

## REVIEW OPEN



# Increasing use of artificial intelligence in genomic medicine for cancer care- the promise and potential pitfalls

Olivia O'Connor<sup>1</sup>✉ and Terri P. McVeigh<sup>2</sup>

© The Author(s) 2025

The field of genomic medicine produces large datasets, which need to be rapidly analysed to produce clinically actionable insights in cancer care. Artificial intelligence thrives on data, processing and learning from datasets with a degree of accuracy and efficiency that traditional computing algorithms can not achieve. Based on a patient's genome sequence, AI could allow earlier detection of cancer, inform personalised treatment plans and provide insights into prognostication. However, this valuable tool is met with skepticism, with stakeholders concerned over data security, liability for AI's mistakes due to hallucination and the threat to clinical jobs. This review highlights both the benefits and potential problems of using AI in genomic medicine for cancer care, with the aim to lessen the knowledge gap between clinicians and data scientists and facilitate the future deployment of AI in cancer care.

BJC Reports; <https://doi.org/10.1038/s44276-025-00135-4>

## INTRODUCTION

'Artificial intelligence (AI) has emerged as a powerful tool in the field of cancer genomics, revolutionising our understanding of cancer and its treatment.' At least, that is the first line of the response given by ChatGPT, arguably the most discussed AI entity at present, when asked about the use of AI in genomics. In recent years, AI has improved beyond recognition. From essay writing to genome analysis, it has become indispensable in an academic's toolbox. Genomics is one of many fields in recent years which has heralded the arrival of the big data era [1], principally as a result of the advent of next generation sequencing (NGS). With major advancements in the speed of sequencing comes swathes of raw genomic data, which requires difficult and time-consuming analysis to convert into clinically actionable insights. The need for an AI assistant has never been greater, and thankfully by their very nature AI algorithms are driven to self-improve using the large amount of data that necessitates their creation. The integration of AI into industry is already upon us, with the Topol Review laying out how AI can be responsibly rolled out within the NHS [2].

This review will first define AI and its role in genomics before balancing the promises and pitfalls associated with this emerging technology.

## WHAT IS AI?

**Artificial intelligence (AI)** is a term coined nearly 70 years ago to define the science of creating intelligent machines [3]. Using computer processes, it aims to emulate human thought processes to accomplish tasks which typically require objective reasoning and understanding in the same manner as a human subject-matter expert (SME) [4]. AI can be described as weak or strong.

**'Weak' or 'Narrow' AI** describes the application of learning algorithms for specific tasks. Applications include natural language processing tools/chat bots (such as ChatGPT), virtual assistants (such as Siri or Alexa) and image recognition. All existing AI systems are domain-specific and thus can be classified as weak AI [5].

**'Strong' or 'General' AI** is currently theoretical, referring to AI that could progress to a sophisticated level surpassing human intellect with the abilities of agile reaction and forward planning in the same manner as humans [5]. This gives rise to concerns regarding the potential for AI dominance over human intelligence [6].

**Machine Learning (ML)** is a subset of artificial intelligence, referring to computer systems that learn automatically from experience without being explicitly programmed [7]. Machine learning systems identify patterns in datasets and create an algorithm encompassing their findings. They can then apply this to new data, extrapolating knowledge to unfamiliar situations [8].

**Deep Learning (DL)** is a further evolution of machine learning which uses artificial neural networks to recognise patterns in data and provide a suitable output [9]. Deep learning layers algorithms into an artificial neural network, taking inspiration from the structure of the human brain; the output of a previous layer will act as the input for the next layer much like the transmission of information through neurons [10]. The relationship between deep learning, machine learning and AI is demonstrated in Fig. 1.

**Big data** refers to datasets too complex to be processed using traditional data processing methods. Data can be "big" in terms of the volume of data (the scalability), the velocity at which the data is created and the variety of data sets (the dimensionality) [11]. AI differs from traditional computing algorithms in that it can learn from data and tackle new information, excelling at pattern

<sup>1</sup>Frimley Health NHS Foundation Trust, England, UK. <sup>2</sup>Cancer Genetics Unit, The Royal Marsden NHS Foundation Trust, England, UK. ✉email: [Ooconnor262@gmail.com](mailto:Ooconnor262@gmail.com)

