Medical Science

To Cite:
Gajowniczek K, Błaszkiewicz B, Paprocki J, Cyzio M, Chrzanowska N, Nowakowska M, Szustek OG, Łazicka P. Artificial Intelligence in Clinical Diagnostics and Medical Education: Current Applications, Challenges and Future Perspectives - A Narrative Review. *Medical Science* 2026; 30: e206ms3969
doi: https://doi.org/10.54905/disssi.v30i176.e206ms3969

Authors' Affiliation:
[1]Student at the Faculty of Medicine, Maria Sklodowska-Curie Medical Academy in Warsaw, Poland, Mazovia, PL
[2]Medical University of Warsaw: Warsaw, Mazovia, PL

***Corresponding author:**
Katarzyna Gajowniczek,
Student at the Faculty of Medicine, Maria Sklodowska-Curie Medical Academy in Warsaw, Poland, Mazovia, PL
https://orcid.org/0009-0001-9576-2547
email: lokasiagajowniczek@gmail.com

Peer-Review History
Received: 16 July 2026
Reviewed & Revised: 03/August/2026 to 16/October/2026
Accepted: 23 October 2026
Published: 30 October 2026

Peer-review Method
External peer-review was done through double-blind method.

Medical Science
pISSN 2321–7359; eISSN 2321–7367

© The Author(s) 2026. Open Access. This article is licensed under a Creative Commons Attribution License 4.0 (CC BY 4.0)., which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

DISCOVERY SCIENTIFIC SOCIETY

Artificial Intelligence in Clinical Diagnostics and Medical Education: Current Applications, Challenges and Future Perspectives - A Narrative Review

Katarzyna Gajowniczek[1]*, Bianka Błaszkiewicz[1], Jakub Paprocki[1], Magdalena Cyzio[1], Nicola Chrzanowska[1], Maria Nowakowska[1], Oliwia Gabriela Szustek[1], Patrycja Łazicka[2]

ABSTRACT

Background: Artificial intelligence (AI) is increasingly used in areas such as clinical diagnostics, medical imaging and clinical decision support. Successful implementation needs thorough clinical validation and sufficient AI literacy among medical professionals. *Aim*: This narrative review aims to summarize the current clinical uses of AI, examine its role in medical practice and medical education and critically evaluate the practical, ethical, and methodological challenges associated with its implementation. *Materials and methods*: A narrative review of the literature was performed, concentrating on the development of AI technologies, the performance of these technologies in diagnosis across several medical specialties, the usefulness of generative AI and large language models in healthcare, and the frameworks currently in use for clinical validation. *Result and Discussion*: AI systems, in particular convolutional neural networks, have reached a diagnostic performance comparable to that of specialists in radiology, dermatology, and ophthalmology, while the use of machine learning with electronic health records allows for predictive modeling and the development of precision medicine. Generative AI and large language models have potential in clinical documentation, diagnostics, and customized medical education. At the same time, problems such as bias, lack of interpretability, hallucinations, and data privacy problems continue. *Conclusion*: Artificial intelligence has the potential to support doctors rather than replace them. Successful implementation requires continuous innovation, responsible management, and the development of AI-related competencies ensuring that healthcare remains safe, effective, and focused on the patient.

Keywords: Artificial Intelligence; Clinical Decision Support Systems; Diagnostic Imaging; Medical Education; Medical Ethics.

1. INTRODUCTION

The use of computer technology in clinical medicine is by no means a new development; nevertheless, the structure, capabilities, and clinical applications of artificial intelligence (AI) have experienced considerable changes over the last few decades. In the 1970s and 1980s, the early medical AI systems primarily made use of expert systems and rule-based algorithms to help with clinical decision-making (Kaul et al., 2020). An early example was the MYCIN system, which was developed to support the diagnosis of bacterial infections and to recommend treatments by means of rules established by human experts (Kaul et al., 2020; Davenport and Kalakota, 2019). Even though these systems were a major achievement in the field of medical computing, their application in a clinical setting was restricted due to their reliance on rules that had to be manually defined and their inability to deal with the intricacy and variability found in actual healthcare environments (Davenport and Kalakota, 2019). A major turning point in medical AI came with the advent of modern machine learning (ML) and deep learning (DL) methods (Yu et al., 2018).

Owing to greater computational power, the continuously increasing availability of digital health data, and progress in algorithm development, artificial neural networks can now detect complex patterns in large datasets without having to depend entirely on manually designed features (Yu et al., 2018; Topol, 2019). Specifically, convolutional neural networks (CNNs) have shown great promise in the area of medical image analysis and, as a result, have helped lead to the development of AI-based solutions which achieve performance comparable to that of human specialists on certain, well-defined, diagnostic tasks (Topol, 2019).

At the present time, AI applications are being quickly extended across several clinical fields, including diagnostic imaging, clinical decision support, and predictive analytics (Yu et al., 2018; Topol, 2019; Briganti and Le Moine, 2020). There is the possibility that such technologies could improve diagnostic correctness, boost process efficiency, and help doctors in interpreting complicated clinical information (Topol, 2019; Briganti and Le Moine, 2020). However, to ensure a successful introduction of AI into clinical practice, it is necessary to carry out technological development as well as to perform proper validation, to integrate the technologies into medical systems, and to consider the pragmatic constraints associated with real-world implementation (Briganti and Le Moine, 2020; He et al., 2019). Although the number of AI-based medical applications is increasing, significant difficulties remain in introducing these advances into healthcare (He et al., 2019).

In order for AI tools to be used in clinical practice, healthcare professionals must understand both the capabilities and the limitations of these tools as well as know how to correctly interpret the outputs of the algorithms (He et al., 2019; Panch et al., 2019). If healthcare professionals do not have a good grasp of the principles of AI, there is a risk that they will place inappropriate trust in the algorithmic recommendations or will find it difficult to assess their clinical relevance (He et al., 2019). Filling the gap between technological and scientific progress and its application in clinical settings is therefore a major challenge for modern medicine to achieve safe and effective collaboration between healthcare professionals and AI-based systems, it is necessary to carry out a detailed evaluation of current AI applications, the difficulties associated with their implementation, and developments.

