



## OPEN ACCESS

## EDITED BY

Adnan Haider,  
Dongguk University Seoul, Republic of Korea

## REVIEWED BY

Muhammad Toseef,  
City University of Hong Kong, China  
Pradnya Mehta,  
Vishwakarma Institute of Technology, India

## \*CORRESPONDENCE

Ruby Dhar  
✉ [rubydhar@gmail.com](mailto:rubydhar@gmail.com)  
Subhradip Karmakar  
✉ [subhradip.k@aaiims.edu](mailto:subhradip.k@aaiims.edu)

RECEIVED 23 December 2024

ACCEPTED 17 February 2025

PUBLISHED 04 March 2025

## CITATION

Goel I, Bhaskar Y, Kumar N, Singh S,  
Amanullah M, Dhar R and Karmakar S (2025)  
Role of AI in empowering and redefining the  
oncology care landscape: perspective from a  
developing nation.  
Front. Digit. Health 7:1550407.  
doi: 10.3389/fdgth.2025.1550407

## COPYRIGHT

© 2025 Goel, Bhaskar, Kumar, Singh,  
Amanullah, Dhar and Karmakar. This is an  
open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with  
these terms.

# Role of AI in empowering and redefining the oncology care landscape: perspective from a developing nation

Isha Goel<sup>1,2</sup>, Yogendra Bhaskar<sup>3</sup>, Nand Kumar<sup>2</sup>, Sunil Singh<sup>1</sup>,  
Mohammed Amanullah<sup>4</sup>, Ruby Dhar<sup>1\*</sup> and Subhradip Karmakar<sup>1\*</sup>

<sup>1</sup>Department of Biochemistry, All India Institute of Medical Sciences (AIIMS), New Delhi, India,

<sup>2</sup>Department of Psychiatry, All India Institute of Medical Sciences (AIIMS), New Delhi, India,

<sup>3</sup>ICMR Computational Genomics Centre, Indian Council of Medical Research (ICMR), New Delhi, India,

<sup>4</sup>Department of Clinical Biochemistry, College of Medicine, King Khalid University, Abha, Saudi Arabia

Early diagnosis and accurate prognosis play a pivotal role in the clinical management of cancer and in preventing cancer-related mortalities. The burgeoning population of Asia in general and South Asian countries like India in particular pose significant challenges to the healthcare system. Regrettably, the demand for healthcare services in India far exceeds the available resources, resulting in overcrowded hospitals, prolonged wait times, and inadequate facilities. The scarcity of trained manpower in rural settings, lack of awareness and low penetrance of screening programs further compounded the problem. Artificial Intelligence (AI), driven by advancements in machine learning, deep learning, and natural language processing, can profoundly transform the underlying shortcomings in the healthcare industry, more for populous nations like India. With about 1.4 million cancer cases reported annually and 0.9 million deaths, India has a significant cancer burden that surpassed several nations. Further, India's diverse and large ethnic population is a data goldmine for healthcare research. Under these circumstances, AI-assisted technology, coupled with digital health solutions, could support effective oncology care and reduce the economic burden of GDP loss in terms of years of potential productive life lost (YPPLL) due to India's stupendous cancer burden. This review explores different aspects of cancer management, such as prevention, diagnosis, precision treatment, prognosis, and drug discovery, where AI has demonstrated promising clinical results. By harnessing the capabilities of AI in oncology research, healthcare professionals can enhance their ability to diagnose cancers at earlier stages, leading to more effective treatments and improved patient outcomes. With continued research and development, AI and digital health can play a transformative role in mitigating the challenges posed by the growing population and advancing the fight against cancer in India. Moreover, AI-driven technologies can assist in tailoring personalized treatment plans, optimizing therapeutic strategies, and supporting oncologists in making well-informed decisions. However, it is essential to ensure responsible implementation and address potential ethical and privacy concerns associated with using AI in healthcare.

## KEYWORDS

artificial intelligence, digital health, cancer management, personalized medicine, developing nation

# 1 Introduction

## 1.1 Cancer care in India: challenges and solutions

According to The World Bank report, 2024, India has become the most populous country surpassing China in 2022, with over 1.417 billion inhabitants, the majority of whom were under 25 (1). The size and growing pace of the Indian population present various concerns. It puts an enormous strain on healthcare, education, infrastructure, environment, and employment opportunities. Providing proper healthcare services to such a large population is a significant challenge.

Cancer is a foremost cause of death worldwide due to its high mortality rate. According to the International Agency for Research on Cancer (IARC), in the year 2022, mortality and incidence count were 9.7 million and 19.9 million, respectively. Number of new cancer cases are speculated to rise to 35 million in 2050 (2). The Indian population alone accounts for incidence and mortality rates of 1.4 million and 0.9 million, respectively, in 2022 (Figure 1a,b). Most of the deaths reported in India were from breast cancer and cervix uterine cancer in females. Whereas, cancers of the lip/oral cavity, lung, oesophagus and stomach topped the list in Indian males (Figure 1c). Oncology care is, therefore one of the fastest growing therapeutic fields. Early diagnosis and prognosis of a cancer type have become a necessity in oncology research, as it can facilitate the subsequent clinical management and prevention of death in cancer patients.

**Cancer Burden:** With more than one million cancers diagnosed every year in India, early identification and management are critical for developing a robust oncology care system in the country. This urgency is amplified by the vast population and the resultant burden on the healthcare infrastructure. However, population demand for healthcare services outstrips available resources in India, resulting in congested healthcare facilities, excessive waiting times, and often substandard services.

**Rural-Urban Disparities:** Rural communities frequently have insufficient healthcare facilities and poor infrastructure resulting in healthcare inequities across urban and rural populations. In the absence of proper logistics, people, particularly those in remote areas, are often deprived of quality healthcare services. For instance, the lack of transportation options and healthcare centres within a reasonable distance can lead to delayed diagnoses and treatments, significantly affecting patient outcomes.

**Shortage of Healthcare Specialists:** To meet the people's demands, an adequate number of healthcare specialists is also necessary. Yet, nurses, doctors, pathologists, specialized

oncologists, and other medical personnel are in limited supply in India. Considering the number of allopathic doctors, the doctor-to-patient ratio is much lower than the World Health Organisation's (WHO) recommended guidelines of 1 per 1,000 population (3, 4). This scarcity can result in increasing workloads, fatigue among healthcare personnel, and a reduction in the quality of care, impacting the timely diagnosis and treatment of a significant number of patients.

## 1.2 How artificial intelligence can assist

Artificial Intelligence (AI) can assist India and other countries in overcoming healthcare challenges brought on by their large population. With AI and machine learning solutions, any large-scale cancer intervention may be successfully implemented, apart from making it more accurate and accessible. For instance, AI-driven diagnostic tools can analyse vast quantities of medical images and patient data to detect cancer at earlier stages with higher precision than traditional methods. These tools enable healthcare providers to identify subtle patterns and anomalies that might escape the human eye, significantly improving early detection rates. Similarly, AI-assisted surgery, which integrates robotics and machine learning, can enhance precision, reduce recovery times, and improve surgical outcomes.

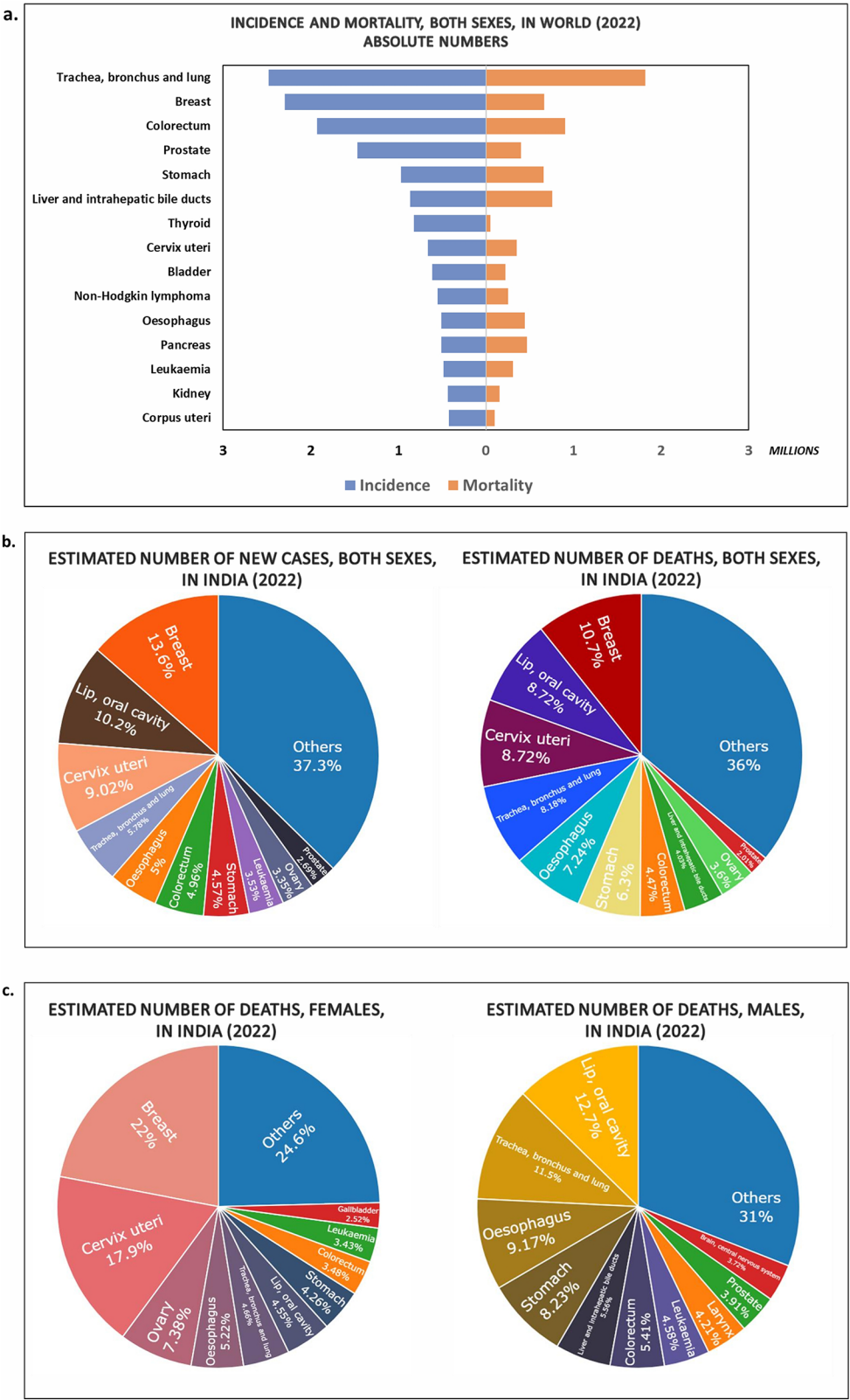
AI also alleviates the burden on healthcare professionals by automating routine and repetitive tasks. Tasks such as patient data entry, appointment scheduling, and medication management can be streamlined through AI, allowing medical staff to focus more on patient care and less on administrative duties. Additionally, AI-powered monitoring tools can track patient conditions in real-time, alerting healthcare providers to critical changes, thereby enhancing patient safety and reducing the workload on medical teams.

AI-powered telemedicine platforms are particularly transformative in bridging the gap between patients and healthcare providers, especially in remote and underserved areas. These platforms enable patients to access medical consultations, follow-ups, and diagnostic services without the need for long-distance travel, ensuring that quality healthcare reaches even the most marginalized populations.

However, realizing AI's transformative potential in healthcare requires addressing several challenges. Key among these are the development of robust data infrastructure, the implementation of stringent data privacy regulations, and the establishment of secure, ethical data-sharing practices to protect patient confidentiality. Additionally, significant investments in

### Abbreviations

AI, artificial intelligence; ANI, artificial narrow intelligence; A3C, asynchronous advantage actor-critic; AR, augmented reality; BCSC, breast cancer surveillance consortium; CBE, clinical breast examination; CDSS, clinical decision support system; CNNs, convolutional neural networks; DL, deep learning; DDR1, discoidin domain receptor 1; EHR, electronic health records; HCC, hepatocellular carcinoma; HPV, human papilloma virus; IARC, International Agency for Research on Cancer; LINC, Library of Integrated Network-Based Cellular Signatures; LLMs, large language models; ML, machine learning; MCTS, Monte Carlo Tree Search; MLPs, multilayer perceptrons; MLLMs, multimodal large language models; NLP, natural language processing; PCAWG, Pan-Cancer Analysis of Whole Genomes; PRS, polygenic risk scores; PCA, principal component analysis; RCTs, randomized controlled trials; RNNs, recurrent neural networks; SNPs, single nucleotide polymorphisms; SVM, support vector machine; TCGA, The Cancer Genome Atlas; VR, virtual reality; WHO, World Health Organisation's; YPLL, years of potential productive life lost.



**FIGURE 1** Charts showing the estimates of new cases and deaths due to different types of cancer during the year 2022 in (a) Worldwide (Top 15 cancer sites); (b) India (both sexes) (Top 10 cancer sites); (c) deaths in India males, females (Top 10 cancer sites) (Data Source: GLOBOCON 2022; International Agency for Research on Cancer).

technological infrastructure and the development of a skilled AI workforce are critical for effective implementation.

This review offers a comprehensive examination of AI's applications across various aspects of cancer management, including prevention, diagnosis, precision treatment, prognosis, and drug discovery. It also provides a detailed discussion on the roadmap and recommendations for integrating AI and telemedicine into India's oncology care landscape, while addressing the inherent challenges of AI adoption and highlighting recent advancements that tackle these critical issues. To refer to the detailed workflow methodology for article collection, please see [Box 1](#).

### 1.3 What is artificial intelligence

AI is a field of research in which smarter algorithms are applied to mimic human intelligence (5). AI seeks to make it possible for machines to emulate and mimic human cognitive functions such as learning, reasoning, problem-solving, perception, and decision-making. Further, it can perform multitasking surpassing human abilities. It is a multifaceted field encompassing numerous specialized areas, such as (i) Natural Language Processing (NLP), it focuses on understanding and generating human language, covering applications like sentiment analysis and chat-bots (6); (ii) Computer Vision, it deals with visual data, enabling tasks such as object detection and facial recognition (7); (iii) Robotics, it

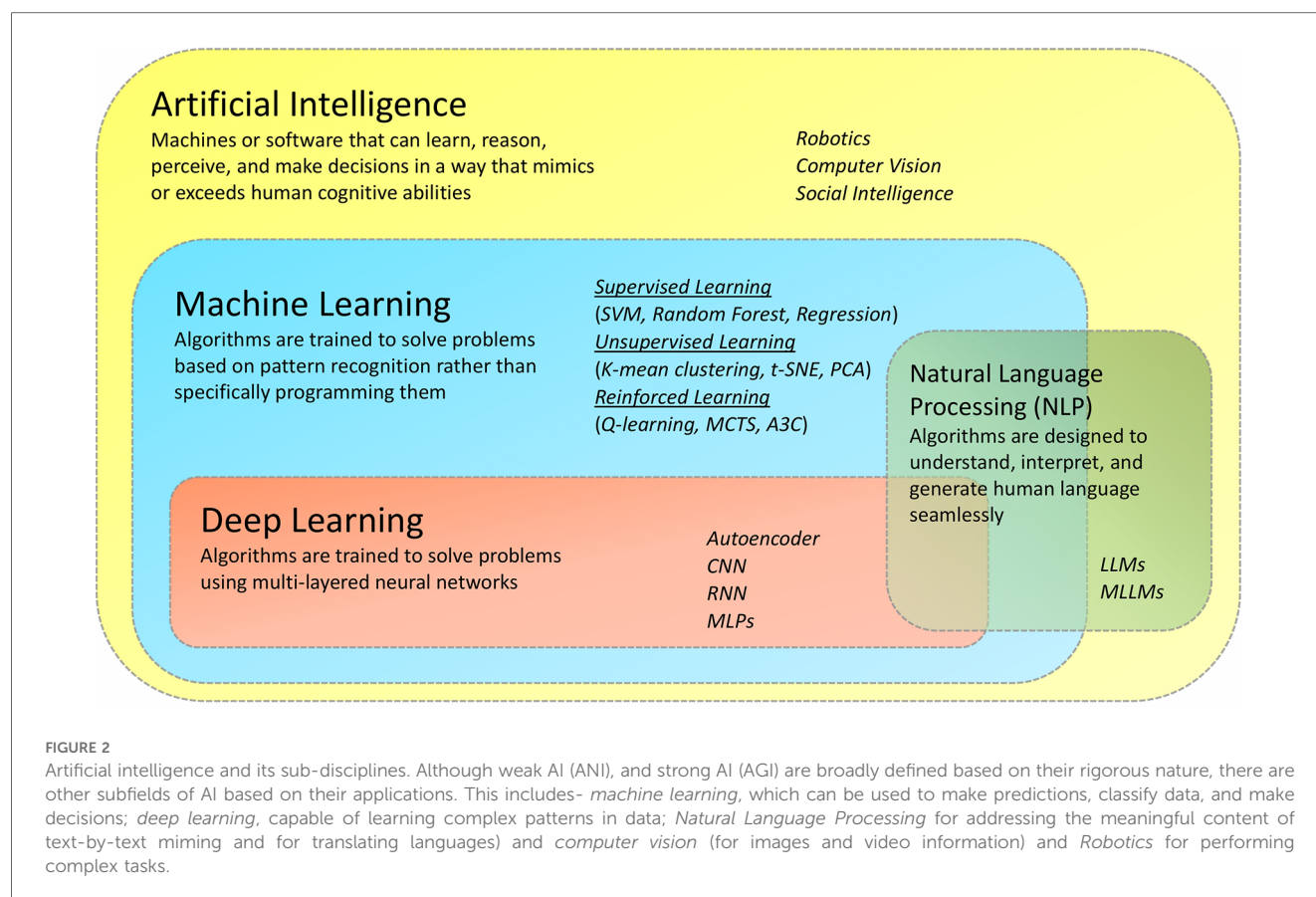
combines hardware and software to create autonomous machines. Robotics has the potential to revolutionize industries like healthcare, transportation, and manufacturing (8); (iv) Social intelligence, is the ability of artificial systems to understand, interact with, and respond to humans in a socially and emotionally intelligent manner such as emotional recognition and feedback & learning (9). Machine learning (ML) is a sub-discipline of AI, that employs sophisticated algorithms on large-scale heterogeneous datasets to uncover valuable patterns that would otherwise be difficult or impossible to identify even for highly trained individuals (10). It is built upon three fundamental approaches or paradigms. Supervised Learning, the most common of these paradigms, involves training algorithms on labelled data to predict outcomes or make classifications on new, unseen data [popular models: Support Vector Machine (SVM), Random Forest and Regression]. On the other hand, unsupervised Learning operates without labelled data, instead uncovering hidden patterns and structures within datasets, frequently employed in clustering and dimensionality reduction tasks (popular models: K-means, t-SNE and PCA). Reinforcement Learning, the third paradigm, focuses on sequential decision-making, where an agent learns to interact with an environment to maximize cumulative rewards through trial and error [popular models: Q-learning, Monte Carlo Tree Search (MCTS) and Asynchronous Advantage Actor-Critic (A3C)] (11). Deep Learning (DL) is a subset of ML that uses neural network-based models to simulate the human brain's capacity for processing enormous amounts of complex data, including, but not limited to, image recognition, language processing, and drug discovery, all of which serve as a decision support system for researchers (Figure 2). Deep learning comprises a wide array of approaches and architectures that have propelled machine learning to new heights. Among these, Multilayer Perceptrons (MLPs) serve as the foundational building blocks, suitable for diverse tasks like regression and classification. Convolutional Neural Networks (CNNs) excel in processing grid-like data such as images, while Recurrent Neural Networks (RNNs) are designed for sequential data like time series and text. Transformers, with their self-attention mechanisms, have revolutionized natural language understanding. Autoencoders, another vital component, are employed for unsupervised learning and feature extraction, finding applications in dimensionality reduction and image denoising (12). Multimodal Large Language Models (MLLMs) and Large Language Models (LLMs) are part of both the Deep Learning and Natural Language Processing (NLP) domains. Large Language Models (LLMs) are designed to understand, generate, and analyse human language, significantly enhancing natural language processing tasks. MLLMs extend beyond LLMs by integrating multiple types of data, such as text, images, and audio, allowing them to handle multimodal information (13). These diverse deep-learning approaches are tailored to various data types and problem domains, marking them as the forefront of modern AI research and applications (Table 1). Since the

