbiased analysis which included data on lower airway inflammation and infection from bronchoalveolar lavage in preschool children with severe wheeze revealed four distinct pathophysiological clusters of approximately equal size: (1) atopic; (2) non-atopic, low infection rate; (3) non-atopic, high infection rate; and (4) non-atopic, low infection rate, no inhaled corticosteroids (ICS),¹⁹⁰ with marked differences in BAL microbial profiles between the clusters. In a multicenter prospective study, authors used clustering on integrated clinical, virus, and serum proteome data to identify a cluster in children with bronchiolitis with a significantly higher risk of developing asthma by age six.¹⁹¹ Multi-omics has also been employed in this domain, such as the novel and open-source method Merged Affinity Network Association Clustering (MANAclust), which provides an automated pipeline to integrate clinical and omics data. The authors identified clinically and molecularly distinct asthma clusters that responded differently to treatment, and substantial heterogeneity in healthy controls.¹⁹² In another recent study, researchers used unsupervised clustering on proteomics data of infants hospitalized with bronchiolitis. They identified two distinct clusters with dysregulated pathways and a higher risk for developing asthma.¹⁹¹ ML approaches have also shown utility for clustering exhaled volatile organic compounds (VOCs) in exhaled breath (breathomics), an exciting noninvasive biomarker for airway disease sensitive to inflammation.¹⁹³

4.2.2 | Pathways and disease mechanisms

Multi-omics and system biology are comprehensive approaches expected to increase insight into the complex biological mechanisms underlying allergic and immunological diseases.¹⁹⁴ The level of detail of such studies can be increased further using single-cell methods, analyzing gene expression profiles, chromatin accessibility, CpG methylation, or the proteome in thousands of cells individually.^{190,191} A fully integrated reference atlas has recently been released for the lung, with consensus annotations for 61 cell types based on data from more than 100 healthy tissue donors.¹⁹⁵ Using a trained model of this fully integrated healthy lung cell atlas, the dataset was expanded by projection and transfer learning using scArches¹⁹⁶ to a dataset of more than 2.4 million cells from more than 480 individuals. This

illustrates the use of deep learning in biology, to define cell types and states. This extended Lung Cell Atlas allowed direct comparison of cell types across datasets based on consensus labels, leading to the identification of disease-associated cell states common to multiple lung diseases.¹⁹⁵

4.3 | Drug and therapy development and precision medicine

AI has the potential to accelerate drug discovery and development throughout the whole pipeline¹⁹⁷ and contribute to precision medicine.¹⁹⁷ Precision medicine promises to enable personalized and more effective treatments based on an individual's genetic variability, environmental exposures, and lifestyle.¹⁹⁸⁻²⁰⁰ We here highlight promising examples for treatment response analysis and drug repurposing.

4.3.1 | Treatment response

In a pediatric cohort, asthma control after 6 months of medication could be accurately predicted using an AdaBoost classification algorithm,²⁰¹ outperforming traditional logistic regression.¹⁸² Wu et al.²⁰² developed a supervised ML model to predict low response to dupilumab in atopic dermatitis patients. The authors identified various indicators of nonresponse, including a high Quan-Charlson Comorbidity Index value, a claim for ibuprofen, or no claims for prednisone medication before dupilumab initiation. Similar approaches have been pursued to analyze nonresponse to Type 2-directed biologics in asthma patients.²⁰³

4.3.2 | Drug repurposing

Artificial intelligence has been used extensively in drug repurposing to overcome the immense time and investments required for new drug development.²⁰⁴⁻²⁰⁷ AI has been applied for virtual drug screening, treatment combination optimization, and drug-target interaction predictions.^{29,208,209} Patrick et al.²¹⁰ developed a workflow to model drug-disease relationships using unsupervised text analysis and supervised classification for cutaneous diseases, including atopic dermatitis. They created word embeddings—a dimensionality reduction method that creates a lower-dimension projection of high-dimensional text data—from 20 million abstracts in PubMed. Some of the strongest identified associations were not directly mentioned in any research article, demonstrating how the analysis of large-scale textual data can unveil novel repurpose opportunities. Despite promising results in other medical fields,²¹¹⁻²¹⁴ we identify a research gap in target discovery and clinical trial optimization applied in the allergy domain. Also, of over 10,000 clinical studies related to allergic diseases, we could only identify five with a fundamental role for artificial intelligence (search [ClinicalTrials.gov](https://clinicaltrials.gov), performed March 13, 2023).

TABLE 2 Guidelines for good AI and ML usage.

| Category | Guiding principle | Additional description | Literature coherence |
|-----------------------|--|--|--|
| Purpose and relevance | (P1) Disclose which clinical problem the model addresses and how it fits in a clinical workflow (P2) Collect modeling data in a consistent, clinically relevant and generalizable manner that aligns with the intended use (P3) Benchmark performance to existing clinical standards of care or previous AI studies or proof of concepts | Choosing a proper benchmark is essential to demonstrate clinical relevance and show the potential for patient care | MI-CLAIM, CONSORT-AI, FUTURE-AI DECIDE-AI, MI-CLAIM, FDA FUTURE-AI, TRIPOD-AI, MI-CLAIM, FUTURE-AI |
| Model development | (M1) Design a conceptual model with a definition of the predicted outcome and its presumed relationship to the input variables (M2) Safeguard appropriate separation between training, validation, and test datasets (M3) Ensure proper documentation and execution of model optimization steps | A conceptual model stimulates inclusion of domain expertise, focus and prioritization in data collection, and alignment with existing hypothesis and knowledge Ensure that model optimization is performed on the training set with tuning of model configuration on the validation set, without affecting the test set The number of configurations steps and decision is generally extensive and requires thorough tracking and documentation. In addition, it is vital that these steps are not performed on the test set | TRIPOD-AI, MI-CLAIM, MINIMAR, FDA MINIMAR, FUTURE-AI |

TABLE 2 (Continued)

| Category | Guiding principle | Additional description | Literature coherence |
|-------------------------|--|---|--|
| Replicability | (M4) Determine the evaluation procedure, metrics and rationale up-front, before starting the modeling procedure | Defining metrics post-analysis is a common pitfall that can lead to overestimation of performance. The evaluation procedure should be in concordance with what it clinically relevant | MI-CLAIM, FUTURE-AI |
| | (R1) Evaluate model performance in a prospective study, randomized trial, or at least an independent replication cohort | Ensure that the testing conditions are clinically relevant and representative for the intended usage context | RISE, MINIMAR, FDA |
| | (R2) Perform sensitivity and robustness checks to assess whether the system is impartial to changing environments or populations | This can further be boosted by training the model on a heterogeneous population | TRIPOD-AI, MI-CLAIM, FUTURE-AI |
| Explainability | (R3) Disclose data preprocessing and the way in which data quality is assessed and ensured | | DECIDE-AI, TRIPOD-AI, MI-CLAIM, MINIMAR, CONSORT-AI, FUTURE-AI |
| | (E1) Determine and provide appropriate levels of interpretability, depending on use case and users | This often guides which algorithm and interpretability tools need to be employed | RISE, TRIPOD-AI, MI-CLAIM, FUTURE-AI |
| System design and usage | (E2) Leverage interpretability toolkits and libraries for black box models | The complexity and black box nature of AI models warrant more focus on interpretability | RISE, FUTURE-AI |
| | (S1) Focus on multi-disciplinary collaboration during the full AI solution lifecycle | Involve a broad range of functional expertise, from AI leads, users, clinicians, study-design experts, in all phases of development and implementation | RISE, FDA, FUTURE-AI |
| | (S2) Invest in the instruction of users on how to interact with the system and predictions | Establishing trust in the solution and explaining how it integrates with the clinical workflow is key for adoption and impact on clinical outcome | DECIDE-AI, MINIMAR, CONSORT-AI, FUTURE-AI, FDA |
| Risks and ethics | (S3) Set up monitoring processes to track technical and analytical performance | | FDA, FUTURE-AI |
| | (S4) Set up a feedback flow to facilitate iterative system improvement | Performance results and user interaction and feedback can be used to improve the system in a targeted way. This can be either periodically or automated in a more self-learning setup, although complexity and existing legislation limit usage of the latter | DECIDE-AI |
| | (R1) Define and evaluate the ethical considerations of the system, e.g. algorithmic fairness | Various frameworks are available for the responsible application of AI and ML, which provide guidance on the relevant components and methods to assess them | DECIDE-AI, CONSORT-AI, FUTURE-AI |
| | (R2) Assess the potential risks involved in the system and outline approach to manage and mitigate them | | DECIDE-AI |

Note: Best practices are coherent with existing frameworks, where the overlap is annotated with Roman numbers as I—CONSORT-AI⁷⁶; II—FUTURE-AI²²⁹; III—DECIDE-AI¹⁰⁶; IV—TRIPOD²⁴⁷; V—TRIPOD-AI (in development)²¹⁸; VI—STARD-AI (in development)²¹⁶; VII—MI-CLAIM²³⁰; VIII—MINIMAR²³¹; IX—FDA checklist²¹⁵; and X—RISE criteria.¹⁰⁵ Reporting guidelines (I–VIII) are generally broader than just the AI and ML component, with substantial overlap with standard protocols for academic reporting, such as encouragement of code availability and well-described participant characteristics. Our focus is on the unique and new elements when applying AI in a biomedical context. Components are referenced to TRIPOD-AI when the Delphi round 1 consensus was larger than 80% for being essential and desirable to include as reporting.

