# **What is AI?** An introduction for clinicians

**David Amoateng, Phillip Good, Anand Ramineni, Reece Chang, Ewan McAndrew, Taylan Gurgenci**

*This article is part of a series of articles on artificial intelligence.*

**Background**
Artificial intelligence (AI) is thought by many to likely underpin the next revolution in clinical medicine. The rapidly evolving terminology in AI – including terms such as 'generative AI', 'machine learning' and 'natural language processing' – can present a barrier to clinical adoption.

**Objective**
The aim of this article is to give a non-technical overview of AI and its medical applications, with particular emphasis on large language models.

**Discussion**
The term 'AI' suggests semantic understanding. This impression is reinforced by text output eerily similar to human writing. However, AI has not yet become intelligent in the common understanding of that word. The apparent semantic understanding is not actually there. Knowing this is crucial to avoiding misapplications of this nevertheless revolutionary technology.

**ARTIFICIAL INTELLIGENCE (AI)** is increasingly present in clinical medicine, offering tools that can enhance practice efficiency and support clinical decisions. However, AI is fundamentally a collection of mathematical procedures that identify patterns in data – there is no true understanding or intelligence as we conceive it. This distinction is crucial for safe clinical implementation.[1]

This article will first define key AI terminology relevant to general practice (Table 1), then it will explain how these technologies work using clinical examples, discuss their current and emerging applications in primary care and address important limitations including bias and the need for clinical oversight.

General practitioners (GPs) are already encountering AI in various forms: clinical documentation assistants (AI scribes), diagnostic support systems, risk prediction tools and administrative applications.[2] Understanding these technologies helps clinicians make informed decisions about their adoption and use.

## Understanding AI fundamentals

Machine learning involves systems that improve their performance by analysing data patterns. Consider electrocardiogram (ECG) interpretation software: it examines features such as ST segments and P waves across thousands of expert-annotated ECGs, gradually learning to make similar assessments.[3] The system adjusts its parameters to minimise differences between its predictions and expert diagnoses. This represents AI at its most transparent – we can understand what features it considers and why.

Neural networks and deep learning represent a more complex approach. These systems can identify patterns directly from raw data without predefined features. In our ECG example, although traditional approaches rely on measurements we specify, deep learning systems can discover subtle patterns in the raw signal that humans might not recognise. This capability comes with a significant trade-off: we cannot explain why the system made a particular decision. This 'black box' problem raises important questions about clinical accountability and patient safety.[4]

Generative AI creates new content on the basis of training data. Large language models (LLMs) are the text-generating subset most relevant to clinical practice. These power everything from clinical note generation to patient education materials, making them increasingly present in GP workflows.[5]

## Current AI applications in clinical practice and research settings

Clinical documentation represents one of the most mature applications. AI scribes transcribe consultations in real time and generate structured notes.[6] Although they reduce documentation burden, they require review for accuracy. In other countries, they are also increasingly used to extract key information to optimise revenue generated from the clinical encounter.[6]

Diagnostic support spans multiple specialties though is yet to achieve routine use in Australian general practice. This delayed uptake is justified, given concerns regarding both accuracy as well as explainability. Dermatology image analysis assesses skin lesions with improving accuracy,

 

**Table 1. Common AI terms, definitions and relevant examples**

| Term | Definition | Medical example |
|---|---|---|
| Algorithm | Instructions that direct a computer to solve a problem | Automated full blood count morphology analysis |
| Machine learning | A type of AI that learns to recognise patterns from data without those patterns being pre-programmed | Brain MRI interpretation in the setting of dementia (eg NeuroQuant) |
| Training data | Dataset used to train a model | Set of brain MRIs (refer above) |
| Large language model (LLM) | AI system trained on vast amounts of text that can 'understand' and generate human-like language | ChatGPT (or similar) drafting a patient letter from patient notes |
| Neural network | Computer system modelled on the human neuronal structure | IDx-DR system (FDA approved) for diabetic retinopathy screening |
| Generative AI | AI systems that generate content based on training data | The GPT family of LLMs (eg GPT-4) are text generators |
| Bias | Different performance between groups based on available training data | Skin lesion identification software across different skin tones |

AI, artificial intelligence; FDA, Food and Drug Administration; GPT, generative pre-trained transformer; MRI, magnetic resonance imaging.