#### BOX 1 Workflow methodology for article collection.

The purpose of this review was to analyse the role and impact of artificial intelligence (AI) in the field of oncology. A comprehensive search was conducted using the PubMed academic database, employing search query with keywords such as “(((Artificial Intelligence) OR (Machine Learning)) AND ((Cancer) OR (Oncology)))” in the title or abstract of the articles published during last 10 years. Initially, 12,856 articles published had been identified. 41 more articles were included from citation search and other sources.

Articles with non-relevant study designs (Interviews, Autobiography, Editorials, Commentaries, Video-Audio Media) were excluded ( $n = 5,955$ ). The collection was further refined based on four exclusion criteria, removing articles in the following categories: (i) non-English language (100 articles), (ii) Restricted access (2,404 articles), (iii) non-relevant articles (3,457 articles) and (iv) Redundant Findings/Applications/Methodology (741 articles). After applying these filters, 240 relevant articles remained, which were then analysed to understand the impact and role of AI in oncology (Figure 3).





human mind is only able to analyse a finite amount of data in a short duration of time, the exponential rise of AI over the last decade implies that it can serve as a sophisticated platform supporting human experts or clinicians to make the best possible decisions.

## 2 Artificial intelligence in oncology care

The AI approaches are already used widely in science and society for everything from clinical trials to robotics to self-driving cars. The power of machine learning has been demonstrated in healthcare by work on the human genome project, initiatives in cancer omics [such as The Cancer Genome Atlas (37), the International Cancer Genome Consortium, and the Clinical Proteomic Tumour Analysis Consortium], as well as numerous international machine learning competitions like the DREAM challenges (38, 39). When it comes to cancer detection, diagnosis, and treatment, the ability to gather and analyse sizable datasets on medical treatments and outcomes holds great promise for transforming medicine into a data-driven, outcome-oriented field. Early usages of machine learning in diagnosis and treatment have shown promise in detecting breast cancer from x-rays (40, 41), discovering new antibiotics (42), predicting gestational diabetes onset from electronic health records (43), and identifying clusters

of patients who share a molecular signature of treatment response (44). AI-based algorithms hold great promise to pave the way to identify genetic mutations and aberrant protein interactions that can lead to disease diagnosis at a very early stage. AI-powered telemedicine and remote diagnosis can also aid in treatment planning and resource optimization for patients based on age, medical history, and genetic makeup. This personalized approach to disease treatment has the potential to improve patient outcomes. Government institutions as well as numerous other healthcare groups, find enormous potential in AI for enhancing cancer care.

In the context of oncology, with the advancement of technology more and more clinical data is being generated every day, and AI has become essential for the automated analysis of massive patient data, such as biomedical images, genetic profiles, health reports, molecular tests, etc. to detect abnormalities/patterns and forecast the possibility of cancer or response to treatment. AI can improve the efficiency and accuracy of personalized risk assessment, early diagnosis, staging, therapeutics, clinical decision-making, prognosis, and survival predictions (Figure 5). In India, institutions like the All-India Institute of Medical Sciences (AIIMS) and various state-run healthcare facilities are exploring AI's potential to enhance cancer care. We provide here an overview of some major areas of oncology and research where AI can have substantial influence (Table 2).

TABLE 1 Comparison of various AI models/algorithms with their applications in healthcare.

Algorithm/Model	Description	Applications in healthcare
<b>Supervised learning</b>		
Decision trees (CART, C4.5)	Decision trees are models that use a tree-like graph of decisions and their possible consequences. Each node represents a decision point, and the branches represent the outcomes. They are intuitive and easy to interpret.	Diagnostic decision support, predicting disease progression (14).
Random forests (RandomForestClassifier)	Random forests are ensembles of decision trees that aggregate the results of multiple trees to improve accuracy and control overfitting. They are robust against noise and have high predictive performance.	Patient risk assessment, predicting treatment outcomes (15).
Support vector machines (SVM)	SVMs are supervised learning models that find the hyperplane that best separates data into classes. They are particularly effective for high-dimensional data.	Disease classification, identifying biomarkers (16, 17).
<b>Unsupervised learning</b>		
K-means algorithm	K-means clustering is an algorithm that partitions data into k clusters, where each data point belongs to the cluster with the nearest mean. It is simple and efficient.	Patient segmentation, clustering symptoms for diagnosis (18).
DBSCAN algorithm	DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is an algorithm that forms clusters based on density. It can find clusters of arbitrary shape.	Anomaly detection, identifying outliers in medical data (19, 20).
Principal component analysis (PCA) algorithm	PCA is a dimensionality reduction technique that transforms data into principal components, capturing the most variance with the fewest components.	Data visualization, noise reduction in medical data (21).
<b>Reinforcement learning</b>		
Q-Learning	Q-Learning is a reinforcement learning algorithm that learns the value of actions in states to achieve the highest cumulative reward.	Personalized treatment planning, optimizing clinical workflows (22).
Deep Q-network (DQN)	DQN combines Q-Learning with deep neural networks to handle high-dimensional data. It is effective in applications where complex decision-making is needed.	Robotic surgery, real-time treatment adaptation (23).
Proximal policy optimization (PPO)	PPO is a policy gradient method for reinforcement learning that balances exploration and exploitation. It provides stable training.	Adaptive therapies, dynamic treatment strategies (24).
<b>Neural networks</b>		
Convolutional neural network (CNN) (LeNet, AlexNet, VGG, Inception)	CNNs are specialized neural networks for processing grid-like data such as images. They consist of convolutional layers that extract features from images and are highly effective in medical image analysis.	Detecting tumors in radiology images (25).
Recurrent neural network (RNN) (LSTM, GRU)	RNNs are neural networks designed for sequential data. They use loops to maintain context over sequences and are used in healthcare for analysing patient records and predicting health outcomes over time.	Time-series health predictions (26).
Multilayer perceptron (MLP) (feedforward neural network, MLP classifier)	MLPs are classic neural networks with multiple layers of neurons. They are versatile and used in various healthcare applications.	Predicting disease outbreak, drug response prediction (27).
Autoencoder (variational, sparse autoencoder)	Autoencoders are neural networks that learn to encode data efficiently, reducing its dimensions and reconstructing it from the encoding.	They are used for detecting anomalies in patient data and compressing large datasets (28).
<b>Natural language processing (NLP)</b>		
Bidirectional encoder representations from transformers (BERT)	BERT is a transformer-based model that pre-trains on large text corpora to understand context.	It enhances clinical documentation and extracts insights from electronic health records (EHRs) (29).
<b>Large language models (LLMs)</b>		
Generative pre-trained transformer (GPT3,4)	GPT is a language model designed for text generation and understanding.	Text-based clinical decision support, medical literature review (30, 31).
Text-to-text transfer transformer (T5)	T5 converts all NLP problems into a text-to-text format.	Summarizing medical documents, translating medical texts (32).
<b>Multimodal large language models (MLLMs)</b>		
Contrastive language-image pre-training (CLIP)	CLIP integrates text and image data, allowing for advanced multimodal analysis.	Analysing medical images with textual reports, multimodal diagnostics (33).
Multi-modal vision and language model (VisualBERT)	VisualBERT combines visual and textual information, enabling tasks like visual question answering and generating image-based medical reports. It captures relationships between images and text effectively.	Visual question answering, generating image-based medical reports (34).
Unified transformer (UniT)	UniT processes multiple data modalities, making it versatile for complex healthcare tasks that require integrating diverse data types.	Multimodal data interpretation, combining patient imaging and clinical notes (35).
<b>Robotics</b>		
Robotics (Da Vinci surgical system, intuitive surgical)	The use of programmable machines to perform tasks autonomously or semi-autonomously.	Surgical assistance, patient rehabilitation, telemedicine, diagnostics, and stroke rehabilitation (36).

TABLE 2 Applications of AI in oncology care.

Applications of AI	Description	Benefits
1. Predictive analytics	Accurate prediction of cancer occurrence in high-risk individuals and timely identification of asymptomatic cancer patients utilizing AI-based algorithms trained on diverse datasets, including: <ol style="list-style-type: none"> <li>Medical imaging (Breast Mammograms, Chest x-Rays),</li> <li>Large-scale genomic datasets (Cancer-related mutations/variants, metabolic biomarkers, signatures of DNA methylation, expression profiles of cancer-related genes or non-coding RNAs),</li> <li>Comprehensive electronic health records data (Patients' medical history, family history, demographics, test reports)</li> </ol>	Enables early intervention, potentially reducing cancer incidence. Improves patient outcomes and survival rates through timely treatment.
2. Cancer diagnosis and staging	Accurate diagnosis of cancer and classification into Grade/Stage & Primary Site/Histological Type utilizing AI-driven platforms that can integrate multimodal datasets such as: <ol style="list-style-type: none"> <li>Radiological Images (CT, x-rays, MRI, PET scans, and ultrasound),</li> <li>Videos (such as colonoscopy for malignant polyps),</li> <li>Fundus images for the identification of retinoblastomas,</li> <li>Digitized H&amp;E-stained slides from core needle biopsies,</li> <li>Dermoscopy images for Skin Cancer,</li> <li>Large-scale genomic data (mutational signatures, gene expression, DNA methylation, miRNA profiles), vii) Peripheral blood biomarkers (protein biomarkers, circulating cell-free tumour DNA) from liquid biopsies</li> </ol>	Enhances early detection, leading to better treatment outcomes. Reduces diagnostic errors and enhances accuracy.
3. Treatment and response assessment	Personalized Treatment: Clinical Decision Support System (CDSS) for data-driven personalized cancer therapy recommendations, integrating automated radiation therapy plans based on radiological data, genomic information, and medical records. AI-Assisted Surgery: Enhanced preoperative surgical planning, intraoperative navigation with real-time feedback using AR/VR technologies, and postoperative monitoring using data feeds from monitoring devices. AI-programmed robotic surgery for increased accuracy and minimizing human error. Prognosis/Survival Prediction: Enhancing precision in forecasts for chemoresistance, cancer recurrence, and survival outcomes using multimodal AI algorithms that can integrate multi-omics data and other complex multi-dimensional datasets.	Optimizes treatment plans for individual patients. Enhances treatment efficacy and minimizes adverse effects. Improves overall patient care and satisfaction.
4. Telemedicine/Remote monitoring	Apps/Websites: AI-powered chatbots and virtual assistants for preliminary assessments, symptom tracking, diagnosis, connecting with a specialised doctor (telemedicine), tele-education, basic medical advice, medication reminders, community support. Wearable Devices: AI-powered real-time remote monitoring of behavioural patterns (Sleeping habits, physical activity, heart rate, stress levels) and Cell phone images for remote patient monitoring, cancer risk prediction, ultra-personalized diagnosis and treatment recommendations.	Improved access to healthcare services, especially in remote or underserved areas. Reduces the need for in-person visits and hospital readmissions.
5. Novel discovery	Drug Targets/Biomarkers: AI-driven prediction of therapeutic targets in oncology utilizing population-specific publicly available mutation/loss-of-function screen profiles, gene expression datasets, and clinical data. Drug Designing: AI-powered tools like AlphaFold, DeepNeuralNetQSAR, DeepChem, DeepTox, and gene2drug for the prediction and identification of drug molecules with enhanced biochemical properties for improved sensitivity and specificity, alongside a reduced risk of adverse effects. Additionally, optimisation of existing drugs with enhanced biochemical properties for better performance. Clinical Trial: Enhanced patient eligibility decisions, automated patient recruitment processes, optimized trial design, and real-time monitoring of patient adherence to protocols for significantly improved overall efficiency and effectiveness of cancer clinical trials.	Speeds up the drug discovery process, leading to innovative treatments. Reduces research and development costs and time with improved success rates.

## 2.1 AI for cancer risk predictive analytics and early intervention

The first step in managing cancer is to prevent it, and the key to that is to identify those who are most at risk of contracting the disease. A quick and early mass screening is required for a population to address its health priorities. As a result, automated AI-based methodologies are now required over manual ones like breast examinations, Pap Smears, HPV tests or visual inspections. Incorporation of AI-based programs in routine healthcare can

transform cancer care in India by identifying high-risk asymptomatic individuals so that timely preventative policies like vaccinations and lifestyle modifications can be devised and recommended.

Researchers globally, including those in India, have successfully developed AI-based programs to analyse diverse forms of biomedical imaging data for the early prediction of cancer. An AI-based algorithm was implemented (53), in a study published in the Journal of the National Cancer Institute to forecast the risk of breast cancer based on mammograms and clinical data.

According to Rodriguez-Ruiz et al., the algorithm had a 70.4% accuracy rate in predicting women at a high risk for developing breast cancer within 5 years. In a recently published study, it was concluded that integrating 5 different AI-based deep learning algorithms outperformed cancer risk as predicted by the Breast Cancer Surveillance Consortium (BCSC). This was further enhanced by combining AI and BCSC models (54). Specifically in Breast Cancer screening, tools like ScreenPoint Medical's Transpara, CureMetrix's cmAssist, Qlarity Imaging's QuantX AI algorithms and BUCAD claim to support the identification of cancer-specific indicators through mammography, MRI or ultrasound images for improved accuracy and efficiency of early detection (55–58). Various other studies have also assessed the capability of mammography-trained AI systems to show significant gains in risk prediction compared to clinical risk models alone (59, 60). For a detailed overview of the AI-workflow for image-based classification of breast mammograms, please refer to **Box 2**. A deep learning-based CNN model could enhance screening by forecasting the likelihood of lung cancer among asymptomatic smokers using chest x-rays and limited EHR data (61). An individual's predisposition to developing certain cancers can also be predicted using AI by analysing large-scale genomic datasets (62). Google created a deep learning-based AI algorithm called DeepVariant to discover genomic variants from sequencing data. It has been used in cancer genomics to accurately identify cancer-related mutations and forecast the likelihood of developing the disease. Breast and lung cancer are just two cancer types where the algorithm has demonstrated promising results (63). In one of their research, Listgarten et al. created an SVM machine learning model to utilize single nucleotide polymorphisms (SNPs) profiles of steroid metabolizing enzymes (CYP450s) to precisely predict the development of "spontaneous" breast cancer. AI-generated Polygenic Risk Scores (PRS) to evaluate a person's chance of acquiring certain cancer has been developed using genetic data from numerous variants. PRS has been successfully employed in the assessment of multiple cancers and is shown to support individualized risk assessment by categorizing people into several risk groups (64–66). Other than genomic variants, metabolic biomarkers, signatures of DNA methylation and expression profiles of cancer-related genes or non-coding RNAs can predict an individual's predisposition to certain cancers or their response to a treatment regime. Employing machine learning strategies to forecast the likelihood of acquiring cancer based on the above metrics is increasingly becoming common in oncology care (67–72). AI can also find patterns and relationships associated with cancer risk using comprehensive electronic health records data. Patients' medical history, family history, medication history, demographics, age, gender, clinical notes, test reports, symptoms, and treatment outcomes can all be fed into an AI model to predict the likelihood of developing cancer. Automated models to predict breast and lung cancer (42, 73–75), using a combination of the above-mentioned clinical records are already being developed and successfully implemented.

**BOX 2** AI-workflow for image-based classification of breast mammograms.

The basic workflow begins with data collection from breast cancer biobanks, ensuring a diverse and comprehensive dataset, such as The Cancer Imaging Archive (TCIA) (45). Next, data preprocessing steps, including normalization using tools like ImageJ (46) and augmentation methods such as rotation, scaling, and flipping via libraries like TensorFlow and Keras, are performed to enhance the quality and variability of the data (47, 48). The dataset is then split into training (80%) and test datasets (20%) using methods such as stratified sampling to enable robust model validation (49). During model selection, various architectures are considered, with a focus on convolutional neural networks (CNNs) like VGG16x (50) or ResNet (51) due to their proficiency in image analysis. Model training involves feature extraction through convolutions and pooling layers, using frameworks like PyTorch (52) or TensorFlow (47), followed by classification. The model outputs a probability score indicating the likelihood of malignancy, which is used to categorize the mammogram as benign, malignant, or requiring further review. The trained model is evaluated against the test set using metrics such as accuracy, sensitivity, specificity, and AUC-ROC curves (49). If the performance is suboptimal, hyperparameter optimization techniques, such as grid search or random search via tools like Scikit-Learn, are applied. Conversely, a well-performing model undergoes clinical validation with real-world datasets to ensure its reliability. Finally, a monitoring and feedback loop is established to continuously assess and refine the model, incorporating new data and improving accuracy over time through continuous integration/continuous deployment (Figure 4).