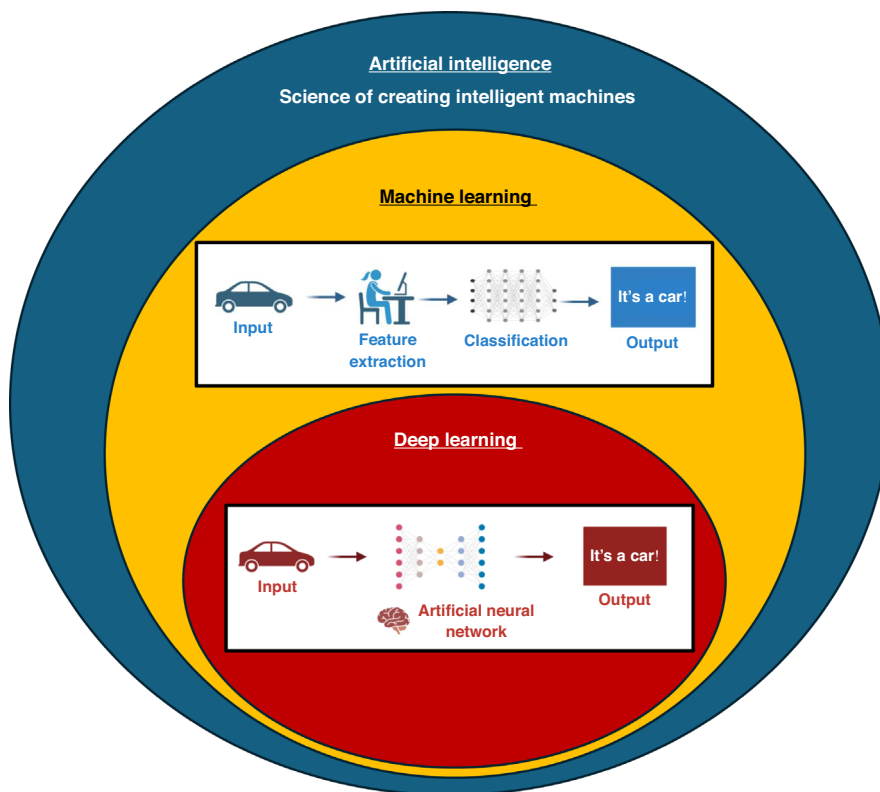


Fig. 1 Venn diagram demonstrating the relationships between AI, ML and DL.

recognition. Where traditional algorithms struggle to keep up when faced with big data, AI algorithms adapt and improve without the need for human interference [12].

#### WHY APPLY AI IN GENOMICS?

Large datasets are a hallmark of genomics. The human genome project produced a reference genome in a period of 13 years at a cost of \$3 billion. This genome alone comprised over 20,000 genes, over 3 billion base pairs [13]. With the advent of Next Generation Sequencing, genomes of individuals can be produced at a fraction of the time for a fraction of the cost. As of 2022 the National Human Genome Research Institute cited the cost of sequencing a genome at \$525 [14]. The Guinness World record for the fastest genome sequencing was just over 5 h [15]. Now, the prospect of sequencing every patient's genome is a realistic one, with the NHS Genomic Medicine service aiming to make the NHS the first national health care system to offer whole genome sequencing as part of routine care [16].

But with big patient populations seeking cheap and fast genome sequencing comes big data, both in terms of scalability and dimensionality. As an example, the Cancer Genome Atlas (TCGA) contains more than 10,000 cancer genomes spanning 33 cancer types. This along with corresponding epigenomic, transcriptomic, and proteomic data yields a total of 2.5 petabytes of raw data [17]. Genomic medicine's challenge now is not the generation of datasets, but the analysis and interpretation of massive amounts of data contained therein. The manual interpretation of genomics data presents a daunting challenge, especially given the number of experts capable of analysing them has not increased commensurately. Furthermore, different individuals may obtain different results while analysing the same dataset (a problem with reproducibility). This is where AI can become invaluable.

Clinical interpretation of genomes relies on accurately identifying significant genetic variants amongst the millions populating

each genome, known as variant calling. To achieve this, raw sequence data is aligned to the reference genome. The quality of the data is improved by removal of duplicates, insertions and deletions, then realignment, base recalibration and finally removal of false positives [18]. DeepVariant, a DL model, outperforms standard tools on some variant calling tasks demonstrating AI's ability to deal with large data sets [19].

#### WHAT ARE THE PROMISES OF AI IN GENOMIC MEDICINE?

##### Early detection of cancer

Improved variant calling facilitated by AI can yield more accurate and efficient identification of cancer-causing variants. This can lead to earlier detection of drivers, more accurate diagnosis of cancer, and facilitate targeted, precision medicine, leading to improved outcomes for the patient [20].

Another way in which AI is facilitating early detection of cancer is by opening up alternative means of investigation. Tumour cells routinely release material into bodily fluids – including blood, pleural, peritoneal or cerebrospinal fluid, nipple aspirate or urine. In many cases, access to these samples as part of a so-called “liquid biopsy” may be less invasive, more easily obtained, and may provide a more holistic overview of the spatially heterogeneous genomic landscape of a tumour compared to a tissue sample. Longitudinal sampling also permits tracking of temporal evolution of genomic changes in a cancer. Material may include circulating tumour cells, extracellular vesicles, or nucleic acids, including tumour-derived cell free DNA (cfDNA). Interest in the clinical utility of the liquid biopsy is accelerating, with applications in profiling of advanced disease [21, 22], early detection of disease relapse [23–25] and early diagnosis in symptomatic individuals [26]. As part of a large-scale national trial, use of a liquid biopsy for detection of methylation patterns indicative of cancer in cfDNA in an asymptomatic population cohort (multi-cancer early detection (MCED) test) is being investigated [27–29].

The use of this technology in this context has been criticised, particularly given the lack of clarity regarding how asymptomatic patients with methylation patterns or other biomarkers indicative of potential underlying cancer should be investigated – and by which type of clinical specialist [30]. There is an opportunity here for artificial intelligence to be applied to help determine most likely underlying primary cancer to help streamline referrals and rationalise investigation of patients [31]. Indeed, the vast amount of data produced by MCED tests necessitate application of machine learning and other AI technologies to enable timely analysis [28] and, therefore, intervention [32] – so called MCED-AI [33]. There is, however, a need for further evaluation to determine whether this translates to a significant survival benefit compared to other types of population screening [30].