The purpose of this review is threefold: First, to provide a summary of the current use of artificial intelligence in the various diagnostic specialties; Second, to examine the possible role of AI in assisting with medical practice and in improving the delivery of healthcare; Third, to critically analyze the practical, ethical, and methodological challenges connected with the implementation of AI in clinical medicine, including worries concerning reliability, transparency, and responsible use.

2. REVIEW METHODS

This is a narrative review that collects the existing evidence regarding the applications, the difficulties, and the educational effects of artificial intelligence (AI) in clinical medicine. A narrative review has been chosen as the method since it allows the evidence from various viewpoints clinical, technological, educational, and ethical to be combined and thus gives a full overview of this field, which is developing very quickly. To find the most pertinent scientific evidence, a structured literature search was carried out.

Search Strategy and Data Sources: A systematic literature search was conducted using the main biomedical and technological databases. The search was performed in Europe PMC (PubMed/MEDLINE records) and covered articles published between 1 January 2016 and 28 February 2026. Inclusion and Exclusion Criteria: Predefined eligibility criteria were used to ensure the inclusion of clinically relevant and scientifically sound evidence. The conditions for being included are the following: (1) original research articles that have been subject to peer review, together with randomized controlled trials, studies on diagnostic validity, systematic reviews, meta-analyses, and consensus or narrative reviews; (2) studies that examine the clinical application of AI, the validation of diagnostic algorithms, generative AI, or large language models in healthcare and in medical education; and (3) articles that were published in English.

The conditions for excluding studies are as follows: (1) those which focus entirely on theoretical computational models as well as have neither any clinical nor any educational value; (2) non-peer-reviewed preprints and conference abstracts which do not give sufficient details about the method, together with editorials, opinion articles and personal comments; and (3) duplicate publications or studies which present data that are substantially overlapping. Literature Screening and Evidence Synthesis. To determine whether the publications obtained were relevant to the review's aims, their titles and abstracts were looked at; those which might have been eligible were then assessed according to their full text to judge their relevance, methodological quality and contribution to the thematic scope of the review. The selection process is presented in a flow diagram adapted from PRISMA 2020 (Figure 1). As it was designed as a narrative review, the study did not carry out a formal quantitative synthesis or include a risk-of-bias assessment. Instead, the literature that was chosen was analyzed qualitatively and grouped within thematic sections that correspond to the objectives of the review, for example, the use of AI in clinical diagnostics; generative AI and natural language processing in healthcare; AI in medical education; the current challenges concerning implementation; ethical issues; and the prospects of AI-assisted healthcare.