Given India's resource limitations and cultural challenges, efforts are underway to create affordable and accessible cancer screening solutions (76). The Indian MegCan Care project, a collaborative AI-driven initiative by the Meghalaya government, Apollo Telemedicine Networking Foundation, and the World Economic Forum, seeks to provide free cancer screening to one million individuals. It focuses on early detection of oral, breast, cervical, lung, and esophageal cancers, aiming to enhance timely diagnosis and patient outcomes through widespread screening programs (77). Innovations like Thermalytix, a radiation-free AI tool using thermal imaging, demonstrate significant promise for population-level breast cancer screening in resource-constrained settings, boasting a positive predictive value of 81% (78).

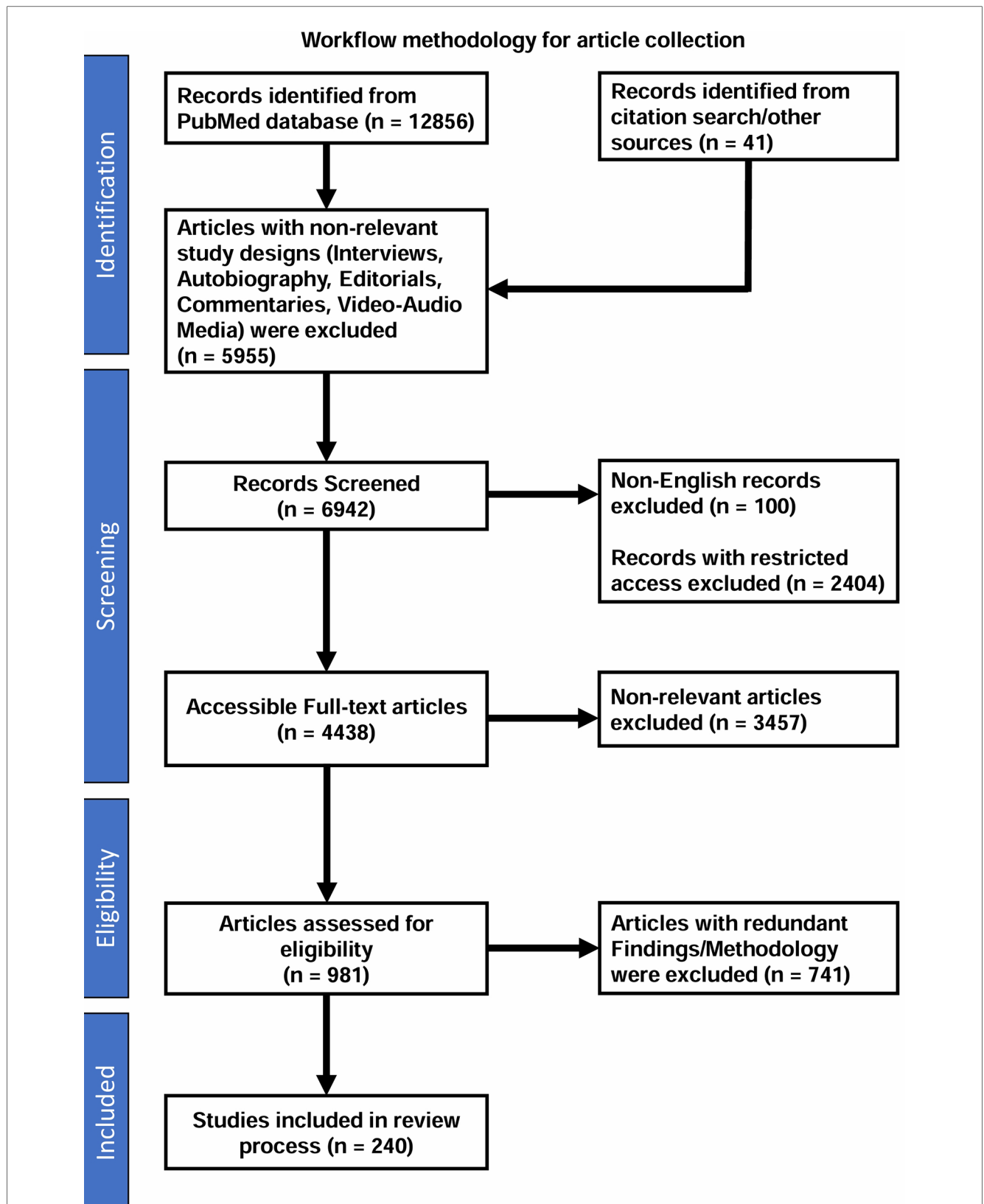


FIGURE 3

Workflow methodology for article collection. This workflow provides a detailed overview of the process used for the article collection in our review on the role and impact of AI in oncology. The workflow highlights the key stages, from identification and screening to eligibility assessment and final inclusion of relevant studies.



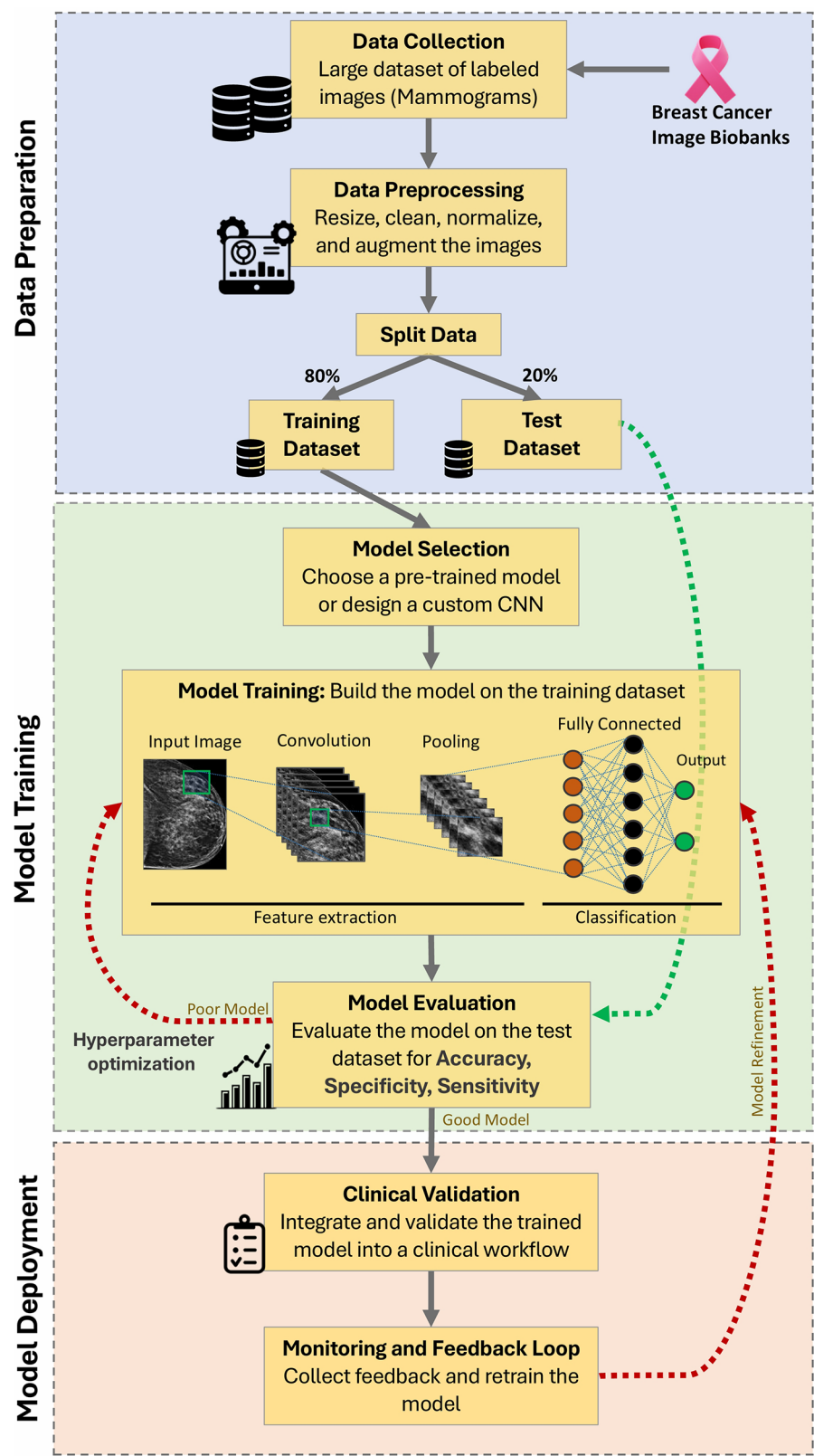


FIGURE 4  
AI workflow for image-based classification of breast mammograms (240). This schematic illustrates the step-by-step workflow of an AI system using a Convolutional Neural Network (CNN) algorithm. Key stages include data collection, preprocessing, training/testing split, model selection, feature extraction, classification, evaluation, hyperparameter optimization, clinical validation, and continuous monitoring for improved accuracy.

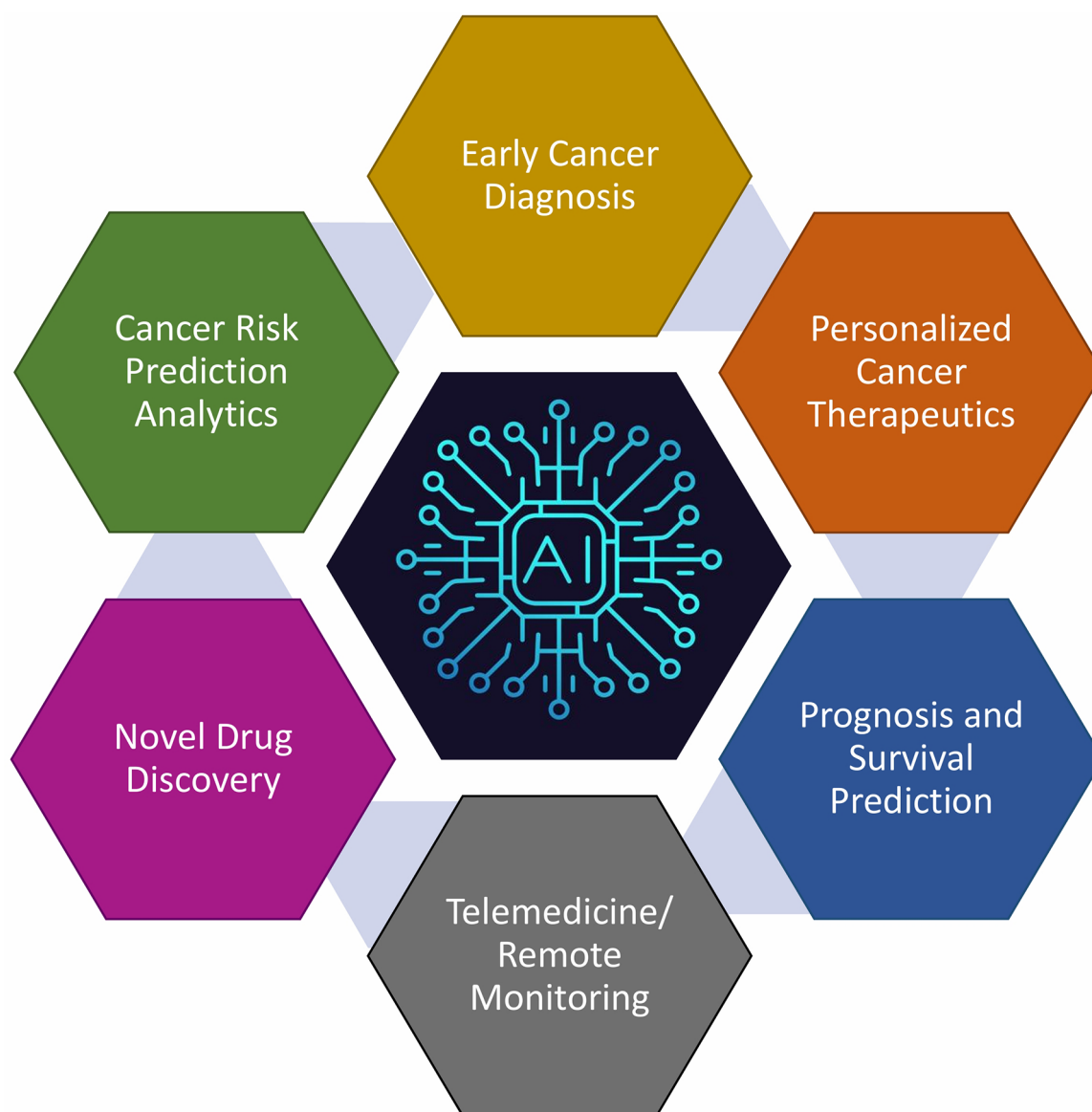


FIGURE 5

Key domains within oncology and research where artificial intelligence (AI) can significantly impact patient care and scientific advancement. While each of these domains may appear disjoint, they are in fact seamlessly integrated with AI playing a significant role in this process. AI-based tools can screen data from an enormous volume of subjects to find patterns and connections between them. This information can not only assist in personalized cancer care but can also recommend the most appropriate therapy as well as digital communications for networking a large number of centres.

## 2.2 AI for early cancer diagnosis and staging

Cancer is a heterogeneous disease harbouring a spectrum of genomic and epigenomic alterations with large variations in tumour fitness, mutational burden and responsiveness to conventional chemotherapy, and biological behaviours (both within and between cancer types). Numerous cancer forms may also exhibit comparable clinical signs or share similar symptoms along with the biomarkers employed to identify and categorize cancers, thus causing uncertainty in the diagnosis and staging of cancer. Further, due to insufficient knowledge or lack of

information, diagnosing rare cancers makes interpretation even more difficult. Additionally, because of differences in how pathologists or radiologists interpret genetic tests, histopathology slides, and imaging results, inter-observer variability affects the consistency and accuracy of cancer diagnosis and staging. The incorporation of AI can enhance the speed, accuracy, and consistency of result interpretation from radiological images, digitized pathology slides, and colonoscopy procedure footage, as well as pictures of skin lesions, to determine the presence of cancer, the subtype categorization, the grade, and the purity of the tumour. AI can assist in mass screening of symptomatic patients

for early cancer detection, galvanizing towards personalized treatment regimens by overcoming the shortage of healthcare professionals and improving accessibility to timely diagnosis.

A study demonstrated by McKinney and others (41) proved the utility of AI in the successful screening of breast cancer using mammograms. A group at Google designed a CNN-based screening model for identifying and classifying abnormal nodules from CT scans for lung cancer diagnosis with more accuracy and efficiency (79). Similarly, x-rays, MRI scans, PET scans, and ultrasound images have also been utilized to improve diagnosis, characterization as well as prediction of metastasis in several cancer types using machine learning approaches (80–83). AI algorithms have also been designed to accurately detect and categorize malignant polyps from clinical colonoscopies in real-time with high sensitivity and specificity (84, 85), and skin cancer from dermoscopy images with a level of competency comparable to dermatologists (86). Additionally, an AI algorithm based on deep learning models like CNNs and RNNs can identify the presence of cancer, grade, and assess tumour purity estimates from digitized H&E-stained slides of prostate tissue from core needle biopsies (87); classify cancer type and assess invasiveness based on breast biopsy images (88–91); identify retinoblastomas from fundus images with high specificity (92). Artificial intelligence (AI) algorithms can also learn to recognize molecular markers suggestive of cancer by comparing genomic patterns between diseased and non-cancerous tissues. These models can subsequently use the genetic profiles of new samples to categorize them, aiding in the correct detection of cancer (93–95). AI algorithms have been used to analyse large-scale genomic data, including mutational signatures (96), gene expression (97–99) DNA methylation (100, 101), and miRNA profiles (102). In 2006, NIH/NCI launched a landmark cancer genomics program “The Cancer Genome Atlas (TCGA) Research Network” with the goal of gathering, profiling, and analysing a sizable number of human clinical tumours representing various tumour types and their subtypes to identify molecular aberrations at the DNA, RNA, protein, and epigenetic levels. Integrative analysis of this data using machine learning has made it possible to identify molecular patterns, common pathways, master regulatory hubs that are activated or deactivated across many tissue types, and molecular phenotypes and biomarkers particular to different cancer stages (103–107). Using hundreds of somatic mutation data, the Pan-Cancer Analysis of Whole Genomes (PCAWG) Consortium utilized a DL model to forecast the origins of 24 cancer types both individually and collectively (108, 109). Automated screening and staging models based on NLP (AI-based) have also been developed to analyse unstructured data from clinical reports for the identification of cancer-specific abnormalities (110–112). Analysis of liquid biopsies for the identification of peripheral blood biomarkers of diagnostic relevance has recently gained interest (113–117). Integrating artificial intelligence with Raman spectroscopy, a group designed a highly accurate DL model to analyse liquid biopsies for lung cancer diagnosis (118). The CancerSEEK model is another ground breaking example of the utilization of AI in the early diagnosis of multiple cancers based on the presence of

specific genetic mutations in circulating cell-free tumour DNA and the levels of protein biomarkers using immunoassays from liquid biopsies. It demonstrated high predictive value for the number of common cancer types; especially for ovarian and liver cancer (119).

From India, iOncology.ai is a cutting-edge AI-driven platform, developed by the All-India Institute of Medical Sciences (AIIMS), New Delhi, in collaboration with the Centre for Development of Advanced Computing (CDAC), Pune, under the auspices of the Ministry of Electronics & Information Technology. Launched in 2024, the platform leverages deep learning algorithms to analyse radiological and histopathological images with high precision, facilitating early detection and personalized treatment of breast and ovarian cancers. The AI model is trained on an extensive dataset comprising approximately 500,000 images from 1,500 patient cases and is currently undergoing validation in district hospitals across India (120). In another significant initiative, the Apollo Cancer/Radiology Centre, in partnership with Google Health, is developing an AI-driven system for the early diagnosis of cancer using imaging data. This collaboration aims to harness AI technology to improve the accuracy and efficiency of cancer detection, addressing the shortage of radiologists and offering specialized screenings to enhance healthcare outcomes (121).