with Australian researchers pioneering efforts to improve explainability and therefore clinician trust.[7] Ophthalmology applications screen for diabetic retinopathy in primary care, identifying at-risk patients who might miss specialist review.[8] Radiology AI provides preliminary interpretations, prioritising urgent cases and flagging subtle findings.[9]

Risk prediction tools synthesise multiple data sources. Machine learning has been used by Australian researchers to improve the accuracy of cardiovascular risk prediction.[10] Hospital readmission prediction identifies patients needing support, whereas falls risk assessment and mental health screening represent emerging preventive care applications.[11]

Clinical decision support extends beyond drug interaction checking to providing guideline-based recommendations tailored to patient characteristics and differential diagnosis suggestions. These remain advisory tools requiring clinical judgment.

Administrative applications often provide quick returns on investment. Appointment optimisation reduces no-shows and improves patient satisfaction with their health system interaction.[12]

We now turn to a specific discussion of LLMs, as they have generated a lot of both public and professional interest, after which we discuss some challenges in clinical AI implementation.

## Large language models

Understanding how LLMs function helps clinicians use them appropriately and recognise their limitations. These models process text by analysing patterns in how words and phrases appear together across vast amounts of training data. When a clinician types 'Management of uncontrolled hypertension requires', the model does not consult a medical textbook or recall physiological principles. Instead, it performs statistical analysis on patterns from millions of medical texts, research papers and clinical guidelines to predict what words typically follow this phrase.[13]

This process can produce remarkably coherent and seemingly knowledgeable responses. The model might complete the sentence with 'careful consideration of patient comorbidities, lifestyle modifications and appropriate pharmacological intervention', which sounds clinically appropriate.

However, this apparent sophistication masks a fundamental limitation: the model has no understanding of what hypertension actually is, how blood pressure affects organs or why certain medications work. LLMs are performing pattern matching at a scale and complexity that can produce text indistinguishable from expert clinical writing but without any comprehension of meaning or consequences. Indeed, modern LLMs could probably write both sides of the classic historic Pickering–Platt hypertension debate with tighter prose than either of the real interlocutors while understanding none of it.[14]

Sometimes the variability clinicians notice when using LLMs – where similar questions yield different responses – seems like a subtle change in model 'understanding'. But these differences actually stem from how these models generate text. Rather than selecting the single most probable next word every time, models introduce controlled randomness to avoid repetitive outputs.[15] This means asking 'What are the treatment options for type 2 diabetes?' might yield a response emphasising metformin and lifestyle changes on one occasion, whereas another query might lead with discussion of newer glucagon-like peptide-1 agonists. Neither response indicates the model 'knows' more in one instance than another; it is simply following different probability paths through its training patterns.

This characteristic explains why prompt engineering – the practice of carefully crafting queries to achieve desired outputs – has become important for clinical users. Small changes in phrasing can yield significantly different responses. For instance, asking for 'evidence-based treatment options' versus 'current treatment options' might activate different pattern associations in the model's training data.[16] Clinicians who understand this can better use these tools while recognising that variation in output quality reflects statistical processes, not genuine understanding or reasoning.

### Understanding and addressing bias in clinical AI

The issue of bias in AI systems represents one of the most critical challenges for clinical implementation. All AI systems can perpetuate or amplify biases present in their training data, with potentially serious implications for patient care and health equity.[17] Understanding these biases helps clinicians recognise when AI recommendations might be unreliable or inappropriate for specific patients.

Consider the use case of a neural network approach to skin lesion analysis.[12] The developers gave many naevus photos to the neural network in addition to the final histopathological diagnosis, allowing the program to develop its own approach to diagnosing the presence or absence of malignancy. The problem was that some of the naevi in the training data had ink marking around them as the photos were taken just prior to excision by a dermatologist, and the ones that were excised were more likely to be malignant. The model learnt that one parameter that was very predictive of malignancy was the presence of skin marker around the lesion – which is clearly not useful and an important example of the risks of unintentionally biasing the dataset.