Application and interpretation of results derived from analysis of liquid biopsies is complex; complicated by co-existing constitutional genetic variation, naturally occurring age-related clonal haematopoiesis or clonal haematopoiesis of indeterminate potential [34], potential identification of driver variants from second undiagnosed primary cancers, or detection of variants of uncertain significance or actionability [35]. As well as this, ctDNA is less likely to be detectable for early-stage disease or in individuals with brain-only disease. Various AI technologies have been applied to try to minimise “noise” from sequencing data, to facilitate signal enhancement to enable detection of variants at very low frequencies, enable pattern/signature recognition, and, crucially, allow integration of other clinicopathological data for comprehensive analysis [36–38].

One of many examples includes a recent study used a novel computational model which integrated genomic and epigenomic data from the cell-free DNA to detect cancer. It did this with a high sensitivity at 95% specificity, with 91% and 98% sensitivity across two training cohorts (versus below 50% sensitivity for the pre-existing model DELFI). By improving the diagnostic power of liquid biopsy, this non-invasive technique could be applied to screen for cancer yielding early intervention opportunities [39].

AlphaMissense, a recently developed AI model constructed on the protein structure prediction model AlphaFold [40], can predict the pathogenicity of all possible missense variants in the human genome at a single amino acid substitution level [41]. Classifying missense variants has been an ongoing challenge in human genetics, but with this new deep learning tool clinically relevant pathogenic variants can be identified and thereafter screened for, leading to earlier detection of diseases.

#### **Variant prediction from phenotypic data for focused genetic testing**

Computer vision, a field of AI which trains computers to interpret information from image data, can inform targeted genomic testing. It has been successfully applied to distinguish cancer cells from non-cancer cells on histopathological specimens across many different cancer types [42], but recent AI models have gone one step further. These models, trained on histopathological images, have not only been able to identify cancer cells but can also predict which mutations are present in the tumour based on phenotypic features. For example Inception V3 was trained on whole slides obtained from TCGA and was able to predict mutations in lung adenocarcinoma and liver cancers based on histopathology of the tumours [43, 44]. Another recent AI innovation using computer vision comes in the form of the CHARM DL model, trained on 2334 brain tumour samples. It is able to assist in making real time decisions intra-operatively during neurosurgical resection of brain tumours. Based on histopathological images, it can predict the borders of the tumour using cellular density to inform surgical decisions of resection margins. It is also able to predict the molecular profile of the tumours, allowing personalised treatment of the brain cancer with suitable medication directly onto the tissue [45].

Similarly, computer vision can be applied to radiological images of tumours to infer genetic changes within the tumour. For instance, CT and PET scan images have been analysed by AI to predict *EGFR* mutation status in non-small-cell lung cancer (NSCLC) [46].

#### **Precision medicine**

Precision medicine involves tailoring treatment to individual patients based on molecular profiling using genomic information [47]. It aims to classify individuals based on their prognosis or their response to specific treatments, and therefore recommend interventions for those who would benefit, and spare expense and side effects for those who would not. AI, when coupled with genomic data, enables the development of highly personalised treatment plans by recommending targeted therapies that are more likely to be effective, thereby minimising trial-and-error treatment approaches. This is represented schematically in Fig. 2.

When recommending immunotherapies, for example, biopsy is the traditional method of collecting information on biomarkers of the immunogenicity tumour microenvironment. This information is used to predict response to immunotherapy. Deep learning methods have been employed to predict biomarkers using radiomic and pathomic data to avoid these invasive procedures, for instance DL trained on histology specimens and clinical data has been used to predict immunotherapy response in advanced melanoma [48].

AI models have also been employed to recommend personalised treatment plans based on the genomes of brain cancer patients. IBM Watson was able to analyse genomes and propose a treatment plan in 10 min where it would take 160 h of work by expert humans. However, dangerous inaccuracies in treatment plans were noted when it was initially trialled in US hospitals [49]. Retrospective analyses of these errors can be impossible due to the lack of interpretability of AI models, or the ‘black box’ phenomenon discussed below.

#### **Drug discovery**

AI-driven drug discovery platforms can identify novel therapeutic targets by analysing vast datasets of genomic information, expediting this previously time-consuming process (this is represented schematically in Fig. 2). Recently, using AI programme AlphaFold, a potential drug to target liver cancer was developed in 30 days, a process that would previously have taken years. AlphaFold was used to predict the structure of a novel protein CDK20 from proteomic data, and to design a molecule which targeted the protein’s weak spots [50]. AI can also be used to predict the feasibility of drug repurposing. The deep learning model CDRscan predicted anti-cancer drug responsiveness based on a large-scale drug screening assay. It identified 14 oncology and 23 non-oncology drugs as having new potential anti-cancer indications [51].

#### **Improved treatment monitoring and prognostication**

AI can monitor a patient’s genomic data over time to assess treatment effectiveness and identify the development of resistance or relapse. This can guide adjustments in treatment plans. For example, the DL software IDEA was used to successfully identify and characterise single nucleotide variations in circulating tumour DNA, which can be used to monitor response and relapse in the blood of patients with metastatic colorectal cancer receiving targeted agents [52].