## 2.3 AI in cancer treatment and response assessment

Following a cancer diagnosis, the treatment outcome varies depending on the individual's genetic make-up and aggressiveness of the tumour. However, the dose and the choice of intervention (neoadjuvant, surgery, chemotherapy, radiotherapy, immunotherapy, etc.) are still empirical without a solid scientific basis. AI can assist clinicians in designing personalized and appropriate treatment plans, monitoring response to the treatment, and in predicting recurrence risk and patient survival, thus lowering healthcare costs, and workload and improving the efficiency of cancer management.

### 2.3.1 Personalized treatment

AI can successfully evaluate complicated datasets to predict cancer prognosis, thereby assisting physicians in making better-informed treatment decisions and enhancing patient care. By analysing a patient's genetic information, omics data, medical history, and treatment response data, AI algorithms can uncover novel biomarkers linked to drug response and can predict which treatments will be most effective for an individual patient. This approach of precision medicine can improve the efficacy of cancer treatments while reducing the risk of adverse drug-associated side effects. Oncologists can choose the most efficient and individualized treatment options for their patients by utilizing the power of AI to calibrate drugs and increase their pharmacological action efficacy (122–126). IBM Watson for Oncology is a well-known instance of an AI-based treatment recommendation system created by IBM in partnership with Memorial Sloan Kettering Cancer Centre to help oncologists

make data-driven treatment decisions for cancer care. This platform examines a patient's medical records, including clinical notes, pathology results, and molecular profiling data, and compares this information with a sizable collection of medical literature, trial data from clinical studies, and professional therapy recommendations. Watson for Oncology develops tailored, real-time treatment recommendations based on this research and predicts disease prognosis, highlighting prospective therapeutic choices (127–129). Another deep learning-based Clinical Decision Support System (CDSS) was created by Printz et al. in 2017, and it is capable of extracting and analysing a significant quantity of clinical data from medical records and generating cancer therapy alternatives highlighting the value of AI technology in aiding physicians in enhancing cancer patient treatment regimens (130). A precision oncology knowledge repository called OncoKB is another example of an AI-based platform that uses clinical data and genomic information, providing details on the therapeutic ramifications of certain mutations. Identifying targeted treatments, forecasting cancer risk, and gauging therapy efficacy are all made easier with the help of OncoKB (131). Based on genetic variants and clinical parameters, AI models can forecast a particular individual's susceptibility to chemotherapy-related toxicity, helping oncologists make informed decisions about dosages and combined treatment plans to improve efficacy and tolerance (132). Dorman et al. developed a machine-learning algorithm to forecast how well breast cancer will respond to chemotherapy. It was able to distinguish between the effects of two chemotherapy medications, taxol or gemcitabine, by investigating the association between chemotherapy treatments and patients' genetic profiles (133). Also, an AI-based platform was developed that could accurately evaluate the effectiveness of immunotherapy in PD-1-sensitive patient with advanced solid tumours (134, 135). Utilizing the HLA mass-spectroscopy database, an AI model could help identify cancer neoantigens more accurately and boost the effectiveness of cancer immunotherapy (136). Machine learning algorithms have also been utilized to associate radiomic biomarkers (quantifiable features like texture patterns, structural changes, or intensity variations) within the biomedical images (CTs/MRIs) of tumour regions with treatment response and survival, thus, helping oncologists in anticipating favourable responses to therapies (137–140). AI algorithms can aid in radiation therapy planning by segmenting tumours and organs at risk from medical imaging scans. AI-based contouring algorithms have been created to automate the delineation of organs and tumour volumes in CT and MRI images, potentially assisting radiologists in mapping out treatment targets or autonomously planning radiation schedules (141, 142). Utilising deep-learning technology, the software can now generate automated radiation therapy plans within a matter of a few hours without human intervention (129).

### 2.3.2 AI-assisted surgery

Using imaging data, AI can help with preoperative planning by giving surgeons accurate 3D representations of the patient's anatomy. This enables surgeons to see the tumour and

surrounding tissues more, allowing for more accurate surgical planning. 3D images based on CT scans and MRIs have been utilized to predict the pathological grade of hepatocellular and brain carcinoma for better preoperative planning and reduced complications (143–147). AI can also help with intraoperative navigation by giving the surgeon feedback in real-time while the surgery is being done by overlaying 3D pictures on the surgical field using augmented reality (AR) or virtual reality (VR) technologies. As an illustration, a 2020 study that was published in *Nature Medicine* employed augmented reality to direct the excision of brain tumours based on real-time stimulated Raman histology images, increasing surgical accuracy (148). One of the newest advancements in minimally invasive surgery is robotic surgery. Artificial intelligence (AI) algorithms can be programmed into robotic surgery systems to identify and avoid harm to healthy tissues during surgery, lowering the chance of inaccuracies and collateral damages (149–151). Researchers from the Children's National Health System in Washington, D.C., developed the SMART system, which is intended to autonomously carry out soft tissue procedures, including cancer surgeries. It uses AI algorithms and image processing to direct the robotic arm in real-time while the procedure is being performed, increasing accuracy and minimizing human error (152). AI can also help with postoperative monitoring by evaluating multiple data feeds from sensors and other monitoring equipment such as blood pressure monitors or ECGs to identify early indicators of complications, enabling quick intervention and treatment (153).

### 2.3.3 Prognosis/survival

Cancer prognosis entails predicting disease recurrence, residual cancer burden, metastasis, patient's overall survival, and progression-free survival with the goal of improving patient care. Utilizing integrated multi-omics data (154–156) and other complex multi-dimensional datasets (157, 158), ML techniques have been proven to increase the precision of forecasts for cancer recurrence, and survival outcomes and aid doctors in the precise estimation of prognosis and treatment plan customization both pre and (159–162) post-intervention (163–165) scenarios, taking into account elements including tumour size, lymph node involvement, and hormone receptor status. A group in China has developed an algorithm that could analyse patient chemoresistance and prognosis and accurately predict the recurrence risk of breast cancer patients from EHR data (166). These approaches can assist doctors in developing customized treatment plans and follow-up care (167).

## 2.4 AI in telemedicine/remote monitoring

Post-COVID-19 pandemic, the healthcare sector has significantly changed and has sped up the global deployment of telemedicine tools and telehealth services. Telemedicine has the potential to bring about transformative changes in cancer care in India. The practice of providing cancer care remotely through video, telephone, and other electronic communication methods is

known as Teleoncology or telemedicine in oncology. Teleoncology applications include improved access to specialized care, cancer clinical trials, cancer telegenetics, telepathology, preoperative counselling, remote chemotherapy supervision, symptom management, follow-ups, survivorship care, and palliative care for people who might live in distant or underserved locations or have limited access to medical facilities support (168).

### 2.4.1 Apps/websites

AI-powered chatbots and virtual assistants can conduct preliminary assessments, diagnose symptoms, and offer basic medical advice. They can aid in classifying patients according to the seriousness of their conditions, allowing medical personnel to set priorities and distribute resources appropriately, thus facilitating effective healthcare delivery. One AI-based hospital pharmacy service was successfully tested in China during the COVID-19 pandemic in 2020 (169, 170). Hopido is an Indian Startup that has developed Cancer Dost, an AI-enabled chatbot to assist patients in connecting with specialized doctors and provides guidance for cancer treatment (171). As a tool for home-based treatment, mobile applications improve symptom management, lifestyle adjustment, and medication adherence. The oncologist can benefit from telemedicine's interactive tele-educational resources. AI-based algorithms for skin cancer diagnosis have recently been included in a number of mHealth apps, making this method available to the general public, allowing laypeople to use an AI-based mHealth app to assess their condition and determine if a doctor visit is necessary (172). An AI-powered app called Cancer Aid was created to help cancer patients along their journey. It gives access to a community of other cancer patients for support and direction, as well as individualized treatment information, symptom tracking, medication reminders, and access to symptom tracking (173). Using AI technology, CancerBASE offers individualized nutrition advice to cancer patients based on their individual medical and genetic profiles. To promote treatment outcomes and general well-being, it attempts to optimize nutrition (174).

### 2.4.2 Wearables/devices

With commercialization and the growing use of wearable technology, mobile health apps, and social media posts, AI can also make use of information on a variety of lifestyle and behavioural patterns, including risky behaviour, physical activity, sound sleep, alcohol consumption, smoking, and stress levels. Integrating these data into AI algorithms can revolutionize cancer care by improving cancer risk prediction and patient monitoring (175–177). Healthcare professionals may monitor the health of cancer patients without having to make regular in-person visits thanks to AI-powered remote monitoring technologies. For instance, wearable technology powered by AI algorithms can continuously track vital signs, spot anomalies, and notify medical staff when intervention is required. Innovative wearable technology makes it possible to track and analyse each patient's data differently based on their physical and genetic risk factors, resulting in a highly ultra-personalized diagnosis and course of treatment and better patient outcomes. This enables

faster treatment and reduces the burden on medical facilities. This facilitates prompt treatment and lightens the load on medical institutions (178, 179). Internationally, more than 70 AI-based devices have already been approved by the FDA with at least 50% of them being related to cancer radiology, clinical oncology, and pathology (180). One well-known example is the handheld device iBreastExam (iBE) (UE Life Sciences, Philadelphia, PA), which was created to do breast examinations for suspected anomalies or lumps in the breast tissue in a painless, non-invasive manner (181, 182). In a study done on women from the Nigerian population, this portable equipment revealed superior sensitivity (63%) for the identification of breast abnormalities as compared to clinical breast examination (CBE) (183). This radiation-free device can be particularly beneficial for early breast cancer detection campaigns in underserved communities or in regions with limited access to state-of-the-art healthcare facilities with only a little training. With increasing subscribers nationwide, surveillance using mobile phones integrated with machine learning programs has the potential to extend the reach of clinicians and vital diagnostic care to remote areas in India. A further possibility to enhance cancer screening may be the use of mobile devices to examine skin lesions. Cell phone-based triage or monitoring are prospective future methods for more efficiently deciding which patients to refer to clinicians or for expanding care in resource-limited areas, even though making this technology clinically feasible is still a work in progress.

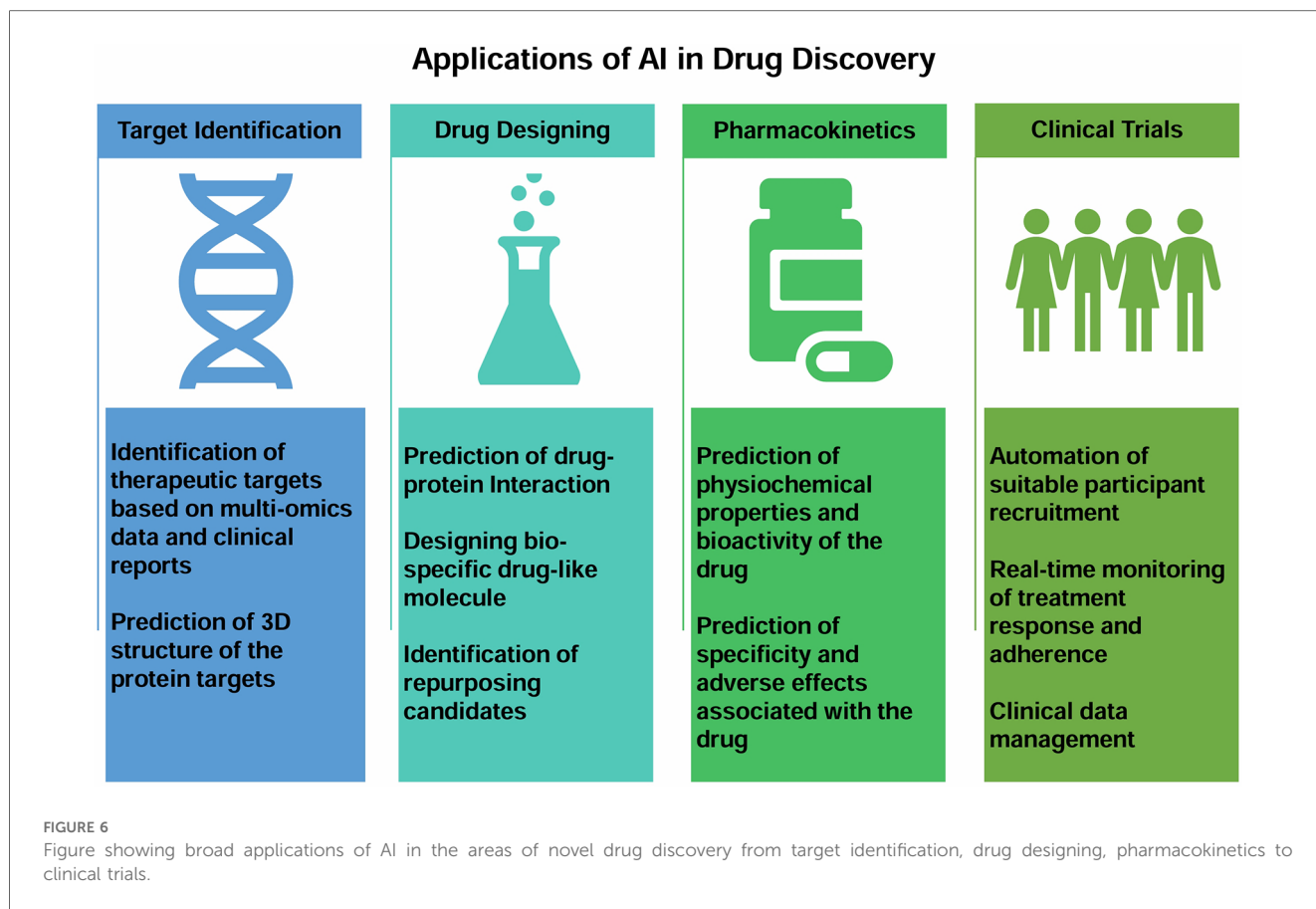
## 2.5 AI in novel discoveries

India has one of the most genetically diverse populations in the world, providing researchers with a unique chance to look into the molecular roots of diseases and create specialized treatment strategies. Additionally, the high prevalence of cancer enables researchers to investigate many subtypes, uncover the natural history and molecular basis of these diseases, and find possible targets for cutting-edge treatments. The growing availability of huge cancer datasets (public or private) and affordable access to various NGS and imaging technologies have sparked an explosion in interest in using AI to speed up discovering novel drugs/targets, which is frequently time and financially expensive. AI methods have been used in several areas of cancer research, including digging into molecular basis, creating anti-cancer drugs, and carrying out randomized controlled trials (RCTs) (Figure 6).

### 2.5.1 Drug targets/biomarkers

For example, a support vector machine model was used in a study to derive features that might predict possible therapeutic targets in liver cancer using clinical data, gene expression patterns, and protein-protein interaction networks (184). Another group developed a deep learning-based categorization strategy to identify proteins linked to the pathogenesis of breast cancer and report promising candidates for biomarkers or drug targets (185). By utilizing both gene-specific and cell-line-specific data, the ECLIPSE machine-learning technique could predict cancer-





specific therapeutic targets based on the publicly available loss-of-function screen datasets from the DepMap consortium (186, 187). Mkrtchyan et al. used the artificial intelligence-powered target discovery platform PandaOmics to investigate alterations in gene expression in rare DNA repair-deficient disorders and found new cancer targets (188). Utilizing the TCGA data, the ML approach established the role of F-box/WD repeat-containing protein 7 (Fbw7) in cancer cell oxidative metabolism (189).

### 2.5.2 Drug designing

AI has also been used to design drugs with desired physiochemical properties and target specificities. Using an AI-powered protein structure database called AlphaFold researchers from the University of Toronto and In-silico Medicine developed a drug that could treat hepatocellular carcinoma (HCC), or liver cancer (190). Instead of the typical length of about a year, a research group developed a DL model and identified potent inhibitors of the discoidin domain receptor 1 (DDR1), a kinase target implicated in several malignancies (191). Another study used the ML approach to improve the sensitivity and specificity of dual inhibitor cyclin-dependent kinases four and human epidermal growth factor receptor 2 (192). AI approaches are also deployed to predict adverse effects linked to drug toxicity and to design molecules with better biochemical properties in terms of

solubility, synthesizability and drug-likeness (193–195). Various AI-based computer tools have been developed that can help identify cancer-related drugs like DeepNeuralNetQSAR, DeepChem, DeepTox, gene2drug, STITCH, etc. (196). The identification of repurposing candidates from medications that can reverse the expression profiles of cancer-specific gene signatures has also been done using transcriptional data sets from LINCS (Library of Integrated Network-Based Cellular Signatures) libraries and other sources (197).

### 2.5.3 Clinical trial

Adoption of novel cancer therapies depends on the success of clinical trials, and finding the right subjects to enrol or recruit is seen to be the most challenging aspect of it. This requires labour-intensive work to match potential subjects with eligibility requirements satisfying inclusion and exclusion criteria. AI can improve cancer clinical trials by locating suitable participants for clinical trials and in the design of trials that are more effective and efficient (198). For instance, to automate the patient recruitment process, the combination of multi-layer perceptron modelling and natural language processing was used to extract pertinent data from patient records to help compile evidence for better patient eligibility decisions (199, 200). Analysis of the clinical trial data can assist in speeding up the discovery of new drugs and help patients receive new cancer treatments more

rapidly by identifying patients who are most likely to benefit from new therapies.