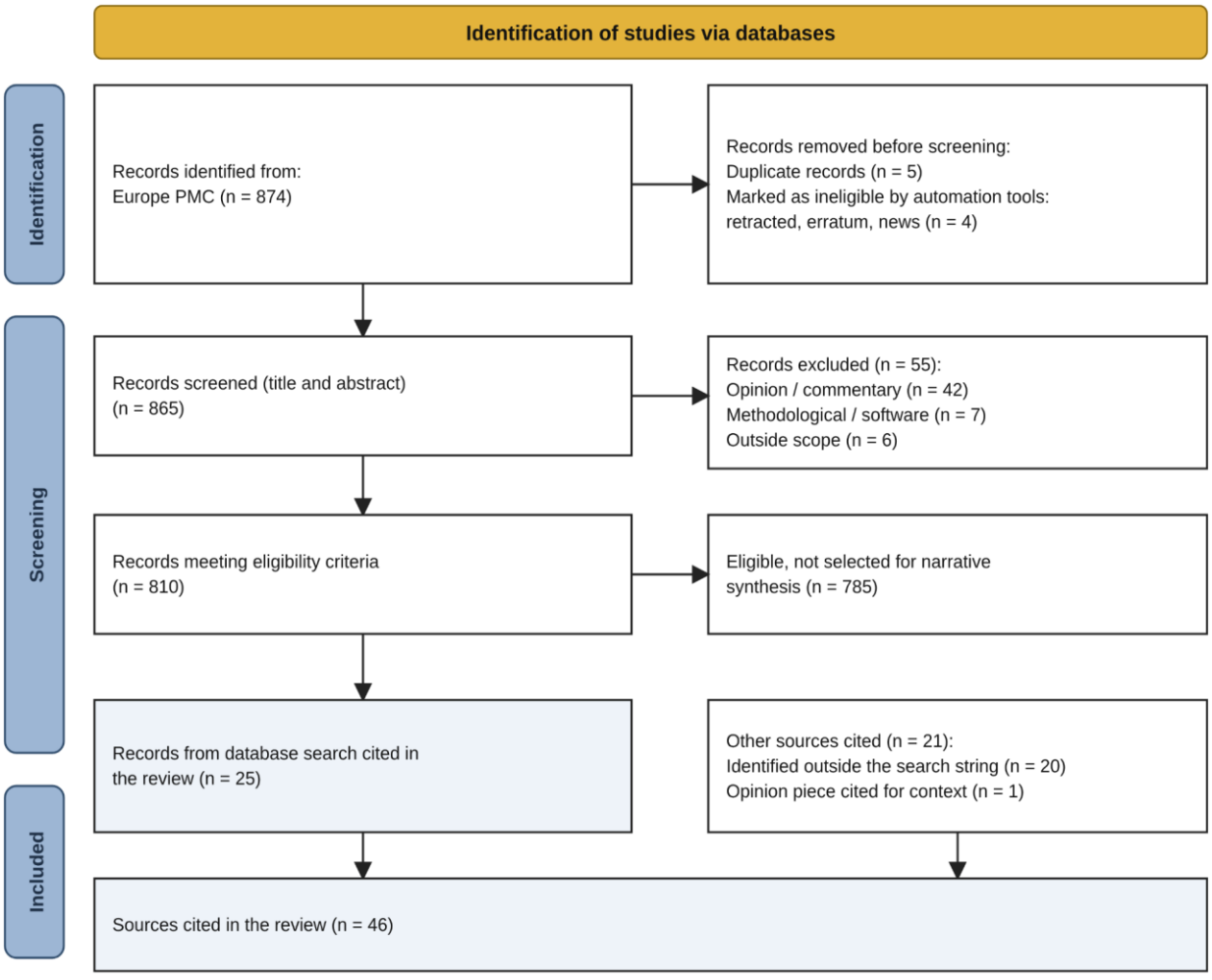


Figure 1. Flow diagram of the literature search and source selection

3. RESULTS AND DISCUSSION

Artificial Intelligence in Clinical Diagnostics

The use of artificial intelligence (AI) in clinical diagnostics is among the most rapidly developing fields in modern medicine, and the shift from conventional rule-based methods to more advanced data-driven techniques has allowed AI systems to analyze complicated clinical information and assist with diagnosis throughout various medical fields.

Radiology and Diagnostic Imaging:

The growing involvement of AI in diagnostic imaging has mainly resulted from progress in deep learning, especially from convolutional neural networks (CNNs), since these are very good at studying the complex dimensional patterns found in medical images. Using CNN methods, relevant imaging features can be automatically picked out from data at the pixel and voxel levels, which in turn allows for a more reproducible interpretation of the images (Lundervold and Lundervold, 2019; Anwar et al., 2018).

Screening and Lesion Detection:

As far as mammography is concerned, artificial intelligence systems have exhibited promising results in the detection of breast cancer. Studies which have compared the performance of AI-based analysis with that of radiologists have found that standalone AI systems can reach a level of diagnostic performance similar to that of experienced specialists and thus support their use in breast cancer screening procedures. The use of AI to assist with the interpretation may help preserve the same level of diagnostic performance while at the same time reducing the amount of workload involved in assessing the images (Rodriguez-Ruiz et al., 2019).

Deep learning techniques have also been studied with a view to analyzing a wide variety of medical imaging methods, including the detection and characterization of pathological findings. These methods could increase the efficiency and consistency of image interpretation by helping clinicians to identify relevant clinical patterns among large amounts of imaging data (Langlotz, 2019). Image Segmentation and Magnetic Resonance Imaging (MRI) AI also has great potential in the area of automated image segmentation, which means identifying and outlining anatomical structures or pathological areas in medical images. In MRI, deep learning techniques have improved quantitative image analysis and, as a result, helped assess anatomical structures and disease-related changes with greater reproducibility (Lundervold and Lundervold, 2019; Anwar et al., 2018).

Dermatology and Ophthalmology:

Image-based specialties have become important areas for developing and evaluating AI diagnostic systems because standardized digital images and large patient datasets are available.

Dermatological Classification:

In the field of dermatology, deep convolutional neural networks trained on dermoscopic images can classify skin lesions, for instance by telling the difference between benign lesions and malignant melanoma as well as other forms of skin cancer. It has been shown that the performance of these models is comparable to that of board-certified dermatologists in respect to specific image classification tasks, which shows their potential as supportive diagnostic tools (Esteva et al., 2017).

Ophthalmological Screening:

In the field of ophthalmology, AI systems have been created to examine fundus photographs and identify diabetic retinopathy as well as other retinal abnormalities. A key clinical trial showed that an autonomous AI system was effective in detecting diabetic retinopathy in primary care settings, indicating that these technologies could help in the earlier identification of patients who need more ophthalmological evaluation (Abràmoff et al., 2018).

Digital Pathology:

The move from traditional pathology to digital pathology, achieved through the use of Whole Slide Images (WSIs), has introduced new potentials for the computational analysis of tissue specimens. Since digital pathology produces large quantities of complex image data, it poses difficulties for conventional manual evaluation and thus amounts to a significant field for analysis with the aid of AI (Niazi et al., 2019).

Detection of Metastases and Histological Assessment:

Research into deep learning algorithms has been carried out with the aim of enabling the automated analysis of digitized histopathological images, especially in relation to the identification of malignant tissue patterns and as a basis for cancer diagnosis (Niazi et al., 2019; Bera et al., 2019). In oncology, AI-based methods can help pathologists improve the consistency and speed of tissue assessment. AI systems are being used increasingly in histological grading and prognostic assessment. In the case of prostate cancer, deep learning models have shown they can assist with the diagnosis and grading of biopsy specimens, including Gleason scoring

(Ström et al., 2020). Such technologies could help to achieve greater standardization and reproducibility in pathological evaluation (Bera et al., 2019; Ström et al., 2020).

Clinical Decision Support Systems (CDSS) and Predictive Modeling:

In addition to being applied to medical imaging, methods of artificial intelligence based on machine learning and natural language processing (NLP) are increasingly being used for the analysis of both the structured and unstructured information contained in electronic health records (EHRs) (Rajkomar et al., 2018; Shickel et al., 2018). With these techniques, it is possible to detect complex patterns in clinical data and to assist in the development of prediction models for a variety of healthcare applications.

Prediction of Clinical Outcomes:

Predictive systems using artificial intelligence have been developed to evaluate the risk of undesirable clinical outcomes through employing routinely collected patient data, such as laboratory results, demographic characteristics, and other clinical variables (Rajkomar et al., 2018; Shickel et al., 2018). Methods that apply deep learning to electronic health record data have shown promise in forecasting clinically important events and influencing decision-making inside medical systems (Rajkomar et al., 2018).