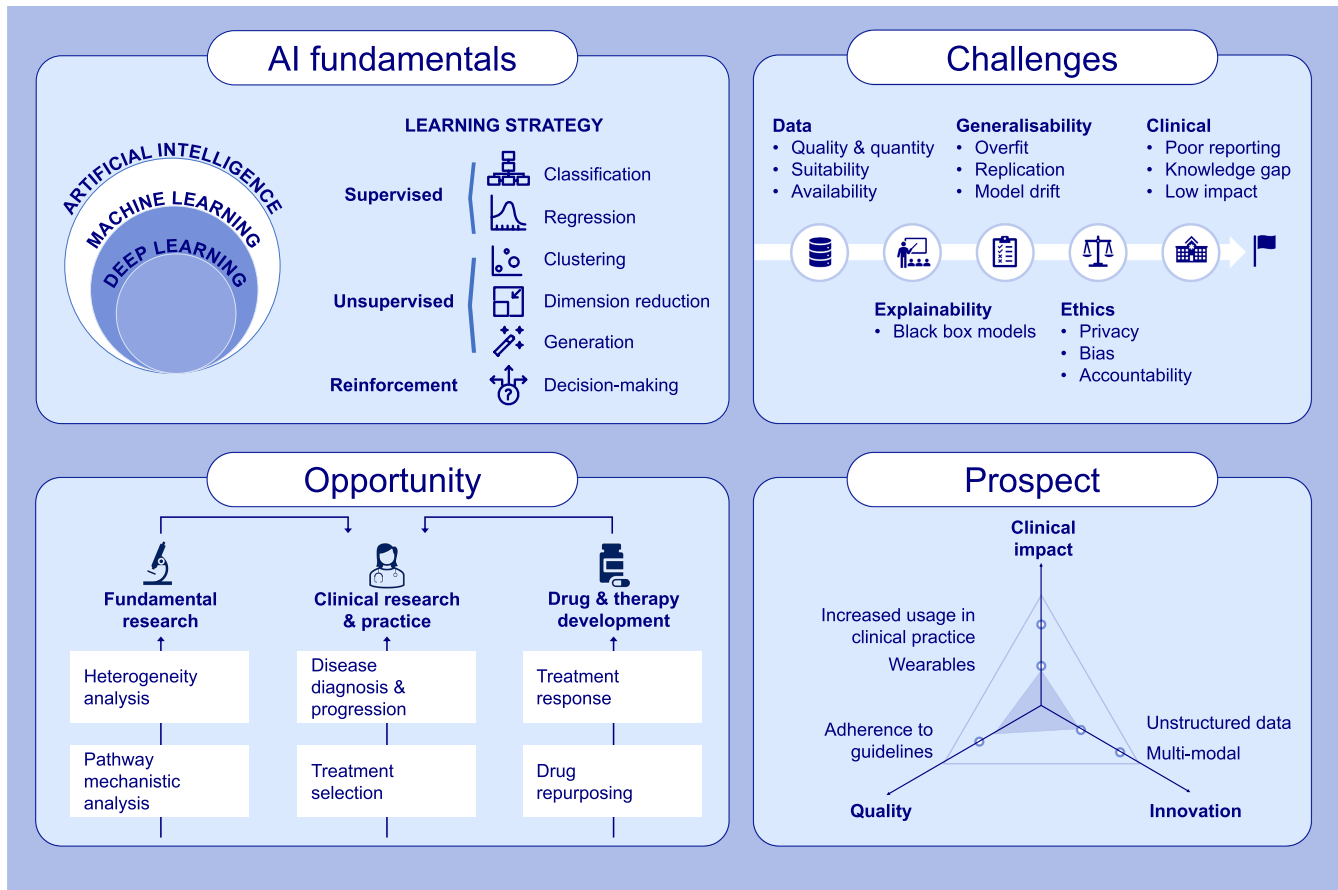


FIGURE 6 Graphical summary of narrative review. This article provides a fundamental understanding of core AI concepts, its critical challenges, and the current and prospect opportunities of its application within the field of allergy.

5 | FUTURE PROSPECTS

Successfully translating AI proof of concepts into clinical practice remains pivotal for fully realizing AI's impact.

While practical guidelines and best practices are emerging in medical AI, they are not always adhered to and require frequent reassessment due to the pace at which the AI field is moving forward. When implementing AI, it is strongly recommended to verify available guidelines to ensure applications are reliable and provide meaningful outcomes. We here propose a set of minimal requirements for good practice in AI (Table 2) based on published guidelines of the FDA,²¹⁵ literature on best-practice model development in biomedicine,⁹² or expert-based checklists for developing and reporting algorithms (e.g., STARD-AI,²¹⁶ TRIPOD checklist,²¹⁷ and awaited TRIPOD-AI adaptation²¹⁸).

In the allergy field, research beyond proofs-of-concept is relatively scarce, let alone meaningful clinical applications. We provide an expert outlook on noteworthy AI trends. Firstly, the ever-increasing accessibility, automation, and transferability of *ML tooling* are expected to drive AI adoption further, enabling non-specialized researchers to apply novel techniques. Secondly, we expect an increase in the use of *unstructured data*. Innovations such as AI-based image analysis, NLP, and generative AI are at the forefront of academic efforts in computer science while being

underutilized in our field. For clinicians, an AI clinical assistant akin to readily available “home assistants,” could quietly listen in on consults and subsequently support in documentation in EHRs, diagnosis, and therapy suggestions. Clinical solutions that leverage speech recognition are entering the market, aiming to improve the clinical workflow and efficiency,²¹⁹ although adoption and showcases of tangible impact are still limited. Thirdly, the emerging trend of multi-modal learning can open new research avenues by integrating multiple data sources and modalities in a singular analytical approach, hereby creating more holistic models and insights. Fourthly, new AI tools and website are being released to the public on a daily basis. These tools will improve, or at least change, the way we work and do research. Table S1 provides a selection of the key AI resources covering topics general productivity, research productivity, modeling resources, and educational resources.

The largest future impact from AI is expected when current proofs-of-concept are translated successfully to clinical practice. The US and the EU are making steps toward developing AI and algorithm regulations, to facilitate updates and improve privacy, security, and transparency.^{220,221} While FDA-approved applications in the field of allergy are still absent, there are initiatives that already use, or could use, AI to positively affect patient care. One recent example is the mobile app EczemaLess,²²² which monitors eczema

progression via uploaded images. For asthma, the a-GPS tool and the MaskAIR^{223,224} app enable patients and clinicians to better monitor their disease and quality of care. Looking forward, we expect the largest contribution of AI from areas where this technology has a significant edge on existing methods, in particular where data volumes are massive, real-life, or unstructured. This makes the development of smart wearables very promising.^{225,226} Notably, these developments would require multi-disciplinary collaboration, as not only the analytics but also technical and clinical development become vital.

The developers of algorithms play a role in clinical translation, and clinicians would need to adapt to the integration of AI within health care. While most AI systems are designed as a support mechanism rather than a replacement, it will change their work and role. Clinician training in the fundamentals of AI is needed to gain trust in these systems and work with them effectively. One of the common concerns regarding AI is that these systems will replace humans in their installment. While many studies position their analytical solution in a head-to-head comparison with humans, most clinical applications are designed as decision-support tools that strengthen and assist experts in their profession rather than replacing them.⁴¹ Lastly, we foresee further developments in dynamic learning systems, which continuously evolve based on clinical usage. Such approaches are rare, and FDA-approved tools are generally “locked,” referring to a fixed algorithm state.²²⁷ The FDA is working on an action plan to better assess and support such applications.²²⁸

In conclusion, the potential of AI to transform clinical medicine is evident, but the steps from a proof of concept to clinical applications are not easily made (Figure 6). Innovations from the field of AI can address many important open questions in allergy; we anticipate that good future utilization of AI (Table 1) will deepen our knowledge of disease mechanisms and contribute to precision medicine in allergy.