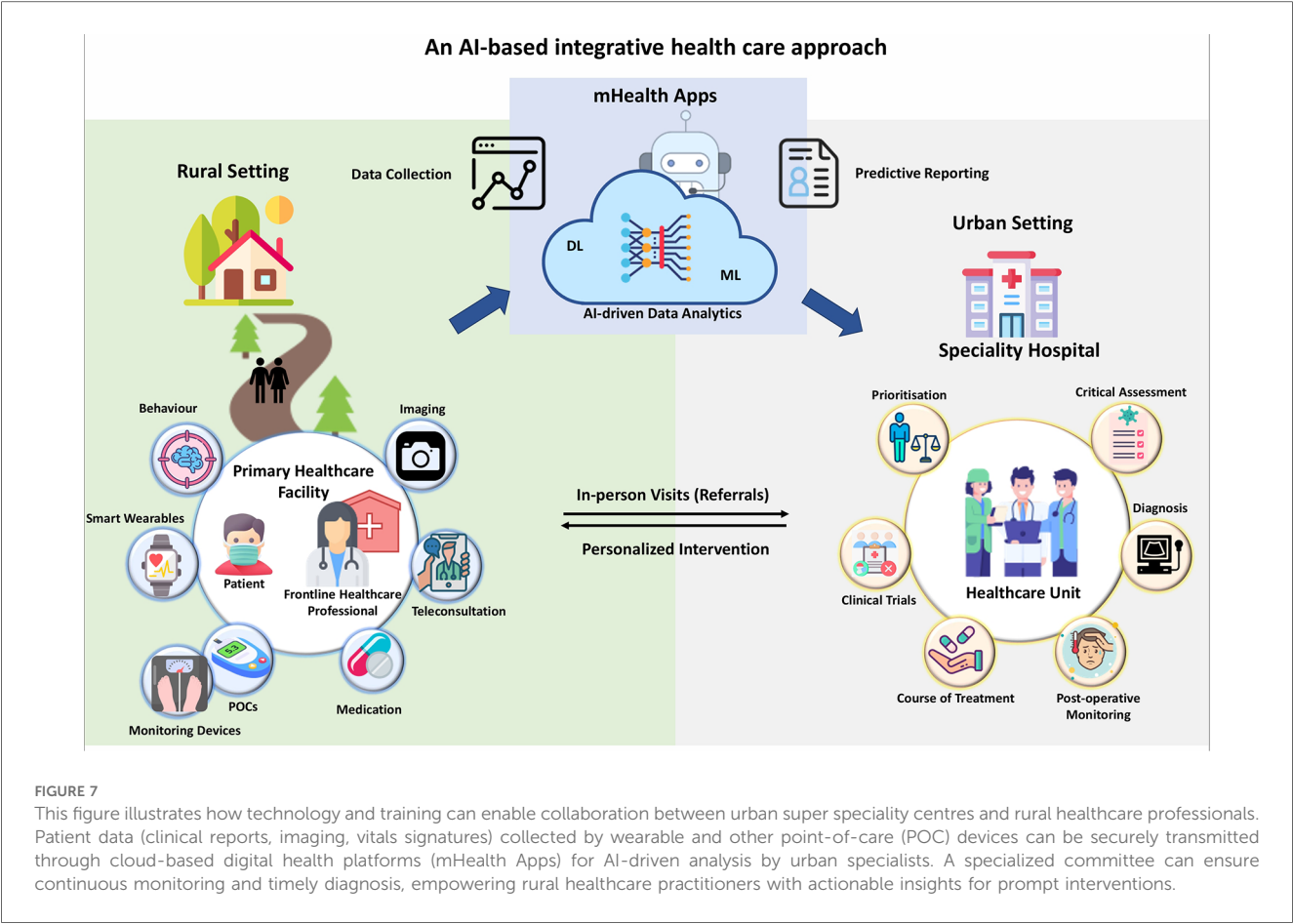
## 2.6 Advancements in AI

Artificial Intelligence (AI) is a rapidly advancing field, with new models being developed and tested daily to improve their effectiveness and accuracy. Innovations like DeepSeek-R1, OpenAI's GPT-4 and Google's Bard exemplify this progress, demonstrating remarkable advancements in understanding and generating human-like language. Among the most notable developments in AI are Multimodal Large Language Models (MLLMs) and Large Language Models (LLMs), which represent the cutting edge of artificial intelligence. These models combine deep learning and natural language processing (NLP) to transform healthcare. By utilizing sophisticated neural network architectures, LLMs and MLLMs can comprehend, generate, and analyse human language while processing diverse data types, such as medical images, audio recordings, and textual information. This capability allows for a comprehensive understanding of complex medical scenarios, enabling applications like disease diagnosis, medical report generation, treatment planning, and mental health support. Such advancements highlight the potential of AI to revolutionize healthcare delivery and improve patient outcomes (13). For instance, LLMs such as ChatDoctor and MLLMs such as Med-PaLM M specifically designed to tackle clinical challenges. These models are capable of processing both textual and visual medical data, significantly enhancing decision support systems in healthcare (201, 202). These advanced models excel in a variety of healthcare applications, such as virtual health assistance, automating administrative tasks, and supporting clinical decision-making. They enable a wide range of functionalities, including remote patient counselling, medical question-answering, dialogue summarization, electronic health record (EHR) generation, and clinical health reasoning. By streamlining healthcare operations and facilitating timely interventions, they play a crucial role in enhancing patient outcomes and creating more efficient healthcare systems (13). Despite their transformative potential, deploying Large Language Models (LLMs) and Multimodal Large Language Models (MLLMs) in clinical settings presents several challenges. Developing these models typically requires extensive annotated medical datasets, which are not only costly to acquire but also demand expert-level annotation and raise significant data privacy concerns. Furthermore, the high computational resources needed for training and deploying these models can limit accessibility, particularly for smaller healthcare institutions. Additionally, these models are prone to hallucinations—generating fabricated or inaccurate outputs—and often lack real-time updates, which can compromise their reliability and safety in fast-paced, dynamic clinical environments (203). Addressing these challenges is essential to unlocking the full potential of LLMs and MLLMs in transforming healthcare. Ongoing

research is focused on mitigating many of these obstacles. For instance, efforts are being made to create high-quality synthetic datasets, often generated using AI-assisted methods. These synthetic datasets reduce reliance on expensive, expert-annotated medical data, making the development of these models more scalable and cost-effective (201). Additionally, fine-tuning techniques such as reinforcement learning from human feedback (RLHF) and AI feedback (RLAIF) are being employed to align model outputs with clinical expectations and reduce hallucinations (204). Innovations in model optimization, such as parameter-efficient tuning and smaller, specialized models are reducing computational costs, benefiting smaller institutions (205). The future of LLMs and MLLMs lies in real-time learning, enhanced data privacy (using federated learning and encryption), and better integration of diverse data types for more comprehensive insights. Further reducing computational overhead through model compression and energy-efficient training will broaden access, making these technologies feasible for smaller healthcare providers. Crucially, collaboration among AI developers, medical professionals, and policymakers is vital to ensure the development of ethical, user-centred, and clinically effective AI solutions that drive personalized and efficient healthcare (13).

## 3 India's roadmap to digital health

Digital health harnesses advanced technologies such as telemedicine, mobile health applications (mHealth apps), electronic health records, and wearable devices to enhance healthcare accessibility, quality, and efficiency. These innovations facilitate remote patient monitoring, virtual consultations, and real-time data sharing, which are essential for overcoming geographical barriers and providing timely medical interventions. In regions with significant rural populations, access to specialized care is often hindered by inadequate infrastructure and a shortage of well-trained physicians. A digital health-based initiative can address this issue by integrating urban super-speciality centres with AI-based technology and training healthcare providers in resource-limited rural centres. This strategy establishes urban hubs for specialized care while empowering rural practitioners with advanced digital health systems. Telemedicine platforms can connect rural patients with urban specialists, ensuring expert care is available without necessitating travel. Mobile health apps support continuous health tracking and personalized health management, enabling patients to take proactive steps in their care. Central to this initiative is the utilization of wearables and point-of-care (POC) devices for remote collection of essential patient data. This data is securely transmitted to cloud-based digital health platforms (mHealth apps) with ultra-high-speed computing capabilities, enabling rapid AI-driven evaluation and analysis by experts at tertiary care urban centres. A dedicated action plan committee, operational around the clock, can monitor this data, conduct critical analyses, and relay prompt diagnoses and actionable



insights back to primary healthcare providers in rural centres (Figure 7). This streamlined process facilitates swift and effective personalized interventions, optimizing patient outcomes even in remote areas with limited access to specialized care. For actionable insights on enhancing AI adoption in rural-urban healthcare integration, please refer to Box 3.

## 4 Challenges of adopting AI in healthcare

While AI has the potential to revolutionize cancer research in India, it is not without its challenges and limitations. The adoption of AI in this field is hindered by issues such as regulations (206),

**BOX 3** Actionable insights for enhanced AI adoption in rural-urban healthcare integration.

These recommendations aim to bridge the gap between rural and urban healthcare, ensuring equitable access to advanced medical care through AI technologies.

<ol style="list-style-type: none"><li>1. Teaching and Training Programs:<ul style="list-style-type: none"><li>• Integrate AI in medical and technical education.</li><li>• Offer workshops and certification programs for healthcare providers.</li><li>• Encourage internships and practical training in AI-driven settings.</li></ul></li><li>2. Electronic Health Records (EHR) and Health Portals:<ul style="list-style-type: none"><li>• Mandate the use of EHRs for accurate patient data.</li><li>• Develop secure patient portals for easy access and management of health records.</li></ul></li><li>3. Telemedicine and Remote Monitoring:<ul style="list-style-type: none"><li>• Implement telemedicine for remote consultations with urban specialists.</li><li>• Use wearable devices for real-time health monitoring.</li><li>• Train healthcare providers and patients on digital health platforms (mHealth Apps).</li></ul></li></ol>	<ol style="list-style-type: none"><li>4. Policy Development and Implementation:<ul style="list-style-type: none"><li>• Establish protocols for AI adoption covering data security and ethics.</li><li>• Provide financial incentives and support for AI integration in rural healthcare.</li></ul></li><li>5. Mobile Health Units and Outreach Programs:<ul style="list-style-type: none"><li>• Deploy mobile health units with AI tools for regular check-ups in rural areas.</li><li>• Educate rural populations about the benefits of AI and telemedicine.</li></ul></li><li>6. Continuous Monitoring and Evaluation:<ul style="list-style-type: none"><li>• Establish metrics to evaluate AI interventions.</li><li>• Create feedback systems for continuous improvement.</li></ul></li></ol>
---	---

reliability of knowledge (207), difficulties in execution (208), its ability to adapt to diverse ecosystems (209), lack of transparency interpretability and data security.

## 4.1 Ethical/regulatory challenges

The rapid evolution of artificial intelligence (AI) technologies often outpaces the development of comprehensive regulatory frameworks, resulting in uncertainties surrounding approval processes, compliance standards, and ethical considerations, particularly in healthcare. The absence of clear regulations often leads to hesitation among healthcare providers and developers in fully embracing these innovations due to concerns about legal liabilities and ethical repercussions.

Globally, organizations such as the World Health Organization (WHO) have launched initiatives like the Global Strategy on Digital Health 2020–2025, which outlines recommendations for establishing AI governance frameworks to promote ethical AI adoption in healthcare (210). In India, however, the regulatory environment for AI in healthcare remains fragmented, hindering the integration of AI-powered tools in critical areas such as cancer diagnosis and treatment. To address these challenges, government initiatives like the National Digital Health Blueprint (NDHB) propose a comprehensive regulatory framework to ensure the ethical use of AI, focusing on data privacy, interoperability, and algorithmic accountability (211). Additionally, the Indian Council of Medical Research (ICMR) has issued ethical guidelines for AI applications in biomedical research and healthcare, emphasizing data security and patient privacy to bridge existing regulatory gaps (212).

Despite these efforts, significant work is still needed to establish robust policies governing AI deployment in healthcare while safeguarding patient rights and ethical practices. International collaborations, such as the AI for Health program by the International Telecommunication Union (ITU) and WHO, aim to support member countries in developing consistent global standards and regulatory frameworks for AI in healthcare (213, 214). These initiatives seek to harmonize international standards, enabling countries like India to adopt global best practices and strengthen their regulatory frameworks.

## 4.2 Data quality

One of the primary challenges in developing effective AI models for cancer care is the lack of standardized, diverse, and inclusive datasets related to cancer health. Training algorithms on inadequate, small, or outdated datasets can introduce biases and reduce the accuracy of outcomes in critical procedures such as cancer detection and therapy. In India, patient records often lack comprehensive details on genetic backgrounds, lifestyle factors, or previous treatments. This scarcity and inconsistency of data hinder the training of AI algorithms, leading to biased and potentially erroneous results. For example, algorithms trained on limited datasets—such as breast cancer biopsy images from a single

hospital—may fail to generalize effectively to larger, more diverse populations. Additionally, outdated datasets that do not reflect recent advancements in oncology can compromise the reliability of AI systems. Globally, initiatives like The Cancer Genome Atlas (TCGA), the International Cancer Genome Consortium (ICGC), the Global Alliance for Genomics and Health (GA4GH), and large-scale repositories such as the UK Biobank are addressing these gaps by creating comprehensive and diverse healthcare datasets. These efforts aim to foster interoperability and inclusivity in AI model development, ensuring that algorithms are trained on representative data (37, 215, 216).

India has also recognized the importance of addressing data diversity and inclusivity in healthcare AI. Initiatives like the Indian Cancer Genome Atlas, aim to compile comprehensive cancer-related multi-omics datasets across diverse populations (217). The integration of genomic data into such datasets is being facilitated by organizations like the Council of Scientific and Industrial Research (CSIR) and the Regional Centre for Biotechnology (RCB) through projects such as IndiGenomes and GenomeIndia (218, 219). These efforts, combined with collaborations between global tech companies and healthcare institutions, are critical in reducing biases and ensuring that AI models are inclusive, reliable, and representative of diverse populations. Such initiatives are essential for addressing the healthcare needs of India's heterogeneous population effectively.

## 4.3 Replicability

The replicability of AI findings across healthcare systems is a significant concern. Models trained in one environment often fail to perform consistently in others due to variations in data, infrastructure, and clinical practices. Clinical practice at academic medical centres and community hospitals still differs significantly today, but these differences are decreasing in high-income countries, like the United States, due to the integration of oncology practices by integrated health networks and new businesses that streamline oncology workflows and integrate genomic data, and translate these data into actionable reports (220, 221). In developing nations like India, the pace at which these changes occur depends on several aspects, such as disparities between urban hospitals and rural medical centres. Achieving operational consistency across different geographical locations requires not only protocol standardization but also workforce training in AI tools.

In India, the Ayushman Bharat Digital Mission (ABDM), launched in 2021, aims to create a unified digital health ecosystem (222). By integrating healthcare facilities and standardizing data collection practices, ABDM seeks to reduce discrepancies and improve the replicability of AI systems across the country. Additionally, collaborations between the National Cancer Grid (NCG) and international partners are helping establish uniform oncology practices and enhance the sensitivity and robustness of AI tools in diverse settings (223).

Globally, organizations like the European Commission are driving the development of standardized healthcare protocols

through initiatives such as the European Health Data Space, which aims to harmonize healthcare data across EU member states (224). Similarly, the MIT Critical Data Consortium is addressing replicability challenges by developing open-source healthcare datasets and fostering collaborations between academic and healthcare institutions (225). These efforts are crucial for ensuring the consistent performance of AI models across various healthcare settings and geographies.

## 4.4 Infrastructure

The integration of artificial intelligence (AI) into healthcare systems requires substantial investments in infrastructure, workforce training, and public awareness. In India, significant barriers such as limited computational resources and data storage solutions, and the availability of medical equipment and healthcare personnel make the widespread implementation of AI-driven solutions challenging. Initiatives such as the Digital India program and the establishment of Supercomputing Mission Centres aim to address these infrastructural gaps. Furthermore, the Indian Government's recent approval of the IndiaAI Mission, with a budgetary allocation exceeding Rs. 10,000 Crore, represents a significant step towards establishing a robust AI ecosystem in the country (226, 227). Additionally, developing cost-effective alternatives to imaging devices like mammograms, particularly for resource-limited settings, shows great promise in enhancing healthcare accessibility and efficiency. Ministry of Health and Family Welfare, India, has been implementing an action plan for Cancer screening program in rural areas under the Ayushman Bharat scheme. This initiative focuses on strengthening infrastructure, human resource development, health promotion, early diagnosis, management and referral of three common cancers i.e., oral, breast, and cervical (228).

In the private sector, Apollo Hospitals has established India's first AI-driven Precision Oncology Centre in Bengaluru. This centre integrates AI tools to provide personalized cancer care, improve diagnostic accuracy, and optimize treatment planning (121). Such initiatives highlight the growing adoption of AI-enabled healthcare solutions in India and signify a significant advancement in the country's healthcare landscape.

## 4.5 Transparency and interpretability

Additionally, there may be resistance from healthcare professionals due to a lack of familiarity with AI tools or fear of being replaced, leading to challenges in the adoption and utilization of AI in healthcare. A major concern is the lack of transparency and interpretability in AI systems, often referred to as the "black box" problem, where algorithms reach conclusions without providing clear reasoning. This opacity exacerbates trust issues among clinicians and patients, particularly in critical areas like oncology, where clear justifications for medical decisions are essential. Without understanding how an AI system arrives at a specific conclusion, it becomes difficult to trust and act upon its

recommendations. This lack of transparency raises ethical concerns and can hinder the clinical adoption of AI technologies. Furthermore, inherent biases in AI models toward specific demographic groups raise additional ethical and practical concerns, especially in a multicultural country like India.

To address these challenges, the Ministry of Electronics and Information Technology (MeitY) has launched initiatives such as the Responsible AI for Social Empowerment (RAISE) summit and the National AI Portal, which aim to promote AI awareness and ethical practices (229). Programs like the AI in Healthcare Certification Program by the All-India Institute of Medical Sciences (AIIMS) and collaborations between the Indian Institute of Technology (IIT) and global AI leaders like Microsoft are pivotal in training healthcare professionals to use AI tools effectively. Research institutions are also focusing on developing explainable AI (XAI) models to ensure that clinicians and patients can understand and trust AI-generated recommendations.

On an international scale, the Global Digital Health Partnership (GDHP), a collaborative effort involving multiple countries, including India, focuses on sharing best practices and resources for AI integration in healthcare (230). These efforts are crucial for building trust, addressing ethical concerns, and ensuring the responsible deployment of AI in healthcare systems.