Demographic bias poses perhaps the most widespread challenge for primary care applications. Many AI tools have been trained predominantly on data from specific populations, often overrepresenting younger, healthier, more affluent or ethnically homogeneous groups. Pulse oximetry algorithms provide a stark example: devices calibrated primarily on lighter skin tones can overestimate oxygen saturation in patients with darker skin, potentially missing hypoxemia.[18] Voice recognition systems used in clinical documentation may struggle with accents or speech patterns underrepresented in training data, creating barriers to care for immigrant populations or those with speech differences, as well as for physicians whose accents differ from those represented in the training data.[19]

Socioeconomic bias emerges when AI systems train on data that do not reflect the full spectrum of clinical presentations.[20] An AI system trained primarily on data from well-resourced academic medical centres might learn patterns associated with patients who present early with classic symptoms and good health literacy. When deployed in community settings serving populations with limited healthcare access, these systems may miss atypical presentations or fail to account for the delayed presentations common when patients face barriers to care.[17,20] For instance, an AI system trained to recognise diabetic complications might underperform when patients present with advanced disease because they could not afford regular monitoring or medications.

Historical healthcare disparities embedded in training data create another layer of complexity. If an AI system learns from historical data where certain populations received different treatment patterns – whether because of bias, access issues or social determinants – it may perpetuate these disparities. An AI system might learn that certain ethnic groups are less likely to receive aggressive cardiac interventions, not because it is clinically appropriate, but because that is the pattern in historical data.[17] Without careful attention, such systems could recommend continuing these disparate treatment patterns, effectively encoding past injustices into future care recommendations.

Mitigating these biases requires active, ongoing clinician engagement rather than passive acceptance of AI recommendations. Before implementing any AI tool, clinicians should investigate the composition of its training data. Does it include patients like yours? Were multiple ethnic groups, socioeconomic backgrounds and geographic regions represented? Has the tool been validated specifically in primary care settings or only in specialised academic centres?[17] GPs and other medical practitioners already apply such scrutiny to clinical trials that may not adequately reflect their own patient populations. The skills needed here are similar and readily transferrable.

### Practical implementation and clinical oversight

Successfully integrating AI into general practice requires more than simply purchasing software. It demands thoughtful evaluation, careful implementation and ongoing oversight to ensure these tools enhance rather than compromise patient care. The complexity of primary care, with its undifferentiated presentations and longitudinal relationships, creates unique challenges for AI implementation that might not be apparent in more specialised settings.

AI will increasingly be used to provide point-of-care clinical advice.[21] Understanding the training data's relevance to your specific patient population is crucial – an AI trained on urban emergency department presentations may perform poorly in rural general practice. Validation studies should specifically include primary care settings, not just specialist clinics or hospital environments. The transparency of the AI's decision-making process varies significantly between tools; those that provide some explanation for their recommendations, even if simplified, allow for better clinical oversight than completely opaque systems.[21]

Workflow integration often determines whether an AI tool succeeds or creates more problems than it solves. The most sophisticated AI becomes a burden if it disrupts established clinical workflows or requires extensive workarounds. Successful implementation typically starts with identifying specific pain points in current practice – perhaps the hours spent on documentation or the challenge of keeping up with changing guidelines – then selecting AI tools that address these specific issues.[6]

Training requirements extend beyond simple software operation. Clinical staff need to understand not just how to use AI tools, but also their limitations and potential failure modes. A nurse using an AI-powered triage system must recognise when the AI's recommendations do not align with clinical judgment and feel empowered to override them. Reception staff using AI-powered scheduling scheduling must understand when to involve clinical staff in complex booking decisions the AI system cannot handle appropriately.

### Limitations and future directions

Current AI systems in medicine face fundamental limitations that will not be resolved simply through incremental improvements. The lack of true understanding or reasoning ability means AI cannot genuinely comprehend patient narratives, recognise subtle clinical patterns outside its training or apply judgment to novel situations.[2] When a patient presents with an unusual constellation of symptoms, an experienced GP draws on years of clinical experience, understanding of local disease patterns and knowledge of the individual patient. AI systems cannot replicate this integrative thinking, regardless of their training data volume.

The potential for convincing but incorrect outputs poses particular risks in healthcare. Unlike errors that are obviously wrong, AI can generate recommendations that seem plausible and are presented with apparent confidence, making them harder to identify.[22] A medication recommendation might be pharmacologically reasonable but inappropriate for a specific patient's circumstances in ways the AI cannot appreciate. This 'confident incorrectness' requires constant vigilance from clinicians.