AUTHOR CONTRIBUTIONS

MvB and GHK devised and set out the overall outline of this review. RSNF, MB, and MCN have provided critical input on the outline and figure drafts for improved positioning and relevance for the readership. FIR, JWH, SF, and AC have been invited to the author group to bring in additional domain expertise, for which they have written and/or sharpened the respective sections. All authors have provided feedback on the first complete draft of the manuscript. MvB and GHK wrote this first draft and have revised it based on inputs from all authors. All authors have read and approved the final version of this manuscript.

ACKNOWLEDGMENTS

The authors acknowledge the team members from the broader research collaboration between UMCG and Mlcompany for their critical feedback on the positioning and narrative of this article.

CONFLICT OF INTEREST STATEMENT

All authors declare to have no conflict of interests.

FUNDING INFORMATION

The authors have not received financial support for the research, authorship, and publication of this manuscript.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Merlijn van Breugel  <https://orcid.org/0000-0002-3130-5611>

Rudolf S. N. Fehrmann  <https://orcid.org/0000-0002-7516-315X>

John W. Holloway  <https://orcid.org/0000-0001-9998-0464>

REFERENCES

1. FDA. Artificial intelligence and machine learning (AI/ML)-enabled medical devices. Accessed December 11, 2022. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices>
2. Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. *Lancet Digit Health*. 2021;3(3):e195–e203. doi:10.1016/S2589-7500(20)30292-2
3. Zhu S, Gilbert M, Chetty I, Siddiqui F. The 2021 landscape of FDA-approved artificial intelligence/machine learning-enabled medical devices: an analysis of the characteristics and intended use. *Int J Med Inform*. 2022;165:104828. doi:10.1016/J.IJMEDINF.2022.104828
4. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230–243. doi:10.1136/SVN-2017-000101
5. Seetharam K, Shrestha S, Sengupta PP. Artificial intelligence in cardiovascular medicine. *Curr Treat Options Cardiovasc Med*. 2019;21:25. doi:10.1007/s11936-019-0728-1
6. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med Inform Decis Mak*. 2019;19(1):1–16. doi:10.1186/S12911-019-1004-8/FIGURES/12
7. Bzdok D, Altman N, Krzywinski M. Points of significance: statistics versus machine learning. *Nat Methods*. 2018;15(4):233–234. doi:10.1038/NMETH.4642
8. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng*. 2018;2(10):719–731. doi:10.1038/S41551-018-0305-Z
9. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:7553. doi:10.1038/nature14539
10. Shahinfar S, Meek P, Falzon G. “How many images do I need?” Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring. *Ecol Inform*. 2020;57:101085. doi:10.1016/J.ECOINF.2020.101085
11. Otter DW, Medina JR, Kalita JK. A survey of the usages of deep learning for natural language processing. *IEEE Trans Neural Netw Learn Syst*. 2021;32(2):604–624. doi:10.1109/TNNLS.2020.2979670
12. Shrestha A, Mahmood A. Review of deep learning algorithms and architectures. *IEEE Access*. 2019;7:53040–53065. doi:10.1109/ACCESS.2019.2912200
13. Piccialli F, di Somma V, Giampaolo F, Cuomo S, Fortino G. A survey on deep learning in medicine: why, how and when? *Inf Fusion*. 2021;66:111–137. doi:10.1016/J.INFFUS.2020.09.006
14. Wang F, Casalino LP, Khullar D. Deep learning in medicine—promise, progress, and challenges. *JAMA Intern Med*. 2019;179(3):293–294. doi:10.1001/JAMAINTERNMED.2018.7117

15. Ting DSW, Cheung CYL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *Jama*. 2017;318(22):2211-2223. doi:10.1001/JAMA.2017.18152
16. FDA. FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. Accessed December 13, 2022. <https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye>
17. van der Heijden AA, Abramoff MD, Verbraak F, van Hecke M, Liem A, Nijpels G. Validation of automated screening for referable diabetic retinopathy with the IDx-DR device in the Hoorn diabetes care system. *Acta Ophthalmol*. 2018;96(1):63-68. doi:10.1111/AOS.13613
18. Samala RK, Chan HP, Hadjiiski L, Helvie MA, Wei J, Cha K. Mass detection in digital breast tomosynthesis: deep convolutional neural network with transfer learning from mammography. *Med Phys*. 2016;43(12):6654-6666. doi:10.1118/1.4967345
19. Lotter W, Diab AR, Haslam B, et al. Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. *Nat Med*. 2021;27(2):244-249. doi:10.1038/s41591-020-01174-9
20. Potnis KC, Ross JS, Aneja S, Gross CP, Richman IB. Artificial intelligence in breast cancer screening: evaluation of FDA device regulation and future recommendations. *JAMA Intern Med*. 2022;182(12):1306-1312. doi:10.1001/JAMAINTERNMED.2022.4969
21. DALL-E 2. Accessed January 12, 2023. <https://openai.com/dall-e-2/>
22. ChatGPT: optimizing language models for dialogue. Accessed January 12, 2023. <https://openai.com/blog/chatgpt/>
23. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning? A primer for the epidemiologist. *Am J Epidemiol*. 2019;188(12):2222-2239. doi:10.1093/AJE/KWZ189
24. Barinov L, Jairaj A, Becker M, et al. Impact of data presentation on physician performance utilizing artificial intelligence-based computer-aided diagnosis and decision support systems. *J Digit Imaging*. 2019;32(3):408-416. doi:10.1007/S10278-018-0132-5/TABLES/6
25. Mango VL, Sun M, Wynn RT, Ha R. Should we ignore, follow, or biopsy? Impact of artificial intelligence decision support on breast ultrasound lesion assessment. *AJR Am J Roentgenol*. 2020;214(6):1445. doi:10.2214/AJR.19.21872
26. Lalmuanawma S, Hussain J, Chhachhuak L. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: a review. *Chaos Solitons Fractals*. 2020;139:110059. doi:10.1016/J.CHAOS.2020.110059
27. Albahri AS, Hamid RA, Alwan JK, et al. Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review. *J Med Syst*. 2020;44:1-11. doi:10.1007/s10916-020-01582-x
28. Kadioglu O, Saeed M, Greten HJ, Efferth T. Identification of novel compounds against three targets of SARS CoV-2 coronavirus by combined virtual screening and supervised machine learning. *Comput Biol Med*. 2021;133:104359. doi:10.1016/J.COMPBIOMED.2021.104359
29. Kowalewski J, Ray A. Predicting novel drugs for SARS-CoV-2 using machine learning from a >10 million chemical space. *Heliyon*. 2020;6(8):e04639. doi:10.1016/J.HELİYON.2020.E04639
30. Ho TK. Random decision forests. *Proc Int Conf Doc Anal Recognit*. 1995;1:278-282. doi:10.1109/ICDAR.1995.598994
31. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. *Proc ACM SIGKDD Int Conf Knowl Discov Data Min*. 2016;785-794. doi:10.1145/2939672.2939785
32. Ke G, Meng Q, Finley T, et al. LightGBM: a highly efficient gradient boosting decision tree. *Adv Neural Inf Process Syst*. 2017;30:3147-3156.
33. Oksef C, Haider S, Fontanella S, Frainay C, Custovic A. Classification of pediatric asthma: from phenotype discovery to clinical practice. *Front Pediatr*. 2018;6:258. doi:10.3389/FPED.2018.00258
34. Brew BK, Chiesa F, Lundholm C, Örtqvist A, Almqvist C. A modern approach to identifying and characterizing child asthma and wheeze phenotypes based on clinical data. *PLoS One*. 2019;14(12):e0227091. doi:10.1371/JOURNAL.PONE.0227091
35. Wu W, Bleecker E, Moore W, et al. Unsupervised phenotyping of severe asthma research program participants using expanded lung data. *J Allergy Clin Immunol*. 2014;133(5):1280-1288. doi:10.1016/J.JACI.2013.11.042
36. Hagenaars JA, McCutcheon AL. *Applied Latent Class Analysis*. Cambridge University Press; 2002.
37. MacQueen J. Some methods for classification and analysis of multivariate observations. *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability Source: Project Euclid Journals - Open Access Supplement* 1967.
38. Abdi H, Williams LJ. Principal component analysis. *Wiley Interdiscip Rev Comput Stat*. 2010;2(4):433-459. doi:10.1002/WICS.101
39. Kruskal JB. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*. 1964;29(1):1-27. doi:10.1007/BF02289565/METRICS
40. Cheplygina V, de Bruijne M, Pluim JPW. Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. *Med Image Anal*. 2019;54:280-296. doi:10.1016/J.MEDIA.2019.03.009
41. Yu C, Liu J, Nemati S, Yin G. Reinforcement learning in healthcare: a survey. *ACM Comput Surv*. 2021;55(1):1-36. doi:10.1145/3477600
42. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med*. 2018;24(11):1716-1720. doi:10.1038/s41591-018-0213-5
43. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol*. 2019;103(2):167-175. doi:10.1136/BJOPHTHALMOL-2018-313173
44. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7
45. Rajpurkar P, Chen E, Banerjee O, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med*. 2022;28:1-8. doi:10.1038/S41591-021-01614-0
46. Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface*. 2018;15(141):20170387. doi:10.1098/RSIF.2017.0387
47. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol*. 2019;20(5):e262-e273. doi:10.1016/S1470-2045(19)30149-4
48. Jordon J, Szpruch L, Houssiau F, et al. Synthetic data - what, why and how? doi:10.48550/arxiv.2205.03257
49. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data*. 2019;6(1):1-48. doi:10.1186/S40537-019-0197-0/FIGURES/33
50. Zhuang F, Qi Z, Duan K, et al. A comprehensive survey on transfer learning. *Proc IEEE*. 2021;109(1):43-76. doi:10.1109/JPROC.2020.3004555
51. Bhattacharya S, Hu Z, Butte AJ. Opportunities and challenges in democratizing immunology datasets. *Front Immunol*. 2021;12:846. doi:10.3389/FIMMU.2021.647536/BIBTEX
52. Open-access data and computational resources to address COVID-19 | data science at NIH. Accessed December 28, 2022. <https://datascience.nih.gov/covid-19-open-access-resources>