Drug responses are closely associated with genomic alterations in cancer cells. AI models have been developed for the prediction of drug responses using genomic data. RefDNN improved prediction of resistance and identification of biomarkers related to drug response, outperforming existing methods, and showed a more robust prediction for untrained drugs [53].

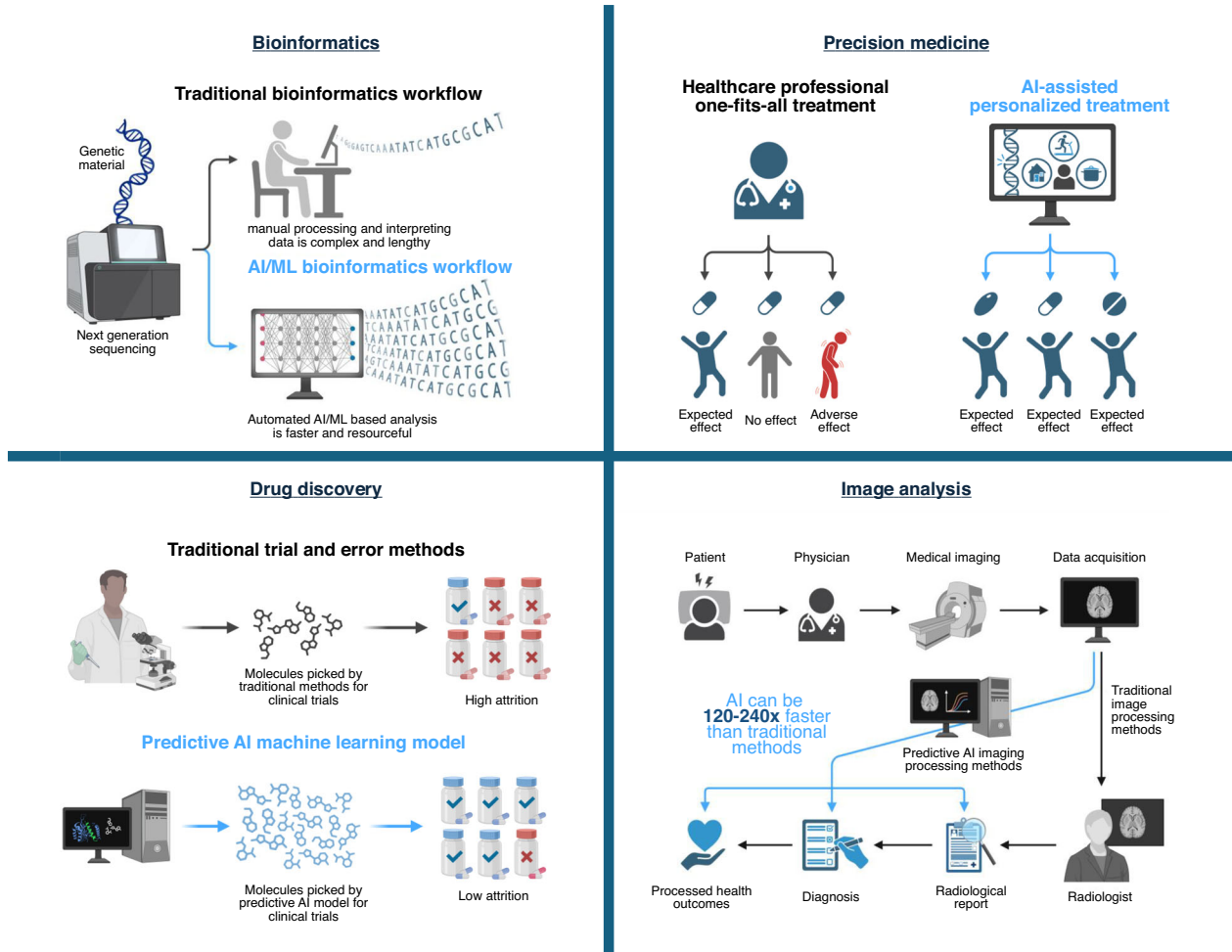


Fig. 2 Schematic showing applications of AI in cancer genomics.

AI software has also been used to improve the accuracy of prognostication, which empowers patients to make informed choices regarding their ongoing treatment. DL software integrating histological and genomic data to predict time to event outcomes showed prediction accuracy surpassing the current clinical paradigm for predicting overall survival of patients diagnosed with glioma [54].

### Improved sample analysis

A final interesting development to touch on is the recent application of FFPEsig AI tool to accurately analyse patterns of DNA mutations from tissues stored at room temperature in paraffin wax. These archival tumour samples, of which there are tens of thousands stored in the pathology archives of a typical large hospital, were previously lost to analysis due to DNA damage associated with FFPE-preservation. FFPEsig enables detection of 9/10 mutational processes by tracking back changes to the DNA that occur in formalin fixation. The analysis of these samples is particularly useful in cancer tissue collected over time which will allow scientists to analyse the first steps in cancer evolution in individual tumours. Mutational analysis is no longer exclusively possible on fresh or frozen tissue samples [55].