Complex, multifaceted clinical scenarios remain beyond current AI capabilities. General practice often involves managing multiple chronic conditions, psychosocial factors and patient preferences simultaneously. An elderly patient with diabetes, mild cognitive impairment, limited family support and financial constraints requires holistic management that AI cannot adequately address. Although AI might optimise each condition individually, it cannot integrate these factors into coherent, person-centred care plans.

Future developments may address some current limitations through improved architectures and training methods. Proposed improvements include better explainability (so-called explainable AI, or 'xAI'), allowing clinicians to understand AI reasoning; reduced bias through more representative training data and algorithmic adjustments; and enhanced ability to indicate uncertainty rather than always providing confident outputs.[22] However, these remain technical improvements to pattern recognition systems, not steps toward genuine understanding or clinical reasoning. The fundamental need for clinical oversight will persist regardless of technical advances.

## Conclusion

AI technologies offer valuable tools for enhancing primary care delivery, from reducing documentation burden to supporting diagnostic decisions and improving practice efficiency.[2] However, AI tools remain sophisticated pattern-recognition systems without true clinical understanding, judgment or ability to appreciate the full context of patient care. This limitation is not a temporary technical hurdle but a fundamental characteristic of current AI approaches.

The GPs who will thrive in an AI-enhanced healthcare landscape are those who embrace these tools thoughtfully, understanding that the art of medicine – with its emphasis on human connection, clinical judgment and individualised care – remains irreplaceable.

### Authors

David Amoateng PhD, Electrical Engineer, Machine Learning Specialist & Software Lead, Renovo Solutions, Brisbane Qld

Phillip Good FRACP, PhD Director, Cancer Services, Mater Adult Hospital, South Brisbane, Qld; Director, Palliative Care, St Vincent's Private Hospital, Brisbane, Qld; Professor, Faculty of Medicine, Mater Research Institute, The University of Queensland, Brisbane, Qld

Anand Ramineni FRACP, Consultant Neonatologist, Leading Steps Paediatric Clinic, Gold Coast Private Hospital, Gold Coast, Qld; Senior Lecturer, Mater Clinical School, The University of Queensland, Brisbane, Qld; Adjunct Assistant Professor, Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Qld

Reece Chang BPhty, CHIA, GradCertHealthcareRedesign, Clinical Informatics Director, Allied Health - Digital Metro North, Metro North Health, Brisbane Qld

Ewan McAndrew BSc (Phys), BSc (Comp. Sc), DipSc, Principal Software Engineer, Mintel Group Ltd, London, United Kingdom

Taylan Gurgenci FRACGP, FAChPM, Palliative care specialist, Department of Cancer Services, Mater Adult Hospital, South Brisbane, Qld; Mater Research Institute, The University of Queensland, Brisbane, Qld

Competing interests: None.

Funding: TG is the chief investigator for the grant 'Betty McGrath Fellowship in Healthcare delivery and Innovation' – he draws a salary from this grant to secure dedicated research time. PG, AR and DA are principal investigators on this grant but do not receive any funds individually from it.

Provenance and peer review: Commissioned, externally peer reviewed.

AI declaration: The authors advise that there was use of artificial intelligence (AI)-assisted technology for assisting in the writing or editing of the manuscript, and accept full responsibility for all content. Details on how AI was used have been declared to the Editors.

**Correspondence to:**
t.gurgenci@uq.edu.au

### Acknowledgements

The ongoing support of Mater Research and the Mater Foundation is gratefully acknowledged.