## 4.6 Misdiagnosis and misuse

The accuracy and consistency of AI systems are heavily dependent on the quality of their training and validation datasets. Due to the inherently predictive and probabilistic nature of AI, these systems can sometimes make incorrect decisions, potentially leading to misdiagnosis, inappropriate treatment plans, or failure to identify critical health conditions, which can negatively impact patient outcomes. Incorporating larger and more diverse datasets in training can help mitigate biases and improve accuracy. Moreover, continuous advancements in AI technology are essential for reducing errors and enhancing predictive performance. While methods exist to detect bias during model training, their effectiveness still requires considerable improvement. Also, continuous validation and rigorous testing of AI systems in real-world scenarios are crucial to identify and rectify errors before deployment, ensuring improved accuracy and reliability.

Deep Learning (DL) methods, such as Convolutional Neural Networks (CNNs), have shown promise in medical applications, with validation rates of approximately 90% for cancer prediction using medical imagery. However, these techniques are often complex and computationally intensive. For instance, the Convolutional Neural Network (CNN) classifier, used in about 41% of experiments, has demonstrated good performance but demands significant computational resources (25, 231). Despite their potential, the practical application of these models in clinical settings remains inadequately validated. Predictions made by AI models often require verification in a clinical context to support medical experts in making diagnostic decisions. Currently, there are no comprehensive regulations or guidelines



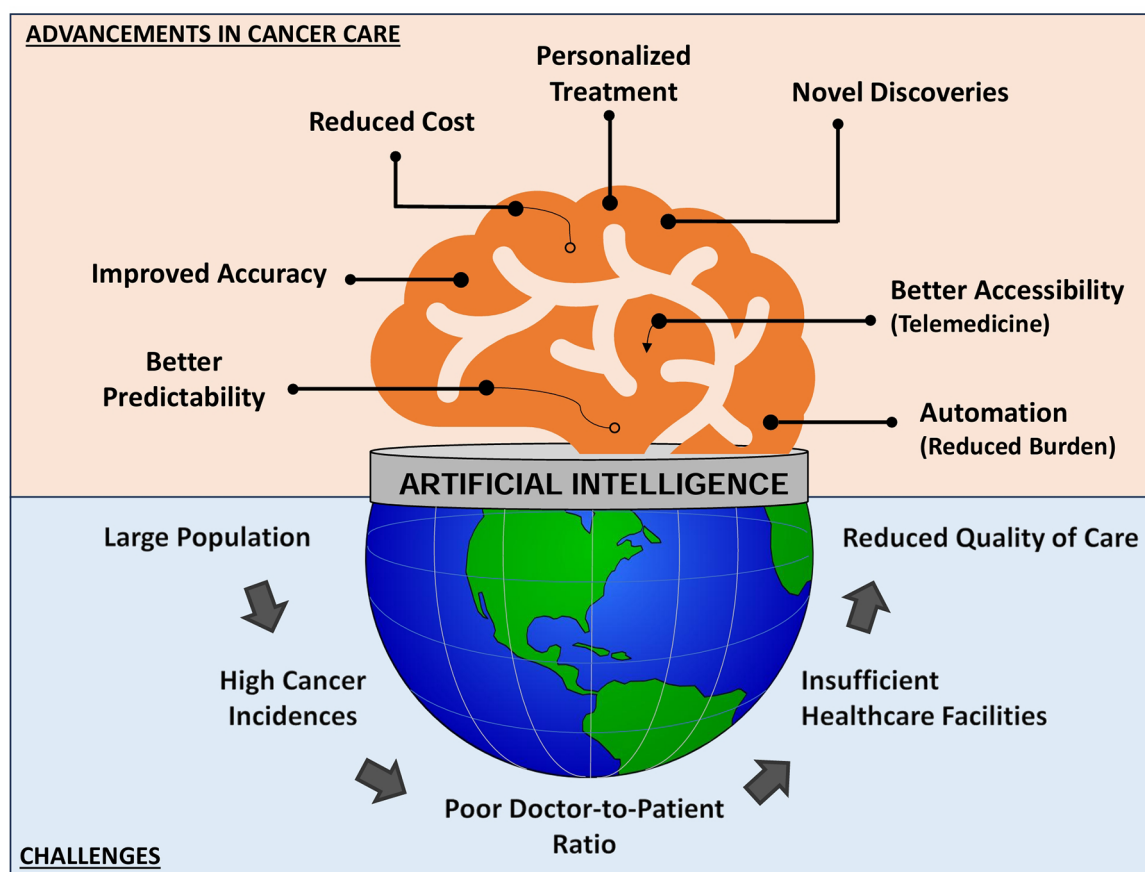


FIGURE 8

Pictorial representation of challenges encountered by populous countries and how advancements driven by artificial intelligence and digital health can enhance the management of cancer care.

to establish legal responsibility when AI systems cause harm or malfunction (232, 233). This underscores the urgent need for well-defined legal frameworks to ensure accountability and enhance patient safety. In the United States, the Food and Drug Administration (FDA) oversees the approval of AI-based medical devices to ensure their safety and efficacy (234–236). Similar regulatory frameworks are needed globally to address the ethical and legal challenges associated with AI in healthcare.

## 4.7 Data security

Addressing the ethical misuse of AI technologies, such as the creation of deepfake medical images or the manipulation of patient data, requires a comprehensive and multi-layered approach to data security and privacy. Strengthening the robustness of AI models through techniques like adversarial training can make them more resilient against manipulation and malicious attacks. Advanced data security measures, such as end-to-end encryption, multi-factor authentication, and secure access controls, are essential to protect sensitive patient information from unauthorized access, tampering, or breaches. Additionally, implementing regular audits, real-time system monitoring, and

anomaly detection mechanisms can help identify and mitigate potential ethical abuses swiftly and effectively.

In India, the proposed Digital Personal Data Protection Bill (2023) introduces stringent regulations aimed at ensuring the secure handling of patient data and preventing misuse (237). Additionally, initiatives led by the National Health Authority, such as the Ayushman Bharat Digital Mission (ABDM), focus on establishing robust accountability frameworks for AI-driven healthcare solutions (222, 238). These efforts emphasize transparency, ethical usage, and patient-centric safeguards. Together with global best practices, these measures provide a strong foundation for ensuring that AI technologies are deployed responsibly, fostering trust while minimizing the risks of ethical misuse.

## 5 Conclusion

Cancer incidence in India for the year 2022 was estimated at 14,61,427 cases with a rough rate of 100.4/100,000 (239) with about 1 in 9 people expected to develop cancer at some point in their lifetime. This poses an enormous burden for the country. Almost all public healthcare facilities and infrastructure are overwhelmed by this disproportionate number of cases and inadequate resources.

Healthcare facilities and healthcare providers (physicians paramedics and nurses) are sparse, particularly in rural and remote areas resulting in healthcare inequities and treatment disparities. Artificial Intelligence can revolutionize cancer care in developing populous nations by filling access gaps, improving healthcare delivery and boosting patient outcomes to better manage the difficulties and engage its large population (Figure 8). AI can leverage an overwhelmed healthcare and still can screen a broader population and intervene early by utilising predictive analytics technology. This review elaborates on the various domains of cancer management including cancer screening, diagnosis, precision treatment, prevention and surveillance using smarter algorithms and predictive tools.

In the future clinical setting, integrating computational input and assistance will become a tangible reality, leading to a significant technological revolution in real-time prediction and diagnosis of human health-related issues. Telemedicine will improve convenience, personalization, and access to specialized care, ultimately leading to better outcomes and a more patient-centred approach to cancer care. Android-based mobile apps and smarter wearables can perform uninterrupted health monitoring and flag any abnormalities at a very early stage. Artificial neural networks and deep learning will serve as optimal decision-making intelligence and evolve continuously, aiding physicians in rapid diagnosis decision-making and exploring treatment regimens.

However, it's essential to understand that AI in clinics only seeks to replace radiologists and other medical professionals partially. Instead, it will function as a novel and potential tool to achieve highly specific treatment performance and identify accurate diagnoses at the highest possible level while maintaining human involvement in final decision-making. As AI technology becomes more prevalent in healthcare, it is crucial to focus on fairness, inclusivity, data security, and adequate training of healthcare personnel. We are at the stage of known knowns (empowering AI in healthcare) but must be cautious of known unknowns (to what extent AI can be a reliable tool in disease diagnosis within its predictive limits) before we step into unknown unknowns (the risk and the ethical turmoil using highly innovative and smarter algorithms may bring in the future). It is best in the interest of humanity to explore the safer and judicious usage of AI for the betterment of society.

## Author contributions

IG: Conceptualization, Data curation, Investigation, Software, Validation, Visualization, Writing – original draft, Writing – review & editing, Formal analysis. YB: Investigation, Software, Visualization, Writing – original draft, Writing – review & editing, Data curation, Formal analysis. NK: Validation, Writing

– review & editing. SS: Writing – original draft, Writing – review & editing, Data curation, Visualization. MA: Validation, Funding acquisition, Writing – review & editing. RD: Conceptualization, Investigation, Supervision, Writing – original draft, Writing – review & editing, Validation, Resources. SK: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing, Investigation, Validation, Writing – original draft.

## Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article. IG acknowledges Indian council for Medical Research (ICMR) (Grant no. File No. 5/4-4/187/M/2020/-NCD-II) for providing fellowship.

## Acknowledgments

MA extends appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP2/138/45. The authors acknowledge All India Institute of Medical Sciences (AIIMS), New Delhi, India for providing infrastructure to carry this work.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. The World Bank. Population (total - India) (2023). Available online at: <https://data.worldbank.org/indicator/SP.POP.TOTL?locations=IN> (accessed May 10, 2024).
2. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide

for 36 cancers in 185 countries. *CA Cancer J Clin.* (2024) 74:229–63. doi: 10.3322/caac.21834

3. Karan A, Negandhi H, Hussain S, Zapata T, Mairambam D, De Graeve H, et al. Size, composition and distribution of health workforce in India: why,

and where to invest? *Hum Resour Health*. (2021) 19:39. doi: 10.1186/s12960-021-00575-2

4. The Global Health Observatory, WHO. Medical doctors (per 10000 population) (2023). Available online at: [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/medical-doctors-\(per-10-000-population\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/medical-doctors-(per-10-000-population)) (accessed May 10, 2024).

5. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. (2019) 380:1347–58. doi: 10.1056/nejmra1814259

6. Khurana D, Koli A, Khatter K, Singh S. Natural language processing: state of the art, current trends and challenges. *Multimed Tools Appl*. (2023) 82:3713–44. doi: 10.1007/s11042-022-13428-4

7. Zhou X, Gong W, Fu W, Du F. Application of deep learning in object detection. *Proceedings - 16th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2017* (2017).

8. Soori M, Arezoo B, Dastres R. Artificial intelligence, machine learning and deep learning in advanced robotics, a review. *Cogn Robot*. (2023) 3:54. doi: 10.1016/j.cogr.2023.04.001

9. Wiltshire TJ, Warta SF, Barber D, Fiore SM. Enabling robotic social intelligence by engineering human social-cognitive mechanisms. *Cogn Syst Res*. (2017) 43:190–207. doi: 10.1016/J.COGLYS.2016.09.005

10. Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. *Clin Infect Dis*. (2018) 66:149–53. doi: 10.1093/cid/cix731

11. Sen PC, Hajra M, Ghosh M. Supervised classification algorithms in machine learning: a survey and review. In: Mandal J, Bhattacharya D, editors. *Emerging Technology in Modelling and Graphics. Advances in Intelligent Systems and Computing*, Vol 937. Singapore: Springer (2020). doi: 10.1007/978-981-13-7403-6\_11

12. Min S, Lee B, Yoon S. Deep learning in bioinformatics. *Brief Bioinform*. (2017) 18:851–69. doi: 10.1093/BIB/BBW068

13. Xiao H, Zhou F, Liu X, Liu T, Li Z, Liu X, et al. A comprehensive survey of large language models and multimodal large language models in medicine. *Inf Fusion*. (2024) 117:102888. doi: 10.1016/j.inffus.2024.102888

14. Snousy MBA, El-Deeb HM, Badran K, Khlil IAA. Suite of decision tree-based classification algorithms on cancer gene expression data. *Egypt Inform J*. (2011) 12(2):73–82. doi: 10.1016/j.eij.2011.04.003

15. Wu Y, Zhu W, Wang J, Liu L, Zhang W, Wang Y, et al. Using machine learning for mortality prediction and risk stratification in atezolizumab? Treated cancer patients: integrative analysis of eight clinical trials. *Cancer Med*. (2022) 12(3):3744–57. doi: 10.1002/cam4.5060

16. Sahoo B, Pinnix Z, Sims S, Zelikovsky A. Identifying biomarkers using support vector machine to understand the racial disparity in triple-negative breast cancer. *J Comput Biol*. (2023) 30(4):502–17. doi: 10.1089/cmb.2022.0422

17. Huang C, Clayton EA, Matyunina LV, McDonald LD, Benigno BB, Vannberg F, et al. Machine learning predicts individual cancer patient responses to therapeutic drugs with high accuracy. *Sci Rep*. (2018) 8(1):16444. doi: 10.1038/s41598-018-34753-5

18. Florensa D, Mateo-Fornés J, Solsona F, Aige TP, Julió MM, Piñol R, et al. Use of multiple correspondence analysis and k-means to explore associations between risk factors and likelihood of colorectal cancer: cross-sectional study. *J Med Internet Res*. (2022) 24(7):e29056. doi: 10.2196/29056

19. Mehmood R, El-Ashram S, Bie R, Dawood H, Kos A. Clustering by fast search and merge of local density peaks for gene expression microarray data. *Sci Rep*. (2017) 7(1):45602. doi: 10.1038/srep45602

20. Ijaz MF, Attique M, Son Y. Data-driven cervical cancer prediction model with outlier detection and over-sampling methods. *Sensors*. (2020) 20(10):2809. doi: 10.3390/s20102809

21. Kabir MF, Chen T, Ludwig SA. A performance analysis of dimensionality reduction algorithms in machine learning models for cancer prediction. *Healthc Anal*. (2022) 3:100125. doi: 10.1016/j.health.2022.100125

22. Zhao Y, Kosorok MR, Zeng D. Reinforcement learning design for cancer clinical trials. *Stat Med*. (2009) 28(26):3294–315. doi: 10.1002/sim.3720

23. Tseng H, Luo Y, Cui S, Chien J, Haken RKT, Naqa IE. Deep reinforcement learning for automated radiation adaptation in lung cancer. *Med Phys*. (2017) 44(12):6690–705. doi: 10.1002/mp.12625

24. Liu M, Shen X, Pan W. Deep reinforcement learning for personalized treatment recommendation. *Stat Med*. (2022) 41(20):4034–56. doi: 10.1002/sim.9491

25. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. (2017) 42:60–88. doi: 10.1016/j.media.2017.07.005

26. Pham T, Tran T, Phung D, Venkatesh S. Predicting healthcare trajectories from medical records: a deep learning approach. *J Biomed Inform*. (2017) 69:218–29. doi: 10.1016/j.jbi.2017.04.001

27. Qin Y, Li C, Shi X, Wang W. MLP-based regression prediction model for compound bioactivity. *Front Bioeng Biotechnol*. (2022) 10:946329. doi: 10.3389/fbioe.2022.946329

28. Sado IT, Fitime LF, Pelap GF, Tinku C, Meudje GM, Bouetou TB. Early multi-cancer detection through deep learning: an anomaly detection approach using

variational autoencoder. *J Biomed Inform*. (2024) 160:104751. doi: 10.1016/j.jbi.2024.104751

29. Zhou S, Wang N, Wang L, Liu H, Zhang R. CancerBERT: a cancer domain-specific language model for extracting breast cancer phenotypes from electronic health records. *J Am Med Inform Assoc*. (2022) 29(7):1208–16. doi: 10.1093/jamia/ocac040

30. Cè M, Chiarpenello V, Bubba A, Felisaz PF, Oliva G, Irmici G, et al. Exploring the role of CHATGPT in oncology: providing information and support for cancer patients. *BioMedInformatics*. (2024) 4(2):877–88. doi: 10.3390/biomedinformatics4020049

31. Andrew A, Tizzard E. Large language models for improving cancer diagnosis and management in primary health care settings. *J Med Surg Public Health*. (2024) 4:100157. doi: 10.1016/j.glmedi.2024.100157

32. Wang M, Xie P, Du Y, Hu X. T5-Based model for abstractive summarization: a semi-supervised learning approach with consistency loss functions. *Appl Sci*. (2023) 13(12):7111. doi: 10.3390/app13127111

33. Zhang H, Yi H, Qin S, Liu X, Liu G. CLIP-based multimodal endorectal ultrasound enhances prediction of neoadjuvant chemoradiotherapy response in locally advanced rectal cancer. *PLoS One*. (2024) 19(12):e0315339. doi: 10.1371/journal.pone.0315339

34. Bazi Y, Rahhal MMA, Bashmal L, Zuair M. Vision-language model for visual question answering in medical imagery. *Bioengineering*. (2023) 10(3):380. doi: 10.3390/bioengineering10030380

35. Zhou HY, Yu Y, Wang C, Zhang S, Gao Y, Pan J, et al. A transformer-based representation-learning model with unified processing of multimodal input for clinical diagnostics. *Nat Biomed Eng*. (2023) 7(6):743–55. doi: 10.1038/s41551-023-01045-x

36. Rivero-Moreno Y, Echevarria S, Vidal-Valderrama C, Stefano-Pianetti L, Cordova-Guilarte J, Navarro-Gonzalez J, et al. Robotic surgery: a comprehensive review of the literature and current trends. *Cureus*. (2023) 15:e42370. doi: 10.7759/cureus.42370

37. Tomczak K, Czerwińska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol*. (2015) 19:A68–77. doi: 10.5114/wo.2014.47136

38. Saez-Rodriguez J, Costello JC, Friend SH, Kellen MR, Mangravite L, Meyer P, et al. Crowdsourcing biomedical research: leveraging communities as innovation engines. *Nat Rev Genet*. (2016) 17:470–86. doi: 10.1038/nrg.2016.69

39. Sage Bionetworks. Digital mammography DREAM challenge - sage bionetworks (2019). Available online at: <https://sagebionetworks.org/research-projects/digital-mammography-dream-challenge/> (accessed December 2, 2023).

40. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature*. (2020) 577:89–94. doi: 10.1038/s41586-019-1799-6

41. Wu Y, Burnside ES, Cox J, Fan J, Yuan M, Yin J, et al. Breast cancer risk prediction using electronic health records. *Proceedings - 2017 IEEE International Conference on Healthcare Informatics, ICHI 2017* (2017).

42. Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, et al. A deep learning approach to antibiotic discovery. *Cell*. (2020) 180:688–702.e13. doi: 10.1016/j.cell.2020.01.021

43. Artzi NS, Shilo S, Hadar E, Rossman H, Barbash-Hazan S, Ben-Haroush A, et al. Prediction of gestational diabetes based on nationwide electronic health records. *Nat Med*. (2020) 26:71–6. doi: 10.1038/s41591-019-0724-8

44. Zitnik M, Nguyen F, Wang B, Leskovec J, Goldenberg A, Hoffman MM. Machine learning for integrating data in biology and medicine: principles, practice, and opportunities. *Inf Fusion*. (2019) 50:71–91. doi: 10.1016/j.inffus.2018.09.012

45. 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 Digit Imaging*. (2013) 26:1045–57. doi: 10.1007/s10278-013-9622-7

46. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. (2012) 9(7):671–5. doi: 10.1038/nmeth.2089

47. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. Tensorflow: a system for large-scale machine learning. *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)* (2016). p. 265–83.

48. Chollet F. Deep learning with python (2017). Available online at: <http://cde.cern.ch/record/2301910> (accessed January 23, 2025).

49. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. *J Mach Learn Res*. (2011) 12:2825–30. doi: 10.5555/1953048.2078195

50. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *Computer vision and pattern recognition* (2014).

51. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016). p. 770–8

52. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: an imperative style, high-performance deep learning library. *Adv Neural Inf Process Syst*. (2019) 32:8026–37. doi: 10.48550/arXiv.1912.01703

53. Rodríguez-Ruiz A, Lång K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P, et al. Stand-alone artificial intelligence for breast cancer detection in mammography: comparison with 101 radiologists. *J Natl Cancer Inst.* (2019) 111:916–22. doi: 10.1093/JNCI/DJY222
54. Arasu VA, Habel LA, Achacoso NS, Buist DSM, Cord JB, Esserman LJ, et al. Comparison of mammography AI algorithms with a clinical risk model for 5-year breast cancer risk prediction: an observational study. *Radiology.* (2023) 307:e222733. doi: 10.1148/radiol.222733
55. Sasaki M, Tozaki M, Rodríguez-Ruiz A, Yotsumoto D, Ichiki Y, Terawaki A, et al. Artificial intelligence for breast cancer detection in mammography: experience of use of the ScreenPoint medical transpara system in 310 Japanese women. *Breast Cancer.* (2020) 27:642–51. doi: 10.1007/s12282-020-01061-8
56. Watanabe AT, Lim V, Vu HX, Chim R, Weise E, Liu J, et al. Improved cancer detection using artificial intelligence: a retrospective evaluation of missed cancers on mammography. *J Digit Imaging.* (2019) 32:625–37. doi: 10.1007/s10278-019-00192-5
57. Hu Q, Giger ML. Clinical artificial intelligence applications: breast imaging. *Radiol Clin.* (2021) 59:1027–43. doi: 10.1016/j.rcl.2021.07.010
58. Mahant SS, Varma AR. Artificial intelligence in breast the emerging future of modern medicine. *Cureus.* (2022) 14:e28945. doi: 10.7759/cureus.28945
59. Eriksson M, Czene K, Strand F, Zackrisson S, Lindholm P, Lång K, et al. Identification of women at high risk of breast cancer who need supplemental screening. *Radiology.* (2020) 297:327–33. doi: 10.1148/RADIOLOGY.2020201620
60. Yala A, Mikhael PG, Strand F, Lin G, Smith K, Wan YL, et al. Toward robust mammography-based models for breast cancer risk. *Sci Transl Med.* (2021) 13:eaba4373. doi: 10.1126/scitranslmed.aba4373
61. Lu MT, Raghu VK, Mayrhofer T, Aerts HJWL, Hoffmann U. Deep learning using chest radiographs to identify high-risk smokers for lung cancer screening computed tomography: development and validation of a prediction model. *Ann Intern Med.* (2020) 173:704–13. doi: 10.7326/M20-1868
62. Mavaddat N, Pharoah PDP, Michailidou K, Tyrer J, Brook MN, Bolla MK, et al. Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst.* (2015) 107:djv036. doi: 10.1093/jnci/djv036
63. 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. doi: 10.1038/nbt.4235
64. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet.* (2018) 50:1219–24. doi: 10.1038/s41588-018-0183-z
65. Mavaddat N, Michailidou K, Dennis J, Lush M, Fachal L, Lee A, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet.* (2019) 104:21–34. doi: 10.1016/j.ajhg.2018.11.002
66. Sipeky C, Talala KM, Tammela TLJ, Taari K, Auvinen A, Schleutker J. Prostate cancer risk prediction using a polygenic risk score. *Sci Rep.* (2020) 10:17075. doi: 10.1038/s41598-020-74172-z
67. Madakkat I, Lumsden AL, Mulugeta A, Olver I, Hyppönen E. Hypothesis? Free discovery of novel cancer predictors using machine learning. *Eur J Clin Investig.* (2023) 53:e14037. doi: 10.1111/eci.14037
68. Ning W, Wu T, Wu C, Wang S, Tao Z, Wang G, et al. Accurate prediction of pan-cancer types using machine learning with minimal number of DNA methylation sites. *J Mol Cell Biol.* (2023) 15:mjad023. doi: 10.1093/jmcb/mjad023
69. Zheng C, Xu R. Predicting cancer origins with a DNA methylation-based deep neural network model. *PLoS One.* (2020) 15:e0226461. doi: 10.1371/journal.pone.0226461
70. Vijaya Lakshmi TR, Krishna Reddy CV. Cancer prediction with gene expression profiling and differential evolution. *Signal Image Video Process.* (2023) 17:1855–61. doi: 10.1007/s11760-022-02396-9
71. Yuan L, Zhao J, Sun T, Shen Z. A machine learning framework that integrates multi-omics data predicts cancer-related LncRNAs. *BMC Bioinformatics.* (2021) 22:332. doi: 10.1186/s12859-021-04256-8
72. Zhang X, Wang J, Li J, Chen W, Liu C. CRlncRC: a machine learning-based method for cancer-related long noncoding RNA identification using integrated features. *BMC Med Genomics.* (2018) 11:120. doi: 10.1186/s12920-018-0436-9
73. Chandran U, Reps J, Yang R, Vachani A, Maldonado F, Kalsekar I. Machine learning and real-world data to predict lung cancer risk in routine care. *Cancer Epidemiol Biomarkers Prev.* (2023) 32:337–43. doi: 10.1158/1055-9965.EPI-22-0873
74. Yeh MCH, Wang YH, Yang HC, Bai KJ, Wang HH, Li YCJ. Artificial intelligence-based prediction of lung cancer risk using nonimaging electronic medical records: deep learning approach. *J Med Internet Res.* (2021) 23:e26256. doi: 10.2196/26256
75. Raghu VK, Walia AS, Zinzuwadia AN, Goiffon RJ, Shepard JAO, Aerts HJWL, et al. Validation of a deep learning-based model to predict lung cancer risk using chest radiographs and electronic medical record data. *JAMA Netw Open.* (2022) 5:E2248793. doi: 10.1001/jamanetworkopen.2022.48793
76. Poli UR, Gudlavalleti AG, Bharadwaj YJ, Pant HB, Agiwal V, Murthy GVS. Development and clinical validation of visual inspection with acetic acid application-artificial intelligence tool using cervical images in screen-and-treat visual screening for cervical cancer in south India: a pilot study. *JCO Glob Oncol.* (2024) 10:e2400146. doi: 10.1200/GO.24.00146
77. ATNF. Meghalaya cancer care project. (2023). Available online at: <https://atnf.org/meghalaya-cancer-care-project/> (accessed January 23, 2025).
78. Adapa K, Gupta A, Singh S, Kaur H, Trikha A, Sharma A, et al. A real world evaluation of an innovative artificial intelligence tool for population-level breast cancer screening. *NPJ Digit Med.* (2025) 8(1):2. doi: 10.1038/s41746-024-01368-2
79. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med.* (2019) 25:954–61. doi: 10.1038/s41591-019-0447-x
80. Schelb P, Kohl S, Radtke JP, Wiesenfarth M, Kickingereder P, Bickelhaupt S, et al. Classification of cancer at prostate MRI: deep learning versus clinical PI-RADS assessment. *Radiology.* (2019) 293:607–17. doi: 10.1148/radiol.2019190938
81. Varghese B, Chen F, Hwang D, Palmer SL, De Castro Abreu AL, Ukimura O, et al. Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images. *Sci Rep.* (2019) 9:1570. doi: 10.1038/s41598-018-38381-x
82. Rodríguez-Ruiz A, Krupinski E, Mordang JJ, Schilling K, Heywang-Köbrunner SH, Schepoulos I, et al. Detection of breast cancer with mammography: effect of an artificial intelligence support system. *Radiology.* (2019) 290:305–14. doi: 10.1148/radiol.2018181371
83. Zhou Q, Zhou Z, Chen C, Fan G, Chen G, Heng H, et al. Grading of hepatocellular carcinoma using 3D SE-DenseNet in dynamic enhanced MR images. *Comput Biol Med.* (2019) 107:47–57. doi: 10.1016/j.compbimed.2019.01.026
84. Wang P, Xiao X, Glissen Brown JR, Berzin TM, Tu M, Xiong F, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat Biomed Eng.* (2018) 2:741–8. doi: 10.1038/s41551-018-0301-3
85. Nagarajan KV, Bhat N. Imaging colonic polyps in 2024. *Indian J Gastroenterol.* (2024) 43(5):954–65. doi: 10.1007/s12664-024-01679-y
86. Esteve A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* (2017) 542:115–8. doi: 10.1038/nature21056
87. Pantanowitz L, Quiroga-Garza GM, Bien L, Heled R, Laifenfeld D, Linhart C, et al. An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study. *Lancet Digit Health.* (2020) 2:e407–16. doi: 10.1016/S2589-7500(20)30159-X
88. Yao H, Zhang X, Zhou X, Liu S. Parallel structure deep neural network using CNN and RNN with an attention mechanism for breast cancer histology image classification. *Cancers.* (2019) 11:1901. doi: 10.3390/cancers11121901
89. Cruz-Roa A, Gilmore H, Basavanthally A, Feldman M, Ganesan S, Shih NNC, et al. Accurate and reproducible invasive breast cancer detection in whole-slide images: a deep learning approach for quantifying tumor extent. *Sci Rep.* (2017) 7:46450. doi: 10.1038/srep46450
90. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med.* (2018) 24:1559–67. doi: 10.1038/s41591-018-0177-5
91. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep.* (2018) 23:181–93.e7. doi: 10.1016/j.celrep.2018.03.086
92. Son J, Shin JY, Kim HD, Jung KH, Park KH, Park SJ. Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. *Ophthalmology.* (2020) 127:85–94. doi: 10.1016/j.ophtha.2019.05.029
93. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature.* (2012) 486:346–52. doi: 10.1038/nature10983
94. Shen R, Olshen AB, Ladanyi M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics.* (2009) 25:2906–12. doi: 10.1093/bioinformatics/btp543
95. Chen R, Yang L, Goodison S, Sun Y. Deep-learning approach to identifying cancer subtypes using high-dimensional genomic data. *Bioinformatics.* (2020) 36:1476–83. doi: 10.1093/bioinformatics/btz769
96. Odhiambo P, Okello H, Wakaanya A, Wekesa C, Okoth P. Mutational signatures for breast cancer diagnosis using artificial intelligence. *J Egypt Natl Canc Inst.* (2023) 35:14. doi: 10.1186/s43046-023-00173-4
97. Gao F, Wang W, Tan M, Zhu L, Zhang Y, Fessler E, et al. DeepCC: a novel deep learning-based framework for cancer molecular subtype classification. *Oncogenesis.* (2019) 8:44. doi: 10.1038/s41389-019-0157-8
98. Hu F, Zhou Y, Wang Q, Yang Z, Shi Y, Chi Q. Gene expression classification of lung adenocarcinoma into molecular subtypes. *IEEE/ACM Trans Comput Biol Bioinform.* (2020) 17:1187–97. doi: 10.1109/TCBB.2019.2905553



99. Wang K, Duan X, Gao F, Wang W, Liu L, Wang X. Dissecting cancer heterogeneity based on dimension reduction of transcriptomic profiles using extreme learning machines. *PLoS One*. (2018) 13:e0203824. doi: 10.1371/journal.pone.0203824
100. Kormaksson M, Booth JG, Figueroa ME, Melnick A. Integrative model-based clustering of microarray methylation and expression data. *Ann Appl Stat*. (2012) 6:1327–47. doi: 10.1214/11-AOAS533
101. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, et al. DNA methylation-based classification of central nervous system tumours. *Nature*. (2018) 555:469–74. doi: 10.1038/NATURE26000
102. Wang A, Hai R, Rider PJ, He Q. Noncoding RNAs and deep learning neural network discriminate multi-cancer types. *Cancers*. (2022) 14:352. doi: 10.3390/cancers14020352
103. Mallavarapu T, Hao J, Kim Y, Oh JH, Kang M. Pathway-based deep clustering for molecular subtyping of cancer. *Methods*. (2020) 173:24–31. doi: 10.1016/j.ymeth.2019.06.017
104. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, et al. The cancer genome atlas pan-cancer analysis project. *Nat Genet*. (2013) 45:1113–20. doi: 10.1038/ng.2764
105. Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, McMichael JF, et al. Comprehensive molecular portraits of human breast tumours. *Nature*. (2012) 490:61–70. doi: 10.1038/nature11412
106. Kamoun A, Idbaih A, Dehais C, Elarouci N, Carpentier C, Letouze E, et al. Integrated multi-omics analysis of oligodendroglial tumours identifies three subgroups of 1p/19q co-deleted gliomas. *Nat Commun*. (2016) 7:11263. doi: 10.1038/ncomms11263
107. Franco EF, Rana P, Calderón VV, Azevedo V, Ramos RTJ, Ghosh P. Performance comparison of deep learning autoencoders for cancer subtype detection using multi-omics data. *Cancers*. (2021) 13:2013. doi: 10.3390/cancers13092013
108. Campbell PJ, Getz G, Korbel JO, Stuart JM, Jennings JL, Stein LD, et al. Pan-cancer analysis of whole genomes. *Nature*. (2020) 578:82–93. doi: 10.1038/s41586-020-1969-6
109. Jiao W, Atwal G, Polak P, Karlic R, Cuppen E, Al-Shahrour F, et al. A deep learning system accurately classifies primary and metastatic cancers using passenger mutation patterns. *Nat Commun*. (2020) 11:728. doi: 10.1038/s41467-019-13825-8
110. Menasalvas Ruiz E, Tuñas JM, Bermejo G, Gonzalo Martín C, Rodríguez-González A, Zanin M, et al. Profiling lung cancer patients using electronic health records. *J Med Syst*. (2018) 42:126. doi: 10.1007/s10916-018-0975-9
111. Moore CR, Farrag A, Ashkin E. Using natural language processing to extract abnormal results from cancer screening reports. *J Patient Saf*. (2017) 13:138–43. doi: 10.1097/PTS.0000000000000127
112. Glaser AP, Jordan BJ, Cohen J, Desai A, Silberman P, Meeks JJ. Automated extraction of grade, stage, and quality information from transurethral resection of bladder tumor pathology reports using natural language processing. *JCO Clin Cancer Inform*. (2018) 2:1–8. doi: 10.1200/cci.17.00128
113. Tao K, Bian Z, Zhang Q, Guo X, Yin C, Wang Y, et al. Machine learning-based genome-wide interrogation of somatic copy number aberrations in circulating tumor DNA for early detection of hepatocellular carcinoma. *EBioMedicine*. (2020) 56:102811. doi: 10.1016/j.ebiom.2020.102811
114. Cucchiara F, Del Re M, Valleggi S, Romei C, Petrini I, Lucchesi M, et al. Integrating liquid biopsy and radiomics to monitor clonal heterogeneity of EGFR-positive non-small cell lung cancer. *Front Oncol*. (2020) 10:593831. doi: 10.3389/fonc.2020.593831
115. Shen H, Liu T, Cui J, Borole P, Benjamin A, Kording K, et al. A web-based automated machine learning platform to analyze liquid biopsy data. *Lab Chip*. (2020) 20:2166–74. doi: 10.1039/d0lc00096e
116. Roth P, Wischhusen J, Happold C, Chandran PA, Hofer S, Eisele G, et al. A specific miRNA signature in the peripheral blood of glioblastoma patients. *J Neurochem*. (2011) 118:449–57. doi: 10.1111/j.1471-4159.2011.07307.x
117. Best MG, In 't Veld SGJG, Sol N, Wurdinger T. RNA sequencing and swarm intelligence-enhanced classification algorithm development for blood-based disease diagnostics using spliced blood platelet RNA. *Nat Protoc*. (2019) 14:1206–34. doi: 10.1038/s41596-019-0139-5
118. Shin H, Oh S, Hong S, Kang M, Kang D, Ji YG, et al. Early-stage lung cancer diagnosis by deep learning-based spectroscopic analysis of circulating exosomes. *ACS Nano*. (2020) 14:5435–44. doi: 10.1021/acsnano.9b09119
119. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science*. (2018) 359:926–30. doi: 10.1126/science.ear3247
120. The Hindu. AIIMS unveils indigenously developed technology for early detection of cancer (2024). Available online at: <https://www.thehindu.com/sci-tech/health/aiims-unveils-indigenously-developed-technology-for-early-detection-of-cancer/article67821512.ece> (accessed January 23, 2025).
121. Apollo Hospitals. Apollo cancer centres has successfully launched India's first AI-precision oncology centre (2024). Available online at: <https://www.apollohospitals.com/apollo-in-the-news/apollo-cancer-centres-has-successfully-launched-indias-first-ai-precision-oncology-centre/> (accessed January 23, 2025).
122. Costello JC, Heiser LM, Georgii E, Gönen M, Menden MP, Wang NJ, et al. A community effort to assess and improve drug sensitivity prediction algorithms. *Nat Biotechnol*. (2014) 32:1202–12. doi: 10.1038/nbt.2877
123. Saito M, Shiraishi K, Matsumoto K, Schetter AJ, Ogata-Kawata H, Tsuchiya N, et al. A three-microRNA signature predicts responses to platinum-based doublet chemotherapy in patients with lung adenocarcinoma. *Clin Cancer Res*. (2014) 20:4784–93. doi: 10.1158/1078-0432.CCR-14-1096
124. Gulhan DC, Lee JJK, Melloni GEM, Cortés-Ciriano I, Park PJ. Detecting the mutational signature of homologous recombination deficiency in clinical samples. *Nat Genet*. (2019) 51:912–9. doi: 10.1038/s41588-019-0390-2
125. Geeleher P, Zhang Z, Wang F, Gruener RF, Nath A, Morrison G, et al. Discovering novel pharmacogenomic biomarkers by imputing drug response in cancer patients from large genomics studies. *Genome Res*. (2017) 27:1743–51. doi: 10.1101/gr.221077.117
126. He X, Liu X, Zuo F, Shi H, Jing J. Artificial intelligence-based multi-omics analysis fuels cancer precision medicine. *Semin Cancer Biol*. (2023) 88:187–200. doi: 10.1016/j.semcancer.2022.12.009
127. Jie Z, Zhiying Z, Li L. A meta-analysis of watson for oncology in clinical application. *Sci Rep*. (2021) 11:5792. doi: 10.1038/s41598-021-84973-5
128. Yun HJ, Kim HJ, Kim SY, Lee YS, Lim CY, Chang HS, et al. Adequacy and effectiveness of watson for oncology in the treatment of thyroid carcinoma. *Front Endocrinol*. (2021) 12:585364. doi: 10.3389/FENDO.2021.585364
129. Babier A, Boutilier JJ, McNiven AL, Chan TCY. Knowledge-based automated planning for oropharyngeal cancer. *Med Phys*. (2018) 45:2875–83. doi: 10.1002/mp.12930
130. Printz C. Artificial intelligence platform for oncology could assist in treatment decisions. *Cancer*. (2017) 123:905. doi: 10.1002/cncr.30655
131. Chakravarty D, Gao J, Phillips S, Kundra R, Zhang H, Wang J, et al. OncoKB: a precision oncology knowledge base. *JCO Precis Oncol*. (2017) 1:1–16. doi: 10.1200/po.17.00011
132. Pantuck AJ, Lee DK, Kee T, Wang P, Lakhota S, Silverman MH, et al. Modulating BET bromodomain inhibitor ZEN-3694 and enzalutamide combination dosing in a metastatic prostate cancer patient using CURATE.AI, an artificial intelligence platform. *Adv Ther*. (2018) 1:1800104. doi: 10.1002/adtp.201800104
133. Dorman SN, Baranova K, Knoll JHM, Urquhart BL, Mariani G, Carcangiu ML, et al. Genomic signatures for paclitaxel and gemcitabine resistance in breast cancer derived by machine learning. *Mol Oncol*. (2016) 10:85–100. doi: 10.1016/j.molonc.2015.07.006
134. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol*. (2018) 19:1180–91. doi: 10.1016/S1470-2045(18)30413-3
135. Jabbari P, Rezaei N. Artificial intelligence and immunotherapy. *Expert Rev Clin Immunol*. (2019) 15:689–91. doi: 10.1080/1744666X.2019.1623670
136. Bulik-Sullivan B, Busby J, Palmer CD, Davis MJ, Murphy T, Clark A, et al. Deep learning using tumor HLA peptide mass spectrometry datasets improves neoantigen identification. *Nat Biotechnol*. (2019) 37:55–63. doi: 10.1038/nbt.4313
137. Trebeschi S, Drago SG, Birkbak NJ, Kurilova I, Ca'lin AM, Delli Pizzi A, et al. Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Ann Oncol*. (2019) 30:998–1004. doi: 10.1093/annonc/mdz108
138. Wang J, Zeng J, Li H, Yu X. A deep learning radiomics analysis for survival prediction in esophageal cancer. *J Healthc Eng*. (2022) 2022:4034404. doi: 10.1155/2022/4034404
139. Xie Y, Liu Q, Ji C, Sun Y, Zhang S, Hua M, et al. An artificial neural network-based radiomics model for predicting the radiotherapy response of advanced esophageal squamous cell carcinoma patients: a multicenter study. *Sci Rep*. (2023) 13:8673. doi: 10.1038/s41598-023-35556-z
140. Cha KH, Hadjiiski L, Chan HP, Weizer AZ, Alva A, Cohan RH, et al. Bladder cancer treatment response assessment in CT using radiomics with deep-learning. *Sci Rep*. (2017) 7:8738. doi: 10.1038/s41598-017-09315-w
141. Men K, Chen X, Zhang Y, Zhang T, Dai J, Yi J, et al. Deep deconvolutional neural network for target segmentation of nasopharyngeal cancer in planning computed tomography images. *Front Oncol*. (2017) 7:315. doi: 10.3389/fonc.2017.00315
142. Lin L, Dou Q, Jin YM, Zhou GQ, Tang YQ, Chen WL, et al. Deep learning for automated contouring of primary tumor volumes by MRI for nasopharyngeal carcinoma. *Radiology*. (2019) 291:677–86. doi: 10.1148/radiol.2019182012
143. Nie P, Yang G, Guo J, Chen J, Li X, Ji Q, et al. A CT-based radiomics nomogram for differentiation of focal nodular hyperplasia from hepatocellular carcinoma in the non-cirrhotic liver. *Cancer Imaging*. (2020) 20:20. doi: 10.1186/s40644-020-00297-z