### WHAT ARE THE PITFALLS OF AI IN GENOMIC MEDICINE?

As established, AI algorithms require vast quantities of data in order to train the algorithms. In genomics, this data is highly

sensitive and personal. Healthcare data is often the target of costly ransomware attacks which generates unease in the patient population that their personal health data is going to be misused [56]. An oft-posed ethical dilemma comes in the form of availability of genomic data to insurance companies, and the question of whether policies would be adjusted based on genetic risk of disease (although it is worth noting that the use of genetic data by insurance companies is currently prohibited in the UK) [57, 58]. A potential solution to these data storage problems comes in the form of privacy preserving distributed DL, which enables multiple parties to jointly learn via DL models without sharing local datasets. Another solution is multi-centre data sharing agreements such as the Cancer Imaging archive [59].

Clinical governance poses an ethical issue to the users of AI- if the machine were to make prediction errors, who would be responsible? In healthcare, where decisions made can have a large impact on patient outcomes and carry a large risk of litigation, this decision is particularly important and the phenomenon of AI hallucination, when AI infers patterns which are non-existent in the training data to generate inaccurate outputs [60], can make AI an unreliable co-worker. The responsible party for incorrect decisions made by DL tools, whether that be the referring clinician, software engineer or someone else, must be determined before their deployment. A paper by Crigger et al aims to provide a roadmap for the responsible deployment of AI in healthcare by exploring the principles necessary for establishing trust in AI systems, key amongst those addressed being establishment of clear accountability [61]. They suggest a shared accountability

**Table 1.** Summary of potentials and pitfalls as laid out in this review article.

Potentials	Pitfalls
Early Cancer Detection: AI enhances variant calling and supports early cancer diagnosis through improved analysis of liquid biopsy results.	Data Privacy Concerns: Genomic data is highly sensitive, raising issues of security, potential misuse, and ethical dilemmas particularly in insurance.
Precision Medicine: AI enables highly personalised treatment plans, reducing trial-and-error and improving therapeutic outcomes e.g. by predicting individual immunotherapy responses.	Clinical Governance: Uncertainty over responsibility in case of AI prediction errors, with potential for litigation and accountability issues.
Drug Discovery: AI accelerates drug development and repurposing, exemplified by AlphaFold's role in discovering liver cancer therapies.	Interpretability: AI often operates as a "black box," lacking transparency in decision-making, which is problematic for high-risk healthcare decisions.
Improved Treatment Monitoring: AI tracks genomic changes over time, offering real-time insights into treatment effectiveness and potential relapses.	Bias in AI Models: Under-representation in training data can lead to biased AI models, affecting accuracy for certain populations, e.g., racial biases in medical diagnostics.
Improved Sample Analysis: AI can analyse archived tissue samples, enabling the study of cancer evolution in stored biopsies.	Job Displacement: AI's rapid development could potentially replace human workers in some roles, leading to job insecurity despite new opportunities.
Image Analysis in Genomic Medicine: AI models can analyse radiological and histopathological images to predict genomic alterations and support personalised treatment plans.	Lack of Retraining in Healthcare: Clinicians may struggle with the rapid integration of AI, especially if they lack the skills to operate new AI technologies effectively.

model, with the ultimate responsibility for decisions which have involved AI resting with the human healthcare provider. They argue physicians should base decisions on professional judgement rather than blindly relying on AI recommendations. Developers share accountability by ensuring machines are reliable and free from biases and healthcare institutions bear the responsibility of integrating AI systems safely into clinical workflows. Other solutions posed to this problem, aimed at allowing AI to work more autonomously, include giving AI legal personhood, similar to how corporations are occasionally afforded human rights, or holding people who manage AI responsible through vicarious liability in the same way in which an employer is responsible for their employee's actions [62].

A further issue with AI algorithms which becomes particularly troublesome in healthcare settings is interpretability. AI tools are often a 'black box' in that they produce an output without any explanation or justification. It is not appropriate to blindly follow the AI tool when making high risk decisions associated with healthcare. Interpretable DL is a trend aiming to alleviate this limitation, for example using the heat-map like class activation algorithm. This allows visualisation of the image regions which are used by DL models when making a decision [63].

Whilst automation is generally thought to be a method in which to eliminate biases, under-representation in training data sets can lead to machine bias. For instance, when using AI to detect Down syndrome from facial features, detection was much more accurate in European populations as opposed to African populations, due to the underrepresentation of African patients in the training data [64]. AI tools can be retrained on more representative datasets to help remedy this problem (Table 1).

A final limitation to be discussed is a perceived threat to jobs with the advent of AI. The technology has evolved so quickly that it is conceivable it will soon surpass human capabilities and make workers redundant. Five hundred and forty million years ago, a rapid burst of evolution, described as a "Cambrian Explosion" led to rapid and dramatic appearance of diverse metazoans with biomineralized skeletons, promoted by external changes in the environment [65]. Many authors have described the precipitous development of and growth in use of AI, robotics and other digital technologies in healthcare, and indeed in daily life, using the same term [66, 67]. Others feel that this rapid expansion of the technology is more akin to the Industrial Revolution [68]. Where the Industrial Revolution mechanised labor to increase productivity, AI automates cognitive and repetitive tasks, enhancing efficiency across sectors. However, both have raised concerns

about widening gaps between those who adopt the technology and those left behind.