### References

*References are available online only.*

correspondence **ajgp@racgp.org.au**

## References


1. Jindal JA, Lungren MP, Shah NH. Ensuring useful adoption of generative artificial intelligence in healthcare. J Am Med Inform Assoc 2024;31(6):1441–44. doi: 10.1093/jamia/ocae043.
2. Sarkar U, Bates DW. Using artificial intelligence to improve primary care for patients and clinicians. JAMA Intern Med 2024;184(4):343–44. doi: 10.1001/jamainternmed.2023.7965.
3. Al-Zaiti SS, Martin-Gill C, Zègre-Hemsey JK, et al. Machine learning for ECG diagnosis and risk stratification of occlusion myocardial infarction. Nat Med 2023;29(7):1804–13. doi: 10.1038/s41591-023-02396-3.
4. Pavlidis G. Unlocking the black box: Analysing the EU artificial intelligence act's framework for explainability in AI. Law Innov Technol 2024;16(1):293–308. doi: 10.1080/17579961.2024.2313795.
5. Omiye JA, Gui H, Rezaei SJ, Zou J, Daneshjou R. Large language models in medicine: The potentials and pitfalls: A narrative review. Ann Intern Med 2024;177(2):210–20. doi: 10.7326/M23-2772.
6. Agarwal PMDM, Lall RMD, Girdhari RMDMBA. Artificial intelligence scribes in primary care. CMAJ 2024;196(30):E1042. doi: 10.1503/cmaj.240363.
7. Mehta D, Primiero C, Betz-Stablein B, et al. Multi-task AI models in dermatology: Overcoming critical clinical translation challenges for enhanced skin lesion diagnosis. J Eur Acad Dermatol Venereol 2025;jdv.20551. doi: 10.1111/jdv.20551.
8. Dow ER, Chen KM, Zhao CS, et al. Artificial intelligence improves patient follow-up in a diabetic retinopathy screening program. Clin Ophthalmol 2023;17:3323–30. doi: 10.2147/OPTH.S422513.
9. Sridharan S, Seah Xin Hui A, Venkataraman N, et al. Real-world evaluation of an AI triaging system for chest X-rays: A prospective clinical study. Eur J Radiol 2024;181:111783. doi: 10.1016/j.ejrad.2024.111783.
10. Sajeev S, Champion S, Beleigoli A, et al. Predicting Australian adults at high risk of cardiovascular disease mortality using standard risk factors and machine learning. Int J Environ Res Public Health 2021;18(6):3187–14. doi: 10.3390/ijerph18063187.
11. Maher C, Dankiw KA, Singh B, Bogomolova S, Curtis RG. In-home evaluation of the neo care artificial intelligence sound-based fall detection system. Future Internet 2024;16(6):197. doi: 10.3390/fi16060197.
12. Knight DRT, Aakre CA, Anstine CV, et al. Artificial intelligence for patient scheduling in the real-world health care setting: A metanarrative review. Health Policy Technol 2023;12(4):100824. doi: 10.1016/j.hlpt.2023.100824.
13. Tian S, Jin Q, Yeganova L, et al. Opportunities and challenges for ChatGPT and large language models in biomedicine and health. Brief Bioinform 2023;25(1):bbad493. doi: 10.1093/bib/bbad493.
14. Alderman MH. Measures and meaning of blood pressure. Lancet 2000;355(9199):159. doi: 10.1016/S0140-6736(99)00409-2.
15. Hunter RB, Mehta SD, Limon A, Chang AC. Decoding ChatGPT: A primer on large language models for clinicians. Intell Based Med 2023;8:100114. doi: 10.1016/j.ibmed.2023.100114.
16. Giray L. Prompt engineering with ChatGPT: A guide for academic writers. Ann Biomed Eng 2023;51(12):2629–33. doi: 10.1007/s10439-023-03272-4.
17. Norori N, Hu Q, Aellen FM, Faraci FD, Tzovara A. Addressing bias in big data and AI for health care: A call for open science. Patterns (N Y) 2021;2(10):100347. doi: 10.1016/j.patter.2021.100347.
18. Cabanas AM, Martín-Escudero P, Pagán J, Mery D. Technical and regulatory challenges in artificial intelligence-based pulse oximetry: A proposed development pipeline. Br J Anaesth 2025;134(5):1295–99. doi: 10.1016/j.bja.2025.02.014.
19. Shi X, Yu F, Lu Y, et al. The accented English speech recognition challenge 2020: Open datasets, tracks, baselines, results and methods. Paper presented at the IEEE International Conference on Acoustics, Speech and Signal Processing, June 2021.
20. Gu T, Pan W, Yu J, et al. Mitigating bias in AI mortality predictions for minority populations: A transfer learning approach. BMC Med Inform Decis Mak 2025;25(1):30–11. doi: 10.1186/s12911-025-02862-7.
21. Kwong JCC, Khondker A, Lajkosz K, et al. APPRAISE-AI Tool for Quantitative Evaluation of AI Studies for Clinical Decision Support. JAMA Netw Open 2023;6(9):e2335377. doi: 10.1001/jamanetworkopen.2023.35377.
22. Reddy S. Explainability and artificial intelligence in medicine. Lancet Digit Health 2022;4(4):e214–15. doi: 10.1016/S2589-7500(22)00029-2.