144. Ponnoprat D, Inkeaw P, Chaijaruwanich J, Traisathit P, Sripan P, Inmutto N, et al. Classification of hepatocellular carcinoma and intrahepatic cholangiocarcinoma based on multi-phase CT scans. *Med Biol Eng Comput.* (2020) 58:2497–515. doi: 10.1007/s11517-020-02229-2
145. Liu X, Khalvati F, Namdar K, Fischer S, Lewis S, Taouli B, et al. Can machine learning radiomics provide pre-operative differentiation of combined hepatocellular cholangiocarcinoma from hepatocellular carcinoma and cholangiocarcinoma to inform optimal treatment planning? *Eur Radiol.* (2021) 31:244–55. doi: 10.1007/s00330-020-07119-7
146. Mao B, Zhang L, Ning P, Ding F, Wu F, Lu G, et al. Preoperative prediction for pathological grade of hepatocellular carcinoma via machine learning-based radiomics. *Eur Radiol.* (2020) 30:6924–32. doi: 10.1007/s00330-020-07056-5
147. Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, et al. Brain tumor segmentation with deep neural networks. *Med Image Anal.* (2017) 35:18–31. doi: 10.1016/j.media.2016.05.004
148. Hollon TC, Pandian B, Adapa AR, Urias E, Save AV, Khalsa SSS, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nat Med.* (2020) 26:52–8. doi: 10.1038/s41591-019-0715-9
149. Panesar S, Cagle Y, Chander D, Morey J, Fernandez-Miranda J, Kliot M. Artificial intelligence and the future of surgical robotics. *Ann Surg.* (2019) 270:223–6. doi: 10.1097/SLA.0000000000003262
150. Veronesi G, Novellis P, Voulaz E, Alloisio M. Robot-assisted surgery for lung cancer: state of the art and perspectives. *Lung Cancer.* (2016) 101:28–34. doi: 10.1016/j.lungcan.2016.09.004
151. Zhou XY, Guo Y, Shen M, Yang GZ. Application of artificial intelligence in surgery. *Front Med.* (2020) 14:417–30. doi: 10.1007/s11684-020-0770-0
152. Shademan A, Decker RS, Opfermann JD, Leonard S, Krieger A, Kim PCW. Supervised autonomous robotic soft tissue surgery. *Sci Transl Med.* (2016) 8:337ra64. doi: 10.1126/scitranslmed.aad9398
153. Kaieski N, da Costa CA, da Rosa Righi R, Lora PS, Eskofier B. Application of artificial intelligence methods in vital signs analysis of hospitalized patients: a systematic literature review. *Appl Soft Comput J.* (2020) 96:6497. doi: 10.1016/j.asoc.2020.106612
154. Chai H, Zhou X, Zhang Z, Rao J, Zhao H, Yang Y. Integrating multi-omics data through deep learning for accurate cancer prognosis prediction. *Comput Biol Med.* (2021) 134:104481. doi: 10.1016/j.combiomed.2021.104481
155. Ching T, Zhu X, Garmire LX. Cox-nnet: an artificial neural network method for prognosis prediction of high-throughput omics data. *PLoS Comput Biol.* (2018) 14: e1006076. doi: 10.1371/journal.pcbi.1006076
156. Zhang S, Xu Y, Hui X, Yang F, Hu Y, Shao J, et al. Improvement in prediction of prostate cancer prognosis with somatic mutational signatures. *J Cancer.* (2017) 8:3261–7. doi: 10.7150/jca.21261
157. Sun D, Wang M, Li A. A multimodal deep neural network for human breast cancer prognosis prediction by integrating multi-dimensional data. *IEEE/ACM Trans Comput Biol Bioinform.* (2019) 16(3):841–50. doi: 10.1109/TCBB.2018.2806438
158. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. *Cancer Inform.* (2006) 2:59–77. doi: 10.1177/117693510600200030
159. Sun Y, Goodison S, Li J, Liu L, Farmerie W. Improved breast cancer prognosis through the combination of clinical and genetic markers. *Bioinformatics.* (2007) 23:30–7. doi: 10.1093/bioinformatics/btl543
160. Chaudhary K, Poirion OB, Lu L, Garmire LX. Deep learning-based multi-omics integration robustly predicts survival in liver cancer. *Clin Cancer Res.* (2018) 24:1248–59. doi: 10.1158/1078-0432.CCR-17-0853
161. Karami G, Orlando MG, Pizzi AD, Caulo M, Del Gratta C. Predicting overall survival time in glioblastoma patients using gradient boosting machines algorithm and recursive feature elimination technique. *Cancers.* (2021) 13:4976. doi: 10.3390/cancers13194976
162. Gao HX, Huang SG, Du JF, Zhang XC, Jiang N, Kang WX, et al. Comparison of prognostic indices in NSCLC patients with brain metastases after radiosurgery. *Int J Biol Sci.* (2018) 14:2065–72. doi: 10.7150/ijbs.28608
163. Cui Y, Li Z, Xiang M, Han D, Yin Y, Ma C. Machine learning models predict overall survival and progression free survival of non-surgical esophageal cancer patients with chemoradiotherapy based on CT image radiomics signatures. *Radiat Oncol.* (2022) 17:212. doi: 10.1186/s13014-022-02186-0
164. Kim JY, Oh JM, Lee SK, Yu J, Lee JE, Kim SW, et al. Improved prediction of survival outcomes using residual cancer burden in combination with ki-67 in breast cancer patients underwent neoadjuvant chemotherapy. *Front Oncol.* (2022) 12:903372. doi: 10.3389/fonc.2022.903372
165. Dammu H, Ren T, Duong TQ. Deep learning prediction of pathological complete response, residual cancer burden, and progression-free survival in breast cancer patients. *PLoS One.* (2023) 18:e0280148. doi: 10.1371/journal.pone.0280148
166. Han Y, Zhu X, Hu Y, Yu C, Guo Y, Hang D, et al. Electronic health record-based absolute risk prediction model for esophageal cancer in the Chinese population: model development and external validation. *JMIR Public Health Surveill.* (2023) 9:e43725. doi: 10.2196/43725
167. Zeng L, Liu L, Chen D, Lu H, Xue Y, Bi H, et al. The innovative model based on artificial intelligence algorithms to predict recurrence risk of patients with postoperative breast cancer. *Front Oncol.* (2023) 13:1117420. doi: 10.3389/fonc.2023.1117420
168. Pan LC, Wu XR, Lu Y, Zhang HQ, Zhou YL, Liu X, et al. Artificial intelligence empowered digital health technologies in cancer survivorship care: a scoping review. *Asia Pac J Oncol Nurs.* (2022) 9:100127. doi: 10.1016/j.apjon.2022.100127
169. Bu F, Sun H, Li L, Tang F, Zhang X, Yan J, et al. Artificial intelligence-based internet hospital pharmacy services in China: perspective based on a case study. *Front Pharmacol.* (2022) 13:1027808. doi: 10.3389/fphar.2022.1027808
170. Li R, Yang Y, Wu S, Huang K, Chen W, Liu Y, et al. Using artificial intelligence to improve medical services in China. *Ann Transl Med.* (2020) 8:711. doi: 10.21037/atm.2019.11.108
171. New Indian Express. Cancer care at your doorstep with Hospido's new AI-enabled chatbot cancer dost (2021). Available online at: <https://www.newindianexpress.com/cities/delhi/2021/jan/21/cancer-care-at-your-doorstep-withhospidos-new-ai-enabled-chatbot-cancer-dost-2253007.html> (accessed May 10, 2024).
172. Smak Gregoor AM, Sangers TE, Bakker LJ, Hollestein L, Uyl-de Groot CA, Nijsten T, et al. An artificial intelligence based app for skin cancer detection evaluated in a population based setting. *NPJ Digit Med.* (2023) 6:90. doi: 10.1038/s41746-023-00831-w
173. Osara Health. Improved cancer care and support (2023). Available online at: <https://osarahealth.com/> (accessed May 10, 2024).
174. CancerBase. CancerBase (2023). Available online at: <https://www.cancerbase.org/> (accessed May 10, 2024).
175. Menta AK, Subbiah IM, Subbiah V. Bringing wearable devices into oncology practice: fitting smart technology in the clinic. *Discov Med.* (2018) 26:261–70.
176. Ge H, Li L, Zhang D, Ma F. Applications of digital medicine in oncology: prospects and challenges. *Cancer Innov.* (2022) 1:285–92. doi: 10.1002/cai2.37
177. Low CA. Harnessing consumer smartphone and wearable sensors for clinical cancer research. *NPJ Digit Med.* (2020) 3:140. doi: 10.1038/s41746-020-00351-x
178. Beg S, Handa M, Shukla R, Rahman M, Almalki WH, Afzal O, et al. Wearable smart devices in cancer diagnosis and remote clinical trial monitoring: transforming the healthcare applications. *Drug Discov Today.* (2022) 27:103314. doi: 10.1016/j.drudis.2022.06.014
179. Shaik T, Tao X, Higgins N, Li L, Gururajan R, Zhou X, et al. Remote patient monitoring using artificial intelligence: current state, applications, and challenges. *Wiley Interdiscip Rev Data Min Knowl Discov.* (2023) 13:e1485. doi: 10.1002/widm.1485
180. Luchini C, Pea A, Scarpa A. Artificial intelligence in oncology: current applications and future perspectives. *Br J Cancer.* (2022) 126:4–9. doi: 10.1038/s41416-021-01633-1
181. Broach RB, Geha R, Englander BS, DeLaCruz L, Thrash H, Brooks AD. A cost-effective handheld breast scanner for use in low-resource environments: a validation study. *World J Surg Oncol.* (2016) 14:277. doi: 10.1186/s12957-016-1022-2
182. iBreastExam. Breast exam for better breast health (2023). Available online at: <http://www.ibreastexam.com> (accessed May 10, 2024).
183. Mango VL, Olasehinde O, Omisore AD, Wuraola FO, Famurewa OC, Sevilimedu V, et al. The iBreastExam versus clinical breast examination for breast evaluation in high risk and symptomatic Nigerian women: a prospective study. *Lancet Glob Health.* (2022) 10:e555–63. doi: 10.1016/S2214-109X(22)00030-4
184. Tong Z, Zhou Y, Wang J. Identifying potential drug targets in hepatocellular carcinoma based on network analysis and one-class support vector machine. *Sci Rep.* (2019) 9:10442. doi: 10.1038/s41598-019-46540-x
185. López-Cortés A, Cabrera-Andrade A, Vázquez-Naya JM, Pazos A, González-Díaz H, Paz-y-Miño C, et al. Prediction of breast cancer proteins involved in immunotherapy, metastasis, and RNA-binding using molecular descriptors and artificial neural networks. *Sci Rep.* (2020) 10:8515. doi: 10.1038/s41598-020-65584-y
186. Gilvary C, Madhukar N, Gayvert K, Foronda M, Perez A, Leslie C, et al. A machine learning approach predicts essential genes and pharmacological targets in cancer. *Icahn School of Medicine at Mount (2019):692277.*
187. Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, et al. Defining a cancer dependency map. *Cell.* (2017) 170:564–76.e16. doi: 10.1016/j.cell.2017.06.010
188. Mkrtchyan GV, Veviorskiy A, Izumchenko E, Shneyderman A, Pun FW, Ozerov IV, et al. High-confidence cancer patient stratification through multiomics investigation of DNA repair disorders. *Cell Death Dis.* (2022) 13:999. doi: 10.1038/s41419-022-05437-w
189. Davis RJ, Gönen M, Margineantu DH, Handeli S, Swanger J, Hoellerbauer P, et al. Pan-cancer transcriptional signatures predictive of oncogenic mutations reveal that Fbw7 regulates cancer cell oxidative metabolism. *Proc Natl Acad Sci U S A.* (2018) 115:5462–7. doi: 10.1073/pnas.1718338115
190. 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. doi: 10.1039/d2sc05709c

191. Zhavoronkov A, Ivanenkov YA, Aliper A, Veselov MS, Aladinskiy VA, Aladinskaya AV, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat Biotechnol.* (2019) 37:1038–40. doi: 10.1038/s41587-019-0224-x

192. Kleandrova VV, Scotti MT, Scotti L, Speck-Planche A. Multi-target drug discovery via PTML modeling: applications to the design of virtual dual inhibitors of CDK4 and HER2. *Curr Top Med Chem.* (2021) 21:661–75. doi: 10.2174/1568026621666210119112845

193. Gayvert KM, Madhukar NS, Elemento O. A data-driven approach to predicting successes and failures of clinical trials. *Cell Chem Biol.* (2016) 23:1294–301. doi: 10.1016/j.chembiol.2016.07.023

194. Martins IF, Teixeira AL, Pinheiro L, Falcao AO. A