The Topol review outlined how digital technology should best be employed in the NHS, arguing that AI should be used to automate mundane tasks so the workforce can focus on the 'human' elements of patient care. It also suggested that more employment opportunities would be created within the NHS to help develop AI technologies than would be made redundant by AI technologies [2]. However, in order to make these opportunities accessible to clinicians, the knowledge gap between clinical and data science experts must be breached. The Topol review proposed that in order to reap the benefits of AI, the NHS must build a digitally ready workforce with the confidence to adapt new technologies in clinical practice. Efforts by the British Society for Genetic Medicine to enhance awareness and improve skills of the workforce in this area include incorporation of dedicated plenary sessions in annual conference proceedings [69], and establishment of a BSGM Emerging Technologies Committee.

## CONCLUSION

As we have seen, AI offers a promising tool for tractably analysing large datasets in genomics, allowing clinicians and researchers to identify new mutations, drugs, and treatment programmes in patients. However, as this technology grows in sophistication, new guidelines are required to protect patients and maintain the humanity of health care. Nevertheless, AI promises to be transformative for the future of medicine by realising the dream of personalised treatment for patients. Training of the specialist genomics workforce is required ensure that the advantages and utilities of digital health technologies can be maximised to improve patient care.

## DATA AVAILABILITY

No datasets were generated or analysed during the current study.

## REFERENCES

- Noor AM, Holmberg L, Gillett C, Grigoriadis A. Big Data: the challenge for small research groups in the era of cancer genomics. *Br J Cancer*. 2015;113:1405–12.
- Topol E. Preparing the healthcare workforce to deliver the digital future: An independent report on behalf of the Secretary of State for Health and Social Care. Health Education England; 2019;1–48. Available from: <https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf>.