53. European health data space. Accessed March 8, 2023. https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en
54. Rieke N, Hancox J, Li W, et al. The future of digital health with federated learning. *NPJ Digit Med.* 2020;3(1):1-7. doi:10.1038/s41746-020-00323-1
55. Warnat-Herresthal S, Schultze H, Shastry KL, et al. Swarm learning for decentralized and confidential clinical machine learning. *Nature.* 2021;594(7862):265-270. doi:10.1038/S41586-021-03583-3
56. Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. *J Biomed Inform.* 2021;113:103655. doi:10.1016/J.JBI.2020.103655
57. High-tech partnership refines artificial intelligence in health care. Accessed December 12, 2022. <https://www.nature.com/articles/d43978-022-00127-9>
58. van der Velden BHM, Kuijff HJ, Gilhuijs KGA, Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med Image Anal.* 2022;79:102470. doi:10.1016/J.MEDIA.2022.102470
59. Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak.* 2020;20(1):1-9. doi:10.1186/S12911-020-01332-6
60. Nazar M, Alam MM, Yafi E, Su'Ud MM. A systematic review of human-computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques. *IEEE Access.* 2021;9:153316-153348. doi:10.1109/ACCESS.2021.3127881
61. Tjoa E, Guan C. A survey on explainable artificial intelligence (XAI): toward medical XAI. *IEEE Trans Neural Netw Learn Syst.* 2021;32(11):4793-4813. doi:10.1109/TNNLS.2020.3027314
62. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst.* 2017;30:4766-4775. doi:10.48550/ARXIV.1705.07874
63. Ribeiro MT, Singh S, Guestrin C. Why should I trust you? Explaining the predictions of any classifier. *NAACL-HLT 2016 - 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Demonstrations Session* (pp. 97-101) 2016. doi:10.48550/ARXIV.1602.04938
64. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. *Proc IEEE Conf Comput Vis Pattern Recognit.* 2016;2921-2929.
65. Kothalawala DM, Murray CS, Simpson A, et al. Development of childhood asthma prediction models using machine learning approaches. *Clin Transl Allergy.* 2021;11(9):e12076. doi:10.1002/CLT2.12076
66. Wolpert DH. The lack of a priori distinctions between learning algorithms. *Neural Comput.* 1996;8(7):1341-1390. doi:10.1162/NECO.1996.8.7.1341
67. Kaur H, Kumari V. Predictive modelling and analytics for diabetes using a machine learning approach. *Appl Comput Inform.* 2002;18:90-100. doi:10.1016/j.aci.2018.12.004
68. Liu Y, Chen PHC, Krause J, Peng L. How to read articles that use machine learning: users' guides to the medical literature. *Jama.* 2019;322(18):1806-1816. doi:10.1001/JAMA.2019.16489
69. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med.* 2019;380(14):1347-1358. doi:10.1056/NEJMRA1814259/SUPPL_FILE/NEJMRA1814259_DISCLOSURES.PDF
70. Patel D, Hall GL, Broadhurst D, Smith A, Schultz A, Foong RE. Does machine learning have a role in the prediction of asthma in children? *Paediatr Respir Rev.* 2022;41:51-60. doi:10.1016/J.PRRV.2021.06.002
71. Berisha V, Krantsevich C, Hahn PR, et al. Digital medicine and the curse of dimensionality. *NPJ Digit Med.* 2021;4(1):1-8. doi:10.1038/s41746-021-00521-5
72. IBM's Watson recommended "unsafe and incorrect" cancer treatments - STAT. Accessed December 29, 2022. <https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/>
73. Ebrahimian S, Kalra MK, Agarwal S, et al. FDA-regulated AI algorithms: trends, strengths, and gaps of validation studies. *Acad Radiol.* 2022;29(4):559-566. doi:10.1016/J.ACRA.2021.09.002
74. Chi S, Tian Y, Wang F, Zhou T, Jin S, Li J. A novel lifelong machine learning-based method to eliminate calibration drift in clinical prediction models. *Artif Intell Med.* 2022;125:102256. doi:10.1016/J.ARTMED.2022.102256
75. Price WN, Cohen IG. Privacy in the age of medical big data. *Nat Med.* 2019;25(1):37-43. doi:10.1038/s41591-018-0272-7
76. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Lancet Digit Health.* 2020;2:e537. doi:10.1136/bmj.m3164
77. Cruz Rivera S, Liu X, Chan AW, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Lancet Digit Health.* 2020;2(10):e549-e560. doi:10.1016/S2589-7500(20)30219-3
78. Plana D, Shung DL, Grimshaw AA, Saraf A, Sung JY, Kann BH. Randomized clinical trials of machine learning interventions in health care: a systematic review. *JAMA Netw Open.* 2022;5(9):e2233946. doi:10.1001/JAMANETWORKOPEN.2022.33946
79. Sahiner B, Chen W, Samala RK, Petrick N. Data drift in medical machine learning: implications and potential remedies. *Br J Radiol.* 2023;20220878. doi:10.1259/BJR.20220878
80. The impact of the general data protection regulation (GDPR) on artificial intelligence | think tank | European Parliament. Accessed March 13, 2023. [https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU\(2020\)641530](https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)641530)
81. Seastedtid KP, Schwab P, O'brien Z, et al. Global healthcare fairness: we should be sharing more, not less, data. *PLOS Digit Health.* 2022;1(10):e0000102. doi:10.1371/JOURNAL.PDIG.0000102
82. Torkezadehmahani R, Nasirigerdeh R, Blumenthal DB, et al. Privacy-preserving artificial intelligence techniques in biomedicine. *Methods Inf Med.* 2022;61(S 01):E12-E27. doi:10.1055/S-0041-1740630/ID/JR21020001-8
83. Torfi A, Fox EA, Reddy CK. Differentially private synthetic medical data generation using convolutional GANs. *Inf Sci.* 2022;586:485-500. doi:10.1016/J.INS.2021.12.018
84. Jordan J, Yoon J, van der Schaar M. PATE-GAN: Generating Synthetic Data with Differential Privacy Guarantees. *ICLR*; 2019.
85. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. *ACM Comput Surv.* 2021;54(6):1-35. doi:10.1145/3457607
86. US Food and Drug Administration. Software as a medical device (SaMD) action plan. 2021. Accessed January 2, 2023. www.fda.gov
87. Cirillo D, Catuara-Solarz S, Morey C, et al. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. *NPJ Digit Med.* 2020;3(1):1-11. doi:10.1038/s41746-020-0288-5
88. EPRS | European parliamentary research service. Artificial Intelligence in Healthcare 2022. doi:10.2861/568473
89. Dael VJ, Ekechi C, Gardner A, et al. Addressing racial and ethnic inequities in data-driven health technologies. 2022;53. doi:10.25561/94902
90. Santoni De Sio F, Mecacci G, Santoni De Sio F, Mecacci G. Four responsibility gaps with artificial intelligence: why they matter and how to address them. *Philos Technol.* 2021;34:1057-1084. doi:10.1007/s13347-021-00450-x