Oncology and Exact Medicine:

A growing application of AI is to assist in the field of precision oncology through analyzing both molecular and clinical data. By combining information from molecular profiling with state-of-the-art computational techniques, it may be possible to identify biomarkers that have clinical relevance and to gain a better understanding of tumor characteristics. Methods that involve the use of AI could help in the development of customized treatment strategies by aiding in the identification of molecular features linked to disease behavior and response to therapy (Malone et al., 2020).

Generative AI and large language models in healthcare

The advances recently achieved in the application of artificial intelligence to medicine are due to the development of generative AI and large language models (LLMs). In contrast to traditional machine learning systems, which are generally created for specific and predetermined tasks, LLMs are more general in nature and can thus handle and produce complex textual information in a wide range of fields, including in healthcare applications (Clusmann et al., 2023; Thirunavukarasu et al., 2023). Because they are based on transformer architectures, these models generate responses by detecting statistical relationships in substantial datasets, which allows them to produce text that is relevant from a clinical point of view (Clusmann et al., 2023).

Development of General-Purpose and Medical Base Models:

The emergence of advanced models like GPT-4 represented a major advance in the development of medical AI, since it showed that large language models are able to deal with intricate medical data, understand clinical text, and perform well in a variety of medical knowledge tests without demanding extensive adaptations for each particular task (Clusmann et al.,2023; Thirunavukarasu et al., 2023). Research together with that concerning general-purpose models has increasingly been directed towards the development of specialized medical foundation models for use in healthcare (Thirunavukarasu et al., 2023; Moor et al., 2023).

For instance, Med-PaLM shows the potential of large language models modified for medical use in producing responses that are relevant from a clinical point of view and in helping with tasks that demand medical knowledge (Singhal et al., 2023). Such systems could make use of more optimization techniques, including instruction tuning and methods based on human feedback, to improve their suitability for medical applications (Thirunavukarasu et al., 2023). However, despite their promising capabilities, medical LLMs still require assessment of their reliability, accuracy and safety (Clusmann et al., 2023; Thirunavukarasu et al., 2023; Moor et al., 2023; Singhal et al., 2023).

Natural Language Processing (NLP) in Clinical Documentation:

A potential application of large language models in the field of healthcare is to aid with the management and interpretation of clinical records. By making use of advanced natural language processing (NLP), generative models can assist in the analysis of unstructured medical text, in the extraction of relevant clinical information, and in organizing the complex patient data found in electronic health records (Emanuel and Wachter, 2019).

There is the possibility that language models could assist healthcare professionals by helping them to summarize clinical information and assisting in the preparation of medical documentation such as clinical summaries and referral documents (Clusmann et al., 2023; Moor et al., 2023). Although such applications might lead to greater efficiency in workflow, further assessment is needed to assess their accuracy, safety, and effect on clinical practice.

Diagnostic Competencies and Clinical Reasoning Support:

Generative AI models are likewise becoming more and more regarded as possible aids to clinical reasoning and to the process of making diagnostic decisions, in addition to those involving documentation. It has been found that when presented with complex clinical scenarios, LLMs are capable of coming up with possible differential diagnoses and providing structured explanations on the basis of the clinical information available (Thirunavukarasu et al., 2023; Rao et al., 2023). In contrast to conventional decision-support systems which rely on pre-defined algorithms, large language models can analyze unformatted clinical texts that contain symptoms, medical history and laboratory findings and may thus help clinicians to take into account a wider variety of possible diagnoses (Thirunavukarasu et al., 2023; Rao et al., 2023). Such models could act as supportive tools that enhance physicians' expertise by supplying extra viewpoints when engaged in clinical reasoning. However, since the outputs have limitations such as the production of inaccurate information, the presence of uncertainty and the fact that they carry no direct clinical responsibility, they need to be subject to professional verification (Clusmann et al., 2023; Rao et al., 2023).

Artificial Intelligence in Medical Education

The greater the use of artificial intelligence (AI) in clinical practice, the more important it is to advance medical education in parallel with the skills that upcoming healthcare professionals will require. In the past, traditional medical education has placed a strong emphasis on gaining and holding a large quantity of clinical knowledge; but in view of the increasing availability of AI-based tools, there is now a need to extend our educational focus to encompass digital competence, the ability to critically assess information and effective interaction with new technologies (Wartman and Combs, 2018).

Curriculum Transformation and New Digital Competencies:

The shift from the traditional 'information age' to an age increasingly determined by artificial intelligence means that the undergraduate and postgraduate medical courses need to be adapted (Wartman and Combs, 2018; Masters, 2024). Instead of concentrating only on memorizing factual information, medical education nowadays should include the skills that allow doctors to properly assess and use AI-assisted clinical tools (Wartman and Combs, 2018). Healthcare professionals need to develop new digital competencies (Masters, 2024).

Medical students and doctors should gain a basic understanding of medical informatics and the principles of artificial intelligence. It is important that future clinicians be able to critically evaluate information produced by AI, identify possible sources of error, and know the situations in which algorithmic recommendations might not apply to individual patients (Masters, 2024; McCoy et al., 2020). The role of the physician may gradually come to include the assessment and integration of AI-assisted information while still exercising clinical judgment and assuming responsibility for patient care.

LLMs as Teaching Assistants and Personal Learning Tools:

The development of large language models (LLMs) has led to talks about their possible use in medical education. Research involving models such as ChatGPT has shown that they are able to attain a high level of performance with medical knowledge tests, including the United States Medical Licensing Examination (USMLE), thereby indicating that they could be employed in an educational context (Kung et al., 2023). LLMs are becoming more and more regarded as interactive learning tools which can help with personalized medical education (Khan et al., 2023). By means of conversational interaction they can help students by explaining complicated concepts, producing educational materials and allowing them to take part in independent learning. Nevertheless, their use in medical education It needs proper supervision because of the limitations concerning accuracy, reliability, and the likelihood of producing incorrect information (Kung et al., 2023; Khan et al., 2023).