3. National Physical Laboratory (Great Britain). Mechanisation of thought processes. HM Stationery Office; 1959
4. Thomason R. Logic and Artificial Intelligence. Platostanfordedu. 2003; Available from: <https://plato.stanford.edu/archives/sum2020/entries/logic-ai/> [Accessed 17th December 2024]
5. Bory P, Natale S, Katzenbach C. Strong and weak AI narratives: an analytical framework. *AI & SOCIETY* 2024;10:1-1.
6. Goode L. Life, but not as we know it: AI and the popular imagination. *Cult Unbound*. 2018;10:185–207.
7. Matheny M, Israni ST, Ahmed M, Whicher D Artificial intelligence in health care: The hope, the hype, the promise, the peril. Washington, DC: National Academy of Medicine 2019;10.
8. Royal Society Working Group. Machine learning: The power and promise of computers that learn by example. Technical Report. 2017.
9. Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput Sci*. 2021;2:420.
10. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw*. 2015;61:85–117.
11. Laney D 3D Data Management: Controlling Data Volume, Velocity, and Variety [Internet]. META Group; 2001. Available from: <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>.
12. Demigla S The impact of Big Data on AI. 2020 International Conference on Computational Science and Computational Intelligence (CSCI) [Internet]. 2020; Available from: <https://american-cse.org/sites/csci2020proc/pdfs/CSCI2020-6SccvdzjqC7bKupZxFmCoA/762400b395/762400b395.pdf>.
13. Watson JD. The human genome project: past, present, and future. *Science*. 1990;248:44–9.
14. KA W. National human genome research institute. DNA sequencing costs: data from the nhgri genome sequencing program (gsp). Available from: <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>.
15. Stanford Medicine. "Fastest DNA sequencing technique helps undiagnosed patients find answers in mere hours." *ScienceDaily*. ScienceDaily (2022). Available from: [www.sciencedaily.com/releases/2022/01/220113092144.html](http://www.sciencedaily.com/releases/2022/01/220113092144.html). [Accessed 8 Dec 2023].
16. NHS England. NHS Genomic Medicine Service. United Kingdom. NHS. (2023). Available from: <https://www.england.nhs.uk/genomics/nhs-genomic-med-service/>. [Accessed 15 Oct 2024].
17. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, et al. The cancer genome atlas pan-cancer analysis project. *Nat Genet*. 2013;45:1113–20.
18. Dlamini Z, editor. Artificial Intelligence and Precision Oncology: Bridging Cancer Research and Clinical Decision Support. Springer Nature; 2023.
19. Poplin R, Chang PC, Alexander D, Schwartz S, Colthurst T, Ku A, et al. A universal SNP and small-indel variant caller using deep neural networks. *Nat Biotechnol*. 2018;36:983–7.
20. Crosby D, Bhatia S, Brindle KM, Coussens LM, Dive C, Emberton M, et al. Early detection of cancer. *Science*. 2022;375:eaay9040.
21. Popat S, Januszewski A, O'Brien M, Ahmad T, Lewanski C, Dernedde U, et al. Long term efficacy of first-line afatinib and the clinical utility of ctDNA monitoring in patients with suspected or confirmed EGFR mutant non-small cell lung cancer who were unsuitable for chemotherapy. *Br J Cancer*. 2024;132:1–8.
22. NHS England. Thousands more lung cancer patients to get innovative blood test as part of NHS pilot [Internet]. [www.england.nhs.uk](http://www.england.nhs.uk). 2024. Available from: <https://www.england.nhs.uk/2024/03/thousands-more-lung-cancer-patients-to-get-innovative-blood-test-as-part-of-nhs-pilot/>.
23. Abbosh C, Frankell AM, Harrison T, Kisistok J, Garnett A, Johnson L, et al. Tracking early lung cancer metastatic dissemination in TRACERx using ctDNA. *Nature*. 2023;616:553–62.
24. Garcia-Murillas I, Chopra N, Comino-Méndez I, Beaney M, Tovey H, Cutts RJ, et al. Assessment of molecular relapse detection in early-stage breast cancer. *JAMA Oncol*. 2019;5:1473–8.
25. Nakamura Y, Watanabe J, Akazawa N, Hirata K, Kataoka K, Yokota M, et al. ctDNA-based molecular residual disease and survival in resectable colorectal cancer. *Nat Med*. 2024;30:3272–83.
26. Mencil J, Feber A, Begum R, Carter P, Smalley M, Bourmpaki E, et al. Liquid biopsy for diagnosis in patients with suspected pancreatic and biliary tract cancers: PREVAIL ctDNA pilot trial. *J Clin Oncol*. 2022;40:522–2.
27. Detecting cancer early [Internet]. NHS-Galleri Trial. Available from: <https://www.nhs-galleri.org/>.
28. Klein EA, Richards D, Cohn A, Tummala M, Lapham R, Cosgrove D, et al. Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. *Ann Oncol*. 2021;32:1167–77. [https://www.annalsofncology.org/article/S0923-7534\(21\)02046-9/fulltext](https://www.annalsofncology.org/article/S0923-7534(21)02046-9/fulltext).
29. Old R, Pharoah P, Wald N. NHS announces a pilot of a blood test for early detection of many cancers. *J Med Screen*. 2021;28:1–2.
30. Turnbull C, Wald N, Sullivan R, Pharoah P, Houlston RS, Aggarwal A, et al. GRAIL-Galleri: why the special treatment? *Lancet*. 2024;403:431–2.
31. Bahado-Singh R, Vlachos KT, Aydas B, Gordevicius J, Radhakrishna U, Vishweswariah S. Precision oncology: artificial intelligence and DNA methylation analysis of circulating cell-free DNA for lung cancer detection. *Front Oncol*. 2022;12:790645.
32. Vittone J, Gill D, Goldsmith A, Klein EA, Karlitz JJ. A multi-cancer early detection blood test using machine learning detects early-stage cancers lacking USPSTF-recommended screening. *NPJ Precis Oncol*. 2024;8:91.
33. Wang HY, Lin WY, Zhou C, Yang ZA, Kalpana S, Lebowitz MS. Integrating artificial intelligence for advancing multiple-cancer early detection via serum biomarkers: a narrative review. *Cancers*. 2024;16:862.
34. Hoermann G. Clinical significance of clonal hematopoiesis of indeterminate potential in hematology and cardiovascular disease. *Diagnostics*. 2022;12:1613.
35. Mencil J, Rayarel N, Proszek P, Carter P, Feber A, Popat S, et al. Incidental finding of leukaemia in circulating tumour DNA—the importance of a molecular tumour board. *BJC Rep*. 2024;2:12.
36. Widman AJ, Shah M, Frydendahl A, Halmos D, Khamnei CC, Øgaard N, et al. Ultrasensitive plasma-based monitoring of tumor burden using machine-learning-guided signal enrichment. *Nat Med*. 2024;30:1655–66.
37. Ritch EJ, Herberts C, Warner EW, Ng SW, Kwan EM, Bacon JV, et al. A generalizable machine learning framework for classifying DNA repair defects using ctDNA exomes. *NPJ Precis Oncol*. 2023;7:27.
38. Foser S, Maiese K, Digumarthy SR, Puig-Butille JA, Rebhan C. Looking to the future of early detection in cancer: liquid biopsies, imaging, and artificial intelligence. *Clin Chem*. 2024;70:27–32.
39. Bae M, Kim G, Lee TR, Ahn JM, Park H, Park SR, et al. Integrative modeling of tumor genomes and epigenomes for enhanced cancer diagnosis by cell-free DNA. *Nat Commun*. 2023;14:2017.
40. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583–9.
41. Cheng J, Novati G, Pan J, Bycroft C, Žemgulytė A, Applebaum T, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*. 2023;381:eadg7492.
42. Shafi S, Parwani AV. Artificial intelligence in diagnostic pathology. *Diagn Pathol*. 2023;18:109.
43. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24:1559–67.
44. Chen M, Zhang B, Topatana W, Cao J, Zhu H, Juengpanich S, et al. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. *NPJ Precis Oncol*. 2020;4:14.
45. Nasrallah MP, Zhao J, Tsai CC, Meredith D, Marostica E, Ligon KL, et al. Machine learning for cryosection pathology predicts the 2021 WHO classification of glioma. *Med*. 2023;4:526–40.
46. Manafi-Farid R, Askari E, Shiri I, Pirich C, Asadi M, Khateri M et al. [18F] FDG-PET/CT radiomics and artificial intelligence in lung cancer: technical aspects and potential clinical applications. In: *Seminars in nuclear medicine* Nov 1 (Vol. 52, pp. 759–80). WB Saunders. 2022
47. Timmerman L. What's in a name? A lot, when it comes to precision medicine'. *Xconomy*. 2013. Available from: <https://xconomy.com/>
48. Johannet P, Coudray N, Donnelly DM, Jour G, Illa-Bochaca I, Xia Y, et al. Using machine learning algorithms to predict immunotherapy response in patients with advanced melanoma. *Clin Cancer Res*. 2021;27:131–40.
49. Loh E. Medicine and the rise of the robots: a qualitative review of recent advances of artificial intelligence in health. *BMJ Leader*. 2018;2:59–63.
50. Ren F, Ding X, Zheng M, Korzinkin M, Cai X, Zhu W, et al. AlphaFold accelerates artificial intelligence powered drug discovery: efficient discovery of a novel CDK20 small molecule inhibitor. *Chem Sci*. 2023;14:1443–52.
51. Chang Y, Park H, Yang HJ, Lee S, Lee KY, Kim TS, et al. Cancer drug response profile scan (CDRscan): a deep learning model that predicts drug effectiveness from cancer genomic signature. *Sci Rep*. 2018;8:8857.
52. Corti G, Bartolini A, Crisafulli G, Novara L, Rospo G, Montone M, et al. A genomic analysis workflow for colorectal cancer precision oncology. *Clin Color Cancer*. 2019;18:91–101.
53. Choi J, Park S, Ahn J. RefDNN: a reference drug based neural network for more accurate prediction of anticancer drug resistance. *Sci Rep*. 2020;10:1861.
54. Mobadersany P, Yousefi S, Amgad M, Gutman DA, Barnholtz-Sloan JS, Velázquez Vega JE, et al. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci*. 2018;115:E2970–9.
55. Guo Q, Lakatos E, Bakir IA, Curtius K, Graham TA, Mustonen V. The mutational signatures of formalin fixation on the human genome. *Nat Commun*. 2022;13:4487.
56. IT Governance. WannaCry Ransomware attack costs £92 million [Internet]. (2018). Available from: <https://www.itgovernance.co.uk/blog/wannacry-ransomware-attack-costs-92-million>. [Accessed 15 Oct 2024].