91. Daugherty SL, Blair I, Havranek EP, et al. Implicit gender bias and the use of cardiovascular tests among cardiologists. *J Am Heart Assoc.* 2017;6(12):e006872. doi:10.1161/JAHA.117.006872
92. Bradshaw TJ, Boellaard R, Dutta J, et al. Nuclear medicine and artificial intelligence: best practices for algorithm development. *J Nucl Med.* 2022;63(4):500-510. doi:10.2967/jnumed.121.262567
93. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med.* 2019;170(1):51-58. doi:10.7326/M18-1376/SUPPL_FILE/M18-1376_SUPPLEMENT.PDF
94. Briggs E, Hollmén J. Mitigating discrimination in clinical machine learning decision support using algorithmic processing techniques. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 19–33) 2020. doi:10.1007/978-3-030-61527-7_2/FIGURES/1
95. Zhang BH, Lemoine B, Mitchell M. Mitigating unwanted biases with adversarial learning. *AIES.* 2018;18:335-340. doi:10.1145/3278721.3278779
96. Lee MSA, Singh J. The landscape and gaps in open source fairness toolkits. *Conf Hum Factors Comput Syst Proc.* 2021;6:1-13. doi:10.1145/3411764.3445261
97. Bellamy RKE, Mojsilovic A, Nagar S, et al. AI fairness 360: an extensible toolkit for detecting and mitigating algorithmic bias. *IBM J Res Dev.* 2019;63:1-4. doi:10.1147/JRD.2019.2942287
98. Doshi-Velez F, Kortz M, Budish R, et al. Accountability of AI under the law: the role of explanation. *SSRN Electron J.* 2017. doi:10.48550/1711.01134
99. Verdicchio M, Perin A. When doctors and AI interact: on human responsibility for artificial risks. *Philos Technol.* 2022;35(1):1-28. doi:10.1007/S13347-022-00506-6/METRICS
100. Habli I, Lawton T, Porter Z. Artificial intelligence in health care: accountability and safety. *Bull World Health Organ.* 2020;98(4):251. doi:10.2471/BLT.19.237487
101. Emanuel EJ, Wachter RM. Artificial intelligence in health care: will the value match the hype? *Jama.* 2019;321(23):2281-2282. doi:10.1001/JAMA.2019.4914
102. Houry P, Srinivasan R, Kakumanu S, et al. A framework for augmented intelligence in allergy and immunology practice and research—a work group report of the AAAAI health informatics, technology, and education committee. *J Allergy Clin Immunol.* 2022;10(5):1178-1188. doi:10.1016/J.JAIP.2022.01.047
103. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* 2019;17(195):1-9. doi:10.1186/s12916-019-1426-2
104. Roberts M, Driggs D, Thorpe M, et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell.* 2021;3(3):199-217. doi:10.1038/s42256-021-00307-0
105. Varghese J. Artificial intelligence in medicine: chances and challenges for wide clinical adoption. *Visc Med.* 2020;36(6):443-449. doi:10.1159/000511930
106. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med.* 2022;28(5):924-933. doi:10.1038/s41591-022-01772-9
107. Meskó B, Görög M. A short guide for medical professionals in the era of artificial intelligence. *NPJ Digit Med.* 2020;3(1):126. doi:10.1038/s41746-020-00333-Z
108. Fontanella S, Cucco A, Custovic A. Machine learning in asthma research: moving toward a more integrated approach. *Expert Rev Respir Med.* 2021;15(5):609-621. doi:10.1080/17476348.2021.1894133
109. Duverdier A, Custovic A, Tanaka RJ. Data-driven research on eczema: systematic characterization of the field and recommendations for the future. *Clin Transl Allergy.* 2022;12(6):e12170. doi:10.1002/CLT2.12170
110. Nabi FG, Sundaraj K, Lam CK, Palaniappan R. Characterization and classification of asthmatic wheeze sounds according to severity level using spectral integrated features. *Comput Biol Med.* 2019;104:52-61. doi:10.1016/J.COMPBIOMED.2018.10.035
111. Seol HY, Rolfes MC, Chung W, et al. Expert artificial intelligence-based natural language processing characterises childhood asthma. *BMJ Open Respir Res.* 2020;7(1):e000524. doi:10.1136/BMJRESP-2019-000524
112. van Breugel M, Qi C, Xu Z, et al. Nasal DNA methylation at three CpG sites predicts childhood allergic disease. *Nat Commun.* 2022;13(1):1-13. doi:10.1038/s41467-022-35088-6
113. Wi Cl, Sohn S, Ryu E, Liu H, Park MA, Juhn YJ. Automated chart review for asthma ascertainment: an innovative approach for asthma care and research in the era of electronic medical record. *J Allergy Clin Immunol.* 2016;137(2):AB196. doi:10.1016/j.jaci.2015.12.771
114. Il WC, Sohn S, Ali M, et al. Natural language processing for asthma ascertainment in different practice settings. *J Allergy Clin Immunol Pract.* 2018;6(1):126-131. doi:10.1016/J.JAIP.2017.04.041
115. Jiang Z, Li J, Kong N, et al. Accurate diagnosis of atopic dermatitis by combining transcriptome and microbiota data with supervised machine learning. *Sci Rep.* 2022;12(1):290. doi:10.1038/S41598-021-04373-7
116. Alag A. Machine learning approach yields epigenetic biomarkers of food allergy: a novel 13-gene signature to diagnose clinical reactivity. *PLoS One.* 2019;14(6):e0218253. doi:10.1371/JOURNAL.PONE.0218253
117. Fontanella S, Frainay C, Murray CS, Simpson A, Custovic A. Machine learning to identify pairwise interactions between specific IgE antibodies and their association with asthma: a cross-sectional analysis within a population-based birth cohort. *PLoS Med.* 2018;15(11):e1002691. doi:10.1371/JOURNAL.PMED.1002691
118. Roberts G, Fontanella S, Selby A, et al. Connectivity patterns between multiple allergen specific IgE antibodies and their association with severe asthma. *J Allergy Clin Immunol.* 2020;146(4):821-830. doi:10.1016/J.JACI.2020.02.031
119. Toivonen L, Karppinen S, Schuez-Havupalo L, et al. Longitudinal changes in early nasal microbiota and the risk of childhood asthma. *Pediatrics.* 2020;146(4):20200421. doi:10.1542/PEDS.2020-0421
120. Spacova I, Petrova MI, Fremau A, et al. Intranasal administration of probiotic *Lactobacillus rhamnosus* GG prevents birch pollen-induced allergic asthma in a murine model. *Allergy.* 2019;74(1):100-110. doi:10.1111/ALL.13502
121. Bose S, Kenyon CC, Masino AJ. Personalized prediction of early childhood asthma persistence: a machine learning approach. *PLoS One.* 2021;16(3):e0247784. doi:10.1371/JOURNAL.PONE.0247784
122. Custovic A, Simpson BM, Murray CS, Lowe L, Woodcock A. The national asthma campaign manchester asthma and allergy study. *Pediatr Allergy Immunol.* 2002;13(s15):32-37. doi:10.1034/J.1399-3038.13.S.15.3.X
123. Junayed MS, Sakib ANM, Anjum N, Islam MB, Jeny AA. EczemaNet: a deep CNN-based eczema diseases classification. *IPAS.* 2020;174-179. doi:10.1109/IPAS50080.2020.9334929
124. Rasheed A, Umar AI, Shirazi SH, Khan Z, Nawaz S, Shahzad M. Automatic eczema classification in clinical images based on hybrid deep neural network. *Comput Biol Med.* 2022;147:105807. doi:10.1016/J.COMPBIOMED.2022.105807
125. Kalbande D, Naik R, Jatakia J, Khopkar U. An artificial intelligence approach for the recognition of early stages of eczema. *Int J Med Eng Inform.* 2020;12(1):52-61. doi:10.1504/IJMEI.2020.105656
126. Scheurer J, Ferrari C, Berenguer Todo Bom L, Beer M, Kempf W, Haug L. Semantic segmentation of histopathological slides for the classification of cutaneous lymphoma and eczema. *Commun Comput Inf Sci.* 2020;1248:26-42. doi:10.1007/978-3-030-52791-4_3