Simulations, Virtual Patients, and OSCE Preparation:

Generative AI has also opened up new opportunities in the field of interactive medical simulations and virtual patient-learning environments. Different from conventional computer-based simulations which usually depend on pre-determined scenarios, AI-supported systems can provide more flexible communications and dynamic responses when training in a clinical setting (Masters, 2024; Han et al., 2019). Such technologies would give students more chances to improve their skills in communication, history-taking, and clinical reasoning within a controlled educational setting (Masters, 2024; McCoy et al., 2020; Han et al., 2019). Virtual patients powered by AI could supplement current simulation-based education by allowing students to obtain experience through numerous interactions with a wide variety of clinical scenarios before meeting real patients. However, further research will be needed if the effectiveness of these tools is to be evaluated, if the most appropriate methods of implementing them are to be established, and if the effect they have on the development of clinical competence is to be determined.

Current Challenges, Limitations, and Ethical Concerns

There is great potential in using artificial intelligence for clinical diagnosis and in medical education since it could improve the delivery of healthcare; at the same time, it does so by bringing with it a number of important methodological, ethical, legal, and practical challenges. For AI-based technologies to be introduced safely, transparently, and responsibly into healthcare systems it is necessary to overcome these limitations.

Algorithmic Bias and Dataset Representativeness (AI Bias):

Machine learning models do not operate independently of the data and design choices that go into their development, since their performance is significantly influenced by the quality, structure, and representativeness of the training datasets as well as by the method employed to define clinical outcomes when the model is being developed. If the datasets lack sufficient demographic or clinical diversity or if the proxy outcome labels do not correctly reflect the real clinical construct, AI systems may replicate or worsen the existing inequalities in healthcare data (Obermeyer et al., 2019; Larrazabal et al., 2020). Algorithmic bias can occur through a variety of different mechanisms. For example, where certain groups such as a particular gender or ethnic group are underrepresented in the training data, it has been found that the resulting classifiers display different diagnostic performance across various patient subgroups (Larrazabal et al., 2020).

Furthermore, bias can also result from the choice of proxy outcome labels when training a model. In a well-known case, a commercial healthcare risk prediction algorithm used historical healthcare expenditure as a proxy for clinical need and therefore systematically underestimated the health needs of Black patients when compared with White patients who had similar levels of illness since less money had previously been spent on their care (Obermeyer et al., 2019). This example shows that algorithmic bias does not arise solely from the statistical properties of training data but also from the conceptual decisions made during model development. It is important clinically to make this distinction since the problems of dataset imbalance and proxy-label bias call for different approaches to be taken. When it comes to dataset imbalance, improvement can be achieved by increasing the diversity and representativeness of the training datasets, whereas with proxy-label bias it is necessary to carefully verify that the proxy outcome labels chosen do in fact correspond to the intended clinical objective. Accordingly, to ensure responsible AI deployment it is not enough simply to use training datasets that are diverse and representative.

The "Black Box" Problem and Explainable AI (XAI):

Even though advanced artificial intelligence systems, particularly deep learning models, can achieve a high degree of predictive performance, they are still difficult to interpret. This inability to interpret, which is referred to as the "black box" problem, is a major barrier to winning trust and for these AI technologies being accepted in a clinical environment (Amann et al., 2020). Even though AI models can detect clinically significant patterns, the reasons for their predictions are generally difficult for healthcare professionals to understand. This leads to problems when it comes to properly interpreting the recommendations produced by AI and to integrating these systems into physicians' decision-making processes (Amann et al., 2020). To deal with this problem, the field of Explainable Artificial Intelligence (XAI) has come up with a number of methods whose aim is to improve the interpretability of AI models. For example, techniques based on saliency maps and other visualization methods are intended to highlight the features that contribute to the models' predictions (Amann et al., 2020; Ghassemi et al., 2021). Yet the existing explainability techniques are still limited and might not offer a full understanding of the decision-making processes taking place in complex AI systems (Ghassemi et al., 2021).

Hallucinations and Misinformation in Large Language Models:

Even though large language models (LLMs) have promising capabilities when it comes to medical applications, their probabilistic nature implies that there is a risk of them giving out inaccurate or unverified information, a scenario which is generally referred to as AI hallucination. This aspect of the limitation is particularly important in the area of medical education and in clinical decision support, as inaccurate information cannot easily be detected without the relevant medical expertise and verification. It is wrong to deny their potential usefulness, but LLM-based systems have to be carefully supervised, validated and be responsibly incorporated into medical workflows if the risks arising from incorrect outputs are to be reduced (Alkaissi and McFarlane, 2023).

Data Privacy, Security, and GDPR Compliance:

The creation of effective AI systems in the healthcare sector involves obtaining large amounts of clinical data and this in turn leads to problems concerning privacy protection, data governance and regulatory compliance (Price and Cohen, 2019; Minssen et al., 2020). Since patient health information is a kind of very sensitive data, it is necessary to have suitable safeguards when it is being collected, stored, processed and used in model development. Even though anonymization and pseudonymization techniques are frequently employed, progress in data analysis has led to continuing worries regarding the possibility of re-identifying patients and the potential for misuse of medical information (Price and Cohen, 2019). Furthermore, the use of clinical datasets in the development of AI raises complex issues regarding data ownership, informed consent, regulatory supervision, and the way in which previously collected patient information is handled in constantly changing technological systems (Minssen et al., 2020).

Reporting Standards and Clinical Validation Methodologies:

A big difficulty in getting AI research into clinical practice is ensuring that the safety, effectiveness, and real-world usefulness of the systems are carefully evaluated. The fact that most AI systems require extensive clinical validation before they can be widely adopted means that their performance has to be tested in real healthcare settings. International reporting standards have been set up in order that the transparency and methodological quality of research into artificial intelligence in clinical studies may be improved. The CONSORT-AI extension provides recommendations on how to report randomized clinical trials that involve AI-based interventions, ensuring that the key aspects of algorithm performance, human interaction, and clinical outcomes are adequately described (Liu et al., 2020).