57. Prince AE. Insurance risk classification in an era of genomics: is a rational discrimination policy rational? *Neb Law Rev.* 2017;96:624.
58. Department of Health and Social Care. Code on Genetic Testing and Insurance [Internet]. London: GOV.UK; 2018 [accessed 20/12/24]. Available from: <https://www.gov.uk/government/publications/code-on-genetic-testing-and-insurance>.
59. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digital Imaging.* 2013;26:1045–57.
60. Farquhar S, Kossen J, Kuhn L, Gal Y. Detecting hallucinations in large language models using semantic entropy. *Nature.* 2024;630:625–30.
61. Crigger E, Reinbold K, Hanson C, Kao A, Blake K, Irons M. Trustworthy augmented intelligence in health care. *J Med Syst.* 2022;46:12.
62. Greatbatch O, Garrett A, Snape K The impact of artificial intelligence on the current and future practice of clinical cancer genomics. *Genetics Research.* Cambridge University Press; 101:e9 (2019)
63. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A Learning deep features for discriminative localization. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016* (pp. 2921-9).
64. Lumaka A, Cosemans N, Lulebo Mampasi A, Mubungu G, Mvuama N, et al. Facial dysmorphism is influenced by ethnic background of the patient and of the evaluator. *Clin Genet.* 2017;92:166–71.
65. Zhuravlev AY, Wood RA. The two phases of the Cambrian Explosion. *Sci Rep.* 2018;8:16656.
66. Pratt GA. Is a Cambrian explosion coming for robotics? *J Econ Perspect.* 2015;29:51–60.
67. Gupta A, Singla T, Chennatt JJ, David LE, Ahmed SS, Rajput D. Artificial intelligence: a new tool in surgeon's hand. *J Educ Health Promot.* 2022;11:93.
68. Brynjolfsson E, McAfee A *The second machine age: Work, progress, and prosperity in a time of brilliant technologies.* WW Norton & Company; 2014.
69. British Society for Genetic Medicine. *BSGM Annual Conference 2024* [Internet]. (2024). Available from: <https://bsgm.org.uk/events/bsgm-annual-conference-2024/> [Accessed 15 Oct 2024].

## ACKNOWLEDGEMENTS

This paper is based on an article submitted to the UK Cancer Genetics Group (UKCGG) 2023 Essay Prize "Increasing use of artificial intelligence in genomic medicine for cancer care- the promise and potential pitfalls." We are grateful for the support from UKCGG to submit this piece as a journal article. The AI language model Chat-GPT 3.5 was used to generate the opening quote of this review paper. Figure 1

was created using by the author using Biorender.com (<https://www.biorender.com>). Figure 2 was created by Samara Ona, Jordan Ashby and Juno Shemano with images created in biorender.com.

## AUTHOR CONTRIBUTIONS

OO researched and wrote the article. TM conceived the review topic and revised the manuscript. All authors reviewed the manuscript before submission.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Olivia O'Connor.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.