127. Guimarães P, Batista A, Zieger M, Kaatz M, Koenig K. Artificial intelligence in multiphoton tomography: atopic dermatitis diagnosis. *Sci Rep*. 2020;10(1):7968. doi:10.1038/S41598-020-64937-X
128. Newson RB, Van Ree R, Forsberg B, et al. Geographical variation in the prevalence of sensitization to common aeroallergens in adults: the GA2LEN survey. *Allergy*. 2014;69(5):643-651. doi:10.1111/ALL.12397
129. Davis CM, Apter AJ, Casillas A, et al. Health disparities in allergic and immunologic conditions in racial and ethnic underserved populations: a work group report of the AAAAI committee on the underserved. *J Allergy Clin Immunol*. 2021;147(5):1579-1593. doi:10.1016/J.JACI.2021.02.034
130. Cerrato P, Halamka J, Pencina M. A proposal for developing a platform that evaluates algorithmic equity and accuracy. *BMJ Health Care Inform*. 2022;29(1):100423. doi:10.1136/BMJHCI-2021-100423
131. Parisi GI, Kemker R, Part JL, Kanan C, Wermter S. Continual lifelong learning with neural networks: a review. *Neural Netw*. 2019;113:54-71. doi:10.1016/j.neunet.2019.01.012
132. Finkelstein J, Jeong I. Machine learning approaches to personalize early prediction of asthma exacerbations. *Ann N Y Acad Sci*. 2017;1387(1):153. doi:10.1111/NYAS.13218
133. Goto T, Camargo CA, Faridi MK, Yun BJ, Hasegawa K. Machine learning approaches for predicting disposition of asthma and COPD exacerbations in the ED. *Am J Emerg Med*. 2018;36(9):1650-1654. doi:10.1016/J.AJEM.2018.06.062
134. Patel SJ, Chamberlain DB, Chamberlain JM. A machine learning approach to predicting need for hospitalization for pediatric asthma exacerbation at the time of emergency department triage. *Acad Emerg Med*. 2018;25(12):1463-1470. doi:10.1111/ACEM.13655
135. Zein JG, Wu CP, Attaway AH, Zhang P, Nazha A. Novel machine learning can predict acute asthma exacerbation. *Chest*. 2021;159(5):1747-1757. doi:10.1016/J.CHEST.2020.12.051
136. Loymans RJB, Debray TPA, Honkoop PJ, et al. Exacerbations in adults with asthma: a systematic review and external validation of prediction models. *J Allergy Clin Immunol Pract*. 2018;6(6):1942-1952.e15. doi:10.1016/J.JAIP.2018.02.004
137. Zhang O, Minku LL, Gonem S. Detecting asthma exacerbations using daily home monitoring and machine learning. *J Asthma*. 2021;58(11):1518-1527. doi:10.1080/02770903.2020.1802746
138. Juhn Y, Wi CI, Sohn S, et al. Asthma-guidance and prediction system (a-GPS) as a precision asthma care tool. *J Allergy Clin Immunol*. 2020;145(2):AB210. doi:10.1016/j.jaci.2019.12.231
139. Seol HY, Shrestha P, Muth JF, et al. Artificial intelligence-assisted clinical decision support for childhood asthma management: a randomized clinical trial. *PLoS One*. 2021;16(8):e0255261. doi:10.1371/JOURNAL.PONE.0255261
140. Lugogo NL, DePietro M, Reich M, et al. A predictive machine learning tool for asthma exacerbations: results from a 12-week, open-label study using an electronic multi-dose dry powder inhaler with integrated sensors. *J Asthma Allergy*. 2022;15:1623-1637. doi:10.2147/JAA.S377631
141. Bousquet J, Grattan CE, Akdis CA, et al. Highlights and recent developments in allergic diseases in EAACI journals (2019). *Clin Transl Allergy*. 2020;10(1):1-11. doi:10.1186/S13601-020-00366-3/METRICS
142. Baiardini I, Novakova S, Mihaicuta S, Oguzulgen IK, Canonica GW. Adherence to treatment in allergic respiratory diseases. *Expert Rev Respir Med*. 2018;13(1):53-62. doi:10.1080/17476348.2019.1554438
143. Bohlmann A, Mostafa J, Kumar M. Machine learning and medication adherence: scoping review. *JMIR Med*. 2021;2(4):e26993. doi:10.2196/26993
144. Tsang KCH, Pinnock H, Wilson AM, Ahmar SS. Application of machine learning to support self-management of asthma with mHealth. *Proc Ann Int Conf IEEE Eng Med Biol Soc*. 2020;2020:5673-5677. doi:10.1109/EMBC44109.2020.9175679
145. Chaix B, Bibault JE, Pienkowski A, et al. When chatbots meet patients: one-year prospective study of conversations between patients with breast cancer and a chatbot. *JMIR Cancer*. 2019;5(1):e12856. doi:10.2196/12856
146. Kadariya D, Venkataraman R, Yip HY, Kalra M, Thirunarayanan K, Sheth A. KBot: knowledge-enabled personalized chatbot for asthma self-management. *Proc IEEE Int Conf Smart Comput SMARTCOMP*. 2019;138-143. doi:10.1109/SMARTCOMP.2019.00043
147. Tsang KCH, Pinnock H, Wilson AM, Shah SA. Application of machine learning algorithms for asthma management with mHealth: a clinical review. *J Asthma Allergy*. 2022;15:855-873. doi:10.2147/JAA.S285742
148. Shim J, Kim M, Lee SM, et al. An artificial intelligence algorithm-based smartphone application for daily cough monitoring. *Allergy*. 2023;78:1378-1380. doi:10.1111/ALL.15632
149. Fakotakis ND, Nousias S, Arvanitis G, Zacharakis EI, Moustakas K. Revisiting audio pattern recognition for asthma medication adherence: evaluation with the RDA benchmark suite. 2022. Accessed January 9, 2023. <https://gitlab.com/vvr/monitoring-medication-adherence/rda-benchmark>
150. Alam MZ, Simonetti A, Brillantino R, et al. Predicting pulmonary function from the analysis of voice: a machine learning approach. *Front Digit Health*. 2022;4:750226. doi:10.3389/FDGH.2022.750226/FULL
151. Azari E, Renard H, Lauthier JB, et al. A laboratory evaluation of the new automated Pollen sensor Beenose: pollen discrimination using machine learning techniques. *Sensors*. 2023;23(6):2964. doi:10.3390/S23062964
152. Ambee Pollen API. Real-time and forecast pollen count data. Accessed June 20, 2023. <https://www.getambee.com/api/pollen>
153. Oteros J, Pusch G, Weichenmeier I, et al. Automatic and online Pollen monitoring. *Int Arch Allergy Immunol*. 2015;167(3):158-166. doi:10.1159/000436968
154. Zewdie GK, Liu X, Wu D, Lary DJ, Levetin E. Applying machine learning to forecast daily ambrosia pollen using environmental and NEXRAD parameters. *Environ Monit Assess*. 2019;191(2):1-11. doi:10.1007/S10661-019-7428-X/METRICS
155. Cordero JM, Rojo J, Gutiérrez-Bustillo AM, Narros A, Borge R. Predicting the Olea pollen concentration with a machine learning algorithm ensemble. *Int J Biometeorol*. 2021;65(4):541-554. doi:10.1007/S00484-020-02047-Z/METRICS
156. Kuruvilla ME, Lee EH, Lee GB. Understanding asthma phenotypes, endotypes, and mechanisms of disease. *Clin Rev Allergy Immunol*. 2019;56:219-233. doi:10.1007/s12016-018-8712-1
157. Henderson J, Granell R, Heron J, et al. Associations of wheezing phenotypes in the first 6 years of life with atopy, lung function and airway responsiveness in mid-childhood. *Thorax*. 2008;63(11):974-980. doi:10.1136/thx.2007.093187
158. Granell R, Sterne J, Savenije O, et al. Identification and replication of wheezing phenotypes using longitudinal latent class analysis. *Am Thorac Soc*. 2010;2010:A6242. doi:10.1164/ajrccm-conference.2010.181.1_MeetingAbstracts.A6242
159. Belgrave DCM, Simpson A, Semic-Jusufagic A, et al. Joint modeling of parentally reported and physician-confirmed wheeze identifies children with persistent troublesome wheezing. *J Allergy Clin Immunol*. 2013;132(3):575-583. doi:10.1016/j.jaci.2013.05.041
160. Sordillo JE, Coull BA, Rifas-Shiman SL, et al. Characterization of longitudinal wheeze phenotypes from infancy to adolescence in project viva, a prebirth cohort study. *J Allergy Clin Immunol*. 2020;145(2):716-719 e8. doi:10.1016/j.jaci.2019.10.026
161. Hallmark B, Wegienka G, Havstad S, et al. Chromosome 17q12-21 variants are associated with multiple wheezing phenotypes in childhood. *Am J Respir Crit Care Med*. 2021;203(7):864-870. doi:10.1164/rccm.202003-0820OC