The DECIDE-AI framework focuses on the first clinical assessment of AI-based decision-support systems, with a particular emphasis on safety assessment, usability, and the interaction between healthcare professionals and AI technologies before these systems are widely implemented (Vasey et al., 2022).

In order that AI-based medical technologies may be subject to rigorous scientific evaluation and be introduced into clinical practice on the basis of reliable evidence, it is essential to adhere to these frameworks.

Future Perspectives

Towards 2030 and Beyond Over the next ten years in the field of healthcare, medical artificial intelligence is expected to evolve from comprising a number of separate, task-specific applications to becoming more integrated and involving a wider variety of different types; instead of just enhancing the current digital tools, new AI technologies could modify clinical workflows, help towards more personalized decision-making, and affect certain aspects of the relationship between doctors and patients.

The Rise of Generalist Medical Artificial Intelligence (GMAI)

A major technological advancement is the shift from narrow AI to generalist, multimodal foundation models that can handle various kinds of medical information within a single framework (Moor et al., 2023; Singhal et al., 2023; Rajpurkar et al., 2022; Tu et al., 2024). Unlike the task-specific algorithms currently in use, which are generally developed and tested for specific applications such as the interpretation of chest radiographs, the analysis of electronic health records (EHR) or the classification of genomic variants, generalist medical AI is designed to combine heterogeneous clinical data to aid with comprehensive decision-making (Rajpurkar et al., 2022; Tu et al., 2024).

In the future, such systems may be capable of simultaneously analyzing various types of patient data, for example, high-resolution diagnostic imaging (such as CT and MRI), whole-slide digital pathology images, unstructured clinical notes and family histories taken from electronic health records (EHRs), continuous physiological monitoring data, and high-dimensional multi-omic datasets like

genomics and transcriptomics. Rather than generating separate outputs for each individual task, multimodal models could provide clinicians with integrated diagnostic assessments together with therapy recommendations based on sound evidence and tailored to the particular features of each patient. Most importantly, the systems are expected to function as clinical decision-support tools rather than as independent decision makers, and the ultimate responsibility will still lie with healthcare professionals (Rajpurkar et al., 2022; Tu et al., 2024).

The Augmented Physician and the Return to Human-Centred Care:

Worries about artificial intelligence taking the place of doctors have slowly been replaced by a more cooperative approach, which is usually referred to as the Augmented Physician (Merlo et al., 2026; Verghese et al., 2018). In this approach, AI is seen as a cognitive support system that enhances rather than replaces clinical expertise by automating routine administrative and documentation work and by helping with the synthesis of complex information (Merlo et al., 2026).

The recent progress in generative AI and ambient clinical intelligence indicates that such technologies could greatly decrease the amount of work involved in carrying out real-time clinical documentation, medical coding, administrative reporting, and the preparation of referral and discharge documents. Yet it is still believed that human review and supervision will have to remain essential to guarantee accuracy and patient safety (Merlo et al., 2026; Verghese et al., 2018; Sinsky et al., 2016). AI has the possibility of reducing the administrative workload on doctors so that they can allocate more time to direct patient care, to communication, and to shared decision-making. This would mean that the doctor's role could slowly change from that of the main collector of data to that of an empathetic communicator, a clinical integrator, and an expert who assesses the recommendations produced by AI (Merlo et al., 2026; Verghese et al., 2018).

However, in order to achieve this vision it will be necessary not only to make further technological progress but also to ensure transparency, to keep appropriate clinical supervision, and to develop a properly calibrated trust in decisions based on AI (Hatherley, 2020). In

this context, AI should be regarded as a tool that boosts human expertise while at the same time retaining the key principles of clinical judgment, ethical responsibility, and compassionate patient care.

Summary

The rapid introduction of artificial intelligence into clinical practice and medical education represents one of the most significant developments currently occurring in modern healthcare. Artificial intelligence is no longer merely serving as an additional technological tool; it is now becoming a vital component in clinical decision support, in the integration of data, and in the provision of healthcare. Because AI technologies are still in development, they are expected to affect not only the diagnostic and therapeutic processes but also the skills that will be required of future healthcare professionals. A summary of the current and emerging AI technologies, together with their main clinical applications and the present operational obstacles, is given in Table 1.

Table 1. Summary of AI Applications, Technological Frameworks, and Challenges in Modern Medicine.

Clinical / Educational Domain	Primary AI Technologies	Key Applications & Benefits	Core Limitations & Challenges
Diagnostic Imaging & Radiology	Convolutional Neural Networks (CNNs), Deep Learning (DL)	• Breast cancer screening (mammography) • Lesion detection & segmentation • Automated MRI quantification	• Dataset imbalances (demographic bias) • High inter-observer variability in validation
Surface & Retinal Specialties	Deep CNNs, Autonomous AI Classifiers	• Skin lesion classification (melanoma) • Diabetic retinopathy screening in primary care	• Sensitivity to imaging conditions (lighting, camera type) • Limited out-of-distribution generalizability
Digital Pathology	Deep Learning on Whole Slide Images (WSI)	• Automated histopathological analysis	• Technical artifacts in tissue preparation

		• Malignant pattern identification • Prostate cancer grading (Gleason scoring)	• Extremely large file sizes requiring high computing power
Clinical Decision Support & EHRs	Machine Learning (ML), Natural Language Processing (NLP)	• Risk estimation for adverse outcomes • Predictive modeling based on clinical variables • Precision oncology biomarker discovery	• Data fragmentation across different hospital systems • Patient re-identification risks
Generative AI & LLMs	Transformer Architectures (e.g., GPT-4, Med-PaLM)	• Automated clinical documentation & summaries • Differential diagnosis generation • Structuring unstructured clinical narratives	• AI Hallucinations (fabricated facts/references) • Lack of direct clinical responsibility • Data privacy & GDPR compliance
Medical Education	Conversational LLMs, Dynamic Simulation Engines	• Personalized, self-directed learning tools • AI-driven virtual patients for history-taking • Interactive OSCE preparation	• Risk of students over-relying on unverified AI outputs • Rigid or outdated institutional medical curricula
Future Ecosystems (Towards 2030)	Generalist Medical AI (GMAI), Multimodal Foundation Models	• Simultaneous integration of imaging, EHRs, and multi-omic datasets • Realization of the "Augmented Physician" paradigm	• Transitioning from narrow lab algorithms to real-world clinical deployment • Calibrating human-AI trust