162. Odling M, Wang G, Andersson N, et al. Characterization of asthma trajectories from infancy to young adulthood. *J Allergy Clin Immunol Pract.* 2021;9(6):2368-2376. doi:10.1016/j.jaip.2021.02.007
163. Savenije OE, Granell R, Caudri D, et al. Comparison of childhood wheezing phenotypes in 2 birth cohorts: ALSPAC and PIAMA. *J Allergy Clin Immunol.* 2011;127(6):1505.
164. Belgrave D, Simpson A, Custovic A. Challenges in interpreting wheeze phenotypes: the clinical implications of statistical learning techniques. *Am J Respir Crit Care Med.* 2014;189(2):121-123. doi:10.1164/rccm.201312-2206ED
165. Saria S, Goldenberg A. Subtyping: what it is and its role in precision medicine. *IEEE Intell Syst.* 2015;30(4):70-75.
166. Saglani S, Custovic A. Childhood asthma: advances using machine learning and mechanistic studies. *Am J Respir Crit Care Med.* 2019;199(4):414-422. doi:10.1164/rccm.201810-1956CI
167. Oksef C, Granell R, Mahmoud O, et al. Causes of variability in latent phenotypes of childhood wheeze. *J Allergy Clin Immunol.* 2019;143(5):1783-1790. doi:10.1016/j.jaci.2018.10.059
168. Oksef C, Granell R, Haider S, et al. Distinguishing wheezing phenotypes from infancy to adolescence. A pooled analysis of five birth cohorts. *Ann Am Thorac Soc.* 2019;16(7):868-876. doi:10.1513/AnnalsATS.201811-837OC
169. Haider S, Granell R, Curtin J, et al. Modeling wheezing spells identifies phenotypes with different outcomes and genetic associates. *Am J Respir Crit Care Med.* 2022;205(8):883-893. doi:10.1164/rccm.202108-1821OC
170. Haider S, Custovic A. Breaking Down Silos in Asthma Research: the Case for an Integrated Approach. *Innovations* 2019.
171. McCready C, Haider S, Little F, et al. Early childhood wheezing phenotypes and determinants in a south African birth cohort: longitudinal analysis of the Drakenstein child health study. *Lancet Child Adolesc Health.* 2023;7(2):127-135. doi:10.1016/S2352-4642(22)00304-2
172. Hu C, Duijts L, Erler NS, et al. Most associations of early-life environmental exposures and genetic risk factors poorly differentiate between eczema phenotypes: the generation R study. *Br J Dermatol.* 2019;181(6):1190-1197. doi:10.1111/bjd.17879
173. Lopez DJ, Lodge CJ, Bui DS, et al. Establishing subclasses of childhood eczema, their risk factors and prognosis. *Clin Exp Allergy.* 2022;52(9):1079-1090. doi:10.1111/cea.14139
174. Mulick AR, Mansfield KE, Silverwood RJ, et al. Four childhood atopic dermatitis subtypes identified from trajectory and severity of disease and internally validated in a large UK birth cohort. *Br J Dermatol.* 2021;185(3):526-536. doi:10.1111/bjd.19885
175. Paternoster L, Savenije OEM, Heron J, et al. Identification of atopic dermatitis subgroups in children from 2 longitudinal birth cohorts. *J Allergy Clin Immunol.* 2018;141(3):964-971. doi:10.1016/j.jaci.2017.09.044
176. Nakamura T, Haider S, Fontanella S, Murray CS, Simpson A, Custovic A. Modelling trajectories of parentally reported and physician-confirmed atopic dermatitis in a birth cohort study. *Br J Dermatol.* 2022;186(2):274-284. doi:10.1111/bjd.20767
177. Depner M, Fuchs O, Genuneit J, et al. Clinical and epidemiologic phenotypes of childhood asthma. *Am J Respir Crit Care Med.* 2014;189(2):129-138. doi:10.1164/rccm.201307-1198OC
178. Suaini NHA, Yap GC, Bui DPT, et al. Atopic dermatitis trajectories to age 8 years in the GUSTO cohort. *Clin Exp Allergy.* 2021;51(9):1195-1206. doi:10.1111/cea.13993
179. Ziyab AH, Mukherjee N, Zhang H, Arshad SH, Karmaus W. Sex-specific developmental trajectories of eczema from infancy to age 26 years: a birth cohort study. *Clin Exp Allergy.* 2022;52(3):416-425. doi:10.1111/cea.14068
180. Belgrave DC, Granell R, Simpson A, et al. Developmental profiles of eczema, wheeze, and rhinitis: two population-based birth cohort studies. *PLoS Med.* 2014;11(10):e1001748. doi:10.1371/journal.pmed.1001748
181. Clark H, Granell R, Curtin JA, et al. Differential associations of allergic disease genetic variants with developmental profiles of eczema, wheeze and rhinitis. *Clin Exp Allergy.* 2019;49(11):1475-1486. doi:10.1111/cea.13485
182. Howard R, Rattray M, Prosperi M, Custovic A. Distinguishing asthma phenotypes using machine learning approaches. *Curr Allergy Asthma Rep.* 2015;15(7):1-10. doi:10.1007/S11882-015-0542-0
183. Lazic N, Roberts G, Custovic A, et al. Multiple atopy phenotypes and their associations with asthma: similar findings from two birth cohorts. *Allergy.* 2013;68(6):764-770. doi:10.1111/ALL.12134
184. Krautenbacher N, Flach N, Böck A, et al. A strategy for high-dimensional multivariable analysis classifies childhood asthma phenotypes from genetic, immunological, and environmental factors. *Allergy.* 2019;74(7):1364. doi:10.1111/ALL.13745
185. Mersha TB, Afanador Y, Johansson E, et al. Resolving clinical phenotypes into endotypes in allergy: molecular and omics approaches. *Clin Rev Allergy Immunol.* 2021;60(2):200-219. doi:10.1007/S12016-020-08787-5/FIGURES/6
186. Kaur R, Chupp G. Phenotypes and endotypes of adult asthma: moving toward precision medicine. *J Allergy Clin Immunol.* 2019;144(1):1-12. doi:10.1016/J.JACI.2019.05.031
187. Wu W, Bang S, Bleecker ER, et al. Multiview cluster analysis identifies variable corticosteroid response phenotypes in severe asthma. *Am J Respir Crit Care Med.* 2019;199(11):1358-1367. doi:10.1164/RCCM.201808-1543OC/SUPPL_FILE/DISCLOSURES.PDF
188. Meng Y, Lou H, Wang Y, et al. Endotypes of chronic rhinitis: a cluster analysis study. *Allergy.* 2019;74(4):720-730. doi:10.1111/ALL.13640
189. Malizia V, Cilluffo G, Fasola S, et al. Endotyping allergic rhinitis in children: a machine learning approach. *Pediatr Allergy Immunol.* 2022;33(S27):18-21. doi:10.1111/PAI.13620
190. Robinson PFM, Fontanella S, Ananth S, et al. Recurrent severe preschool wheeze: from prespecified diagnostic labels to underlying endotypes. *Am J Respir Crit Care Med.* 2021;204(5):523-535. doi:10.1164/rccm.202009-3696OC
191. Ooka T, Raita Y, Fujiogi M, et al. Proteomics endotyping of infants with severe bronchiolitis and risk of childhood asthma. *Allergy.* 2022;77(11):3350-3361. doi:10.1111/ALL.15390
192. Tyler SR, Chun Y, Ribeiro VM, et al. Merged affinity network association clustering: joint multi-omic/clinical clustering to identify disease endotypes. *Cell Rep.* 2021;35(2):108975. doi:10.1016/J.CELREP.2021.108975
193. Azim A, Rezwani FI, Barber C, et al. Measurement of exhaled volatile organic compounds as a biomarker for personalised medicine: assessment of short-term repeatability in severe asthma. *J Pers Med.* 2022;12(10):1635. doi:10.3390/JPM12101635/S1
194. Han X, Krempski JW, Nadeau K, Parker SN, Kari NC. Advances and novel developments in mechanisms of allergic inflammation. *Allergy.* 2020;75(12):3100-3111. doi:10.1111/ALL.14632
195. Sikkema L, Ramírez-Suástegui C, Strobl DC, et al. An integrated cell atlas of the lung in health and disease. *Nat Med.* 2023;29(6):1563-1577. doi:10.1038/s41591-023-02327-2
196. Lotfollahi M, Naghipourfar M, Luecken MD, et al. Mapping single-cell data to reference atlases by transfer learning. *Nat Biotechnol.* 2022;40(1):121-130. doi:10.1038/S41587-021-01001-7
197. Kim H, Kim E, Lee I, Bae B, Park M, Nam H. Artificial intelligence in drug discovery: a comprehensive review of data-driven and machine learning approaches. *Biotechnol Bioprocess Eng.* 2020;25(6):895. doi:10.1007/S12257-020-0049-Y
198. Chung KF, Adcock IM. Precision medicine for the discovery of treatable mechanisms in severe asthma. *Allergy.* 2019;74(9):1649-1659. doi:10.1111/ALL.13771
199. Radzikowska U, Baerenfaller K, Cornejo-Garcia JA, et al. Omics technologies in allergy and asthma research: an EAACI position paper. *Allergy Eur J Allergy Clin Immunol.* 2022;77(10):2888-2908. doi:10.1111/ALL.15412