4. CONCLUSION

When all the evidence examined in this article is taken into account, it becomes clear that the conventional debate about whether artificial intelligence will replace doctors' rests on a false dilemma. In fact, the points made above point to a more subtle conclusion: AI is more likely to enhance than to replace medical practitioners, since in all the applications looked at it seems capable of supporting complex decision-making, helping to improve diagnostic accuracy in certain areas, and reducing the burden of repetitive, non-clinical tasks. When the literature reviewed here is considered as a whole, it indicates that the healthcare system in the future will most probably be based on a collaborative model involving augmented intelligence.

Acknowledgments

There are no acknowledgments to disclose.

Authors' Contributions

Conceptualization: KG;
Methodology: KG, BB, JP;
Investigation: KG, BB, JP, MC, MN, NC, OS, PL;
Resources: KG, BB, JP, MC, MN, NC, OS, PL;
Writing original draft preparation: KG, BB, JP, MC, MN, NC, OS, PL;
Writing review and editing: KG, BB, JP, MC, MN, NC, OS, PL;
Visualization: KG;
Supervision: KG.

All authors have read and agreed to the published version of the manuscript.

Informed consent

Not applicable.

Ethical approval

Not applicable. This article does not contain any studies with human participants or animals performed by any of the authors.

Funding

This research did not receive any external funding like specific grant from funding agencies in the public, commercial, or nonprofit sectors.

Conflict of interest

The authors declare that they have no conflicts of interest, competing financial interest or personal relationship that could have influenced the work reported in this paper.

Data and materials availability

All data associated with this study will be available based on the reasonable request to the corresponding author.

REFERENCES

1. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med. 2018;1:39. doi: 10.1038/s41746-018-0040-6
2. Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. Cureus. 2023;15(2):e35179. doi: 10.7759/cureus.35179
3. 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):310. doi: 10.1186/s12911-020-01332-6
4. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical image analysis using convolutional neural networks: a review. J Med Syst. 2018;42(11):226. doi: 10.1007/s10916-018-1088-1
5. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology: new tools for diagnosis and precision oncology. Nat Rev Clin Oncol. 2019;16(11):703-16. doi: 10.1038/s41571-019-0252-y
6. Briganti G, Le Moine O. Artificial intelligence in medicine: today and tomorrow. Front Med (Lausanne). 2020;7:27. doi: 10.3389/fmed.2020.00027
7. Clusmann J, Kolbinger FR, Muti HS, Carrero ZI, Eckardt JN, Ghaffari Laleh N, Löffler CML, Schwarzkopf SC, Unger M, Veldhuizen GP, Wagner SJ, Kather JN. The future landscape of large language models in medicine. Commun Med (Lond). 2023;3:141. doi: 10.1038/s43856-023-00370-1
8. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J. 2019;6(2):94-8. doi: 10.7861/futurehosp.6-2-94
9. Emanuel EJ, Wachter RM. Artificial intelligence in health care: will the value match the hype? JAMA. 2019;321(23):2281-2. doi: 10.1001/jama.2019.4914
10. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-8. doi: 10.1038/nature21056
11. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable AI in healthcare. Lancet Digit Health. 2021;3(11):e745-50. doi: 10.1016/S2589-7500(21)00208-9
12. Han ER, Yeo S, Kim MJ, Lee YH, Park KH, Roh H. Medical education trends for future physicians in the era of advanced technology and artificial intelligence: an integrative review. BMC Med Educ. 2019;19(1):460. doi: 10.1186/s12909-019-1891-5
13. Hatherley JJ. Limits of trust in medical AI. J Med Ethics. 2020;46(7):478-81. doi: 10.1136/medethics-2019-105935
14. He J, Baxter SL, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med. 2019;25(1):30-6. doi: 10.1038/s41591-018-0307-0
15. Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. Gastrointest Endosc. 2020;92(4):807-12. doi: 10.1016/j.gie.2020.06.040

16. Khan RA, Jawaid M, Khan AR, Sajjad M. ChatGPT-reshaping medical education and clinical management. Pak J Med Sci. 2023;39(2):605-7. doi: 10.12669/pjms.39.2.7653
17. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, Madriaga M, Aggabao R, Diaz-Candido G, Maningo J, Tseng V. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health. 2023;2(2):e0000198. doi: 10.1371/journal.pdig.0000198
18. Langlotz CP. Will artificial intelligence replace radiologists? Radiol Artif Intell. 2019;1(3):e190058. doi: 10.1148/ryai.2019190058
19. Larrazabal AJ, Nieto N, Peterson V, Milone DH, Ferrante E. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. Proc Natl Acad Sci U S A. 2020;117(23):12592-4. doi: 10.1073/pnas.1919012117
20. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-74. doi: 10.1038/s41591-020-1034-x
21. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. Z Med Phys. 2019;29(2):102-27. doi: 10.1016/j.zemedi.2018.11.002
22. Malone ER, Oliva M, Sabatini PJB, Stockley TL, Siu LL. Molecular profiling for precision cancer therapies. Genome Med. 2020;12(1):8. doi: 10.1186/s13073-019-0703-1
23. Masters K. Ethical use of artificial intelligence in health professions education: AMEE Guide No. 161. Med Teach. 2024;46(5):541-51. doi: 10.1080/0142159X.2023.2186203
24. McCoy LG, Nagaraj S, Morgado F, Harish V, Das S, Celi LA. What do medical students actually need to know about artificial intelligence? NPJ Digit Med. 2020;3:86. doi: 10.1038/s41746-020-0294-7
25. Merlo EM, Sparacino G, Silvestro O, Giacobello ML, Meduri A, Casciaro M, Gangemi S, Martino G. The role of artificial intelligence in shaping the doctor-patient relationship: a narrative review. Healthcare (Basel). 2026;14(4):481. doi: 10.3390/healthcare14040481
26. Minssen T, Gerke S, Aboy M, Price WN 2nd, Cohen IG. Regulatory responses to medical machine learning. J Law Biosci. 2020;7(1):lsaa002. doi: 10.1093/jlb/lsaa002
27. Moor M, Banerjee O, Shakeri Hossein Abad Z, Krumholz HM, Leskovec J, Topol EJ, Rajpurkar P. Foundation models for generalist medical artificial intelligence. Nature. 2023;616(7956):259-65. doi: 10.1038/s41586-023-05881-4
28. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol. 2019;20(5):e253-61. doi: 10.1016/S1470-2045(19)30154-8
29. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447-53. doi: 10.1126/science.aax2342
30. Panch T, Mattie H, Celi LA. The "inconvenient truth" about AI in healthcare. NPJ Digit Med. 2019;2:77. doi: 10.1038/s41746-019-0155-4
31. 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
32. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, Liu PJ, Liu X, Marcus J, Sun M, Sundberg P, Yee H, Zhang K, Zhang Y, Flores G, Duggan GE, Irvine J, Le Q, Litsch K, Mossin A, Tansuwan J, Wang D, Wexler J, Wilson J, Ludwig D, Volchenboum SL, Chou K, Pearson M, Madabushi S, Shah NH, Butte AJ, Howell MD, Cui C, Corrado GS, Dean J. Scalable and accurate deep learning with electronic health records. NPJ Digit Med. 2018;1:18. doi: 10.1038/s41746-018-0029-1
33. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med. 2022;28(1):31-8. doi: 10.1038/s41591-021-01614-0
34. Rao A, Pang M, Kim J, Kamineni M, Lie W, Prasad AK, Landman A, Dreyer K, Succi MD. Assessing the utility of ChatGPT throughout the entire clinical workflow: development and usability study. J Med Internet Res. 2023;25:e48659. doi: 10.2196/48659
35. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P, Helbich TH, Chevalier M, Tan T, Mertelmeier T, Wallis MG, Andersson I, Zackrisson S, Mann RM, Sechopoulos I. Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. J Natl Cancer Inst. 2019;111(9):916-24. doi: 10.1093/jnci/djy222
36. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. IEEE J Biomed Health Inform. 2018;22(5):1589-604. doi: 10.1109/JBHI.2017.2767063
37. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, Scales N, Tanwani A, Cole-Lewis H, Pfohl S, Payne P, Seneviratne M, Gamble P, Kelly C, Babiker A, Schärli N, Chowdhery A, Mansfield P, Demner-Fushman D, Agüera y Arcas B, Webster D, Corrado GS, Matias Y, Chou K, Gottweis J, Tomasev N, Liu Y, Rajkomar A, Barral J, Semturs C, Karthikesalingam A, Natarajan V. Large language models

encode clinical knowledge. Nature. 2023;620(7972):172-80. doi: 10.1038/s41586-023-06291-2

38. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, Westbrook J, Tutty M, Blike G. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. Ann Intern Med. 2016;165(11):753-60. doi: 10.7326/M16-0961
39. Ström P, Kartasalo K, Olsson H, Solorzano L, Delahunt B, Berney DM, Bostwick DG, Evans AJ, Grignon DJ, Humphrey PA, Iczkowski KA, Kench JG, Kristiansen G, van der Kwast TH, Leite KRM, McKenney JK, Oxley J, Pan CC, Samaratunga H, Srigley JR, Takahashi H, Tsuzuki T, Varma M, Zhou M, Lindberg J, Lindskog C, Ruusuvuori P, Wählby C, Grönberg H, Rantalainen M, Egevad L, Eklund M. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. Lancet Oncol. 2020;21(2): 222-32. doi: 10.1016/S1470-2045(19)30738-7
40. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med. 2023;29(8):1930-40. doi: 10.1038/s41591-023-02448-8
41. 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
42. Tu T, Azizi S, Driess D, Schaekermann M, Amin M, Chang PC, Carroll A, Lau C, Tanno R, Ktena I, Palepu A, Mustafa B, Chowdhery A, Liu Y, Kornblith S, Fleet D, Mansfield P, Prakash S, Wong R, Virmani S, Semturs C, Mahdavi SS, Green B, Dominowska E, Agüera y Arcas B, Barral J, Webster D, Corrado GS, Matias Y, Singhal K, Florence P, Karthikesalingam A, Natarajan V. Towards generalist biomedical AI. NEJM AI. 2024;1(3):AIoa2300138. doi: 10.1056/AIoa2300138
43. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, Denniston AK, Faes L, Geerts B, Ibrahim M, Liu X, Mateen BA, Mathur P, McCradden M, Morgan L, Ordish J, Rogers C, Saria S, Ting DSW, Watkinson PJ, Weber W, Wheatstone P, McCulloch P. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022; 28(5):924-33. doi: 10.1038/s41591-022-01772-9
44. Verghese A, Shah NH, Harrington RA. What this computer needs is a physician: humanism and artificial intelligence. JAMA. 2018;319(1):19-20. doi: 10.1001/jama.2017.19198
45. Wartman SA, Combs CD. Medical education must move from the information age to the age of artificial intelligence. Acad Med. 2018;93(8):1107-9. doi: 10.1097/ACM.0000000000002044
46. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018;2(10):719-31. doi: 10.1038/s41551-018-0305-z