200. Ahmed Z, Mohamed K, Zeeshan S, Dong XQ. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database*. 2020;2020:baaa010. doi:10.1093/DATABASE/BAAA010
201. Lovrić M, Banić I, Lacić E, Pavlović K, Kern R, Turkalj M. Predicting treatment outcomes using explainable machine learning in children with asthma. *Children (Basel)*. 2021;8(5):376. doi:10.3390/CHILDREN8050376
202. Wu JJ, Hong C, Merola JF, et al. Predictors of nonresponse to dupilumab in patients with atopic dermatitis: a machine learning analysis. *Ann Allergy Asthma Immunol*. 2022;129(3):354-359.e5. doi:10.1016/J.ANAI.2022.05.025
203. Ray A, Das J, Wenzel SE. Determining asthma endotypes and outcomes: complementing existing clinical practice with modern machine learning. *Cell Rep Med*. 2022;3(12):100857. doi:10.1016/J.XCRM.2022.100857
204. Tanoli Z, Vähä-Koskela M, Aittokallio T. Artificial intelligence, machine learning, and drug repurposing in cancer. 2021. doi:10.1080/17460441.2021.1883585
205. Dara S, Dhamecherla S, Jadav SS, Babu CM, Ahsan MJ. Machine learning in drug discovery: a review. *ArtifIntellRev*. 2022;55(3):1947-1999. doi:10.1007/S10462-021-10058-4/TABLES/2
206. Staszak M, Staszak K, Wieszczycka K, Bajek A, Roszkowski K, Tylkowski B. Machine learning in drug design: use of artificial intelligence to explore the chemical structure–biological activity relationship. *Wiley Interdiscip Rev Comput Mol Sci*. 2022;12(2):e1568. doi:10.1002/WCMS.1568
207. Issa NT, Stathias V, Schürer S, Dakshanamurthy S. Machine and deep learning approaches for cancer drug repurposing. *Semin Cancer Biol*. 2021;68:132. doi:10.1016/J.SEMCANCER.2019.12.011
208. Abdulla A, Wang B, Qian F, et al. Project IDentif.AI: harnessing artificial intelligence to rapidly optimize combination therapy development for infectious disease intervention. *Adv Ther (Weinh)*. 2020;3:2000034. doi:10.1002/ADTP.202000034
209. Beck BR, Shin B, Choi Y, Park S, Kang K. Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. *Comput Struct Biotechnol J*. 2020;18:784-790. doi:10.1016/J.CSBJ.2020.03.025
210. Patrick MT, Raja K, Miller K, et al. Drug repurposing prediction for immune-mediated cutaneous diseases using a word-embedding based machine learning approach. *J Invest Dermatol*. 2019;139(3):683. doi:10.1016/J.JID.2018.09.018
211. Harrer S, Shah P, Antony B, Hu J. Artificial intelligence for clinical trial design. *Trends Pharmacol Sci*. 2019;40(8):577-591. doi:10.1016/J.TIPS.2019.05.005
212. He H, Liu B, Luo H, Zhang T, Jiang J. Big data and artificial intelligence discover novel drugs targeting proteins without 3D structure and overcome the undruggable targets. *Stroke Vasc Neurol*. 2020;5:323. doi:10.1136/svn-2019-000323
213. Rodrigues T, Bernardes GJL. Machine learning for target discovery in drug development. *Curr Opin Chem Biol*. 2019;2020:16-22. doi:10.1016/j.cbpa.2019.10.003
214. Liu B, He H, Luo H, Zhang T, Jiang J. Artificial intelligence and big data facilitated targeted drug discovery. *Stroke Vasc Neurol*. 2019;4:290. doi:10.1136/svn-2019-000290
215. FDA. Good machine learning practice for medical device development: guiding principles. Accessed January 3, 2023. <https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles>
216. Sounderajah V, Ashrafian H, Golub RM, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. *BMJ Open*. 2021;11(6):e047709. doi:10.1136/BMJOPEN-2020-047709
217. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1-W73. doi:10.7326/M14-0698
218. Collins GS, Dhiman P, Andaur Navarro CL, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open*. 2021;11(7):e048008. doi:10.1136/BMJOPEN-2020-048008
219. AI and NLP. Autoscriber is helping healthcare professionals better care for the patient. Accessed January 25, 2023. <https://www.ai.nl/artificial-intelligence/with-ai-and-nlp-autoscriber-is-helping-healthcare-professionals-better-care-for-the-patient/>
220. Regulatory framework proposal on artificial intelligence | Shaping Europe's digital future. Accessed January 11, 2023. <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
221. Legislation Related to Artificial Intelligence. Accessed January 11, 2023. <https://www.ncsl.org/technology-and-communication/legislation-related-to-artificial-intelligence>
222. Eczema App. Eczema signs and symptoms, eczema skin infection. Accessed June 22, 2023. <https://eczemaless.com/>
223. Ventura MT, Giuliano AFM, Buquicchio R, et al. Implementation of the MASK-air® App for rhinitis and asthma in older adults: MASK@Puglia pilot study. *Int Arch Allergy Immunol*. 2022;183(1):45-50. doi:10.1159/000518032
224. MASK-air. Accessed June 22, 2023. <https://www.mask-air.com/>
225. Sabry F, Eltaras T, Labda W, Alzoubi K, Malluhi Q. Machine learning for healthcare wearable devices: the big picture. *J Healthc Eng*. 2022;2022:4653923. doi:10.1155/2022/4653923
226. Greiwe J, Nyenhuis SM. Wearable technology and how this can be implemented into clinical practice. *Curr Allergy Asthma Rep*. 2020;20(8):36. doi:10.1007/S11882-020-00927-3
227. Lee CS, Lee AY. Clinical applications of continual learning machine learning. *Lancet Digit Health*. 2020;2(6):e279-e281. doi:10.1016/S2589-7500(20)30102-3
228. Vokinger KN, Feuerriegel S, Kesselheim AS. Continual learning in medical devices: FDA's action plan and beyond. *Lancet Digit Health*. 2021;3(6):e337-e338. doi:10.1016/S2589-7500(21)00076-5
229. Lekadir K, Osuala R, Gallin C, et al. FUTURE-AI: guiding principles and consensus recommendations for trustworthy artificial intelligence in medical imaging. 2021. doi:10.48550/arxiv.2109.09658
230. Norgeot B, Quer G, Beaulieu-Jones BK, et al. Minimum information about clinical artificial intelligence modeling: the MICLAIM checklist. *Nat Med*. 2020;26(9):1320-1324. doi:10.1038/S41591-020-1041-Y
231. Hernandez-Boussard T, Bozkurt S, Ioannidis JPA, Shah NH. MINIMAR (MINimum information for medical AI reporting): developing reporting standards for artificial intelligence in health care. *J Am Med Inform Assoc*. 2020;27(12):2011-2015. doi:10.1093/JAMIA/OCAA088

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: van Breugel M, Fehrmann RSN, Bügel M, et al. Current state and prospects of artificial intelligence in allergy. *Allergy*. 2023;00:1-21. doi:10.1111/all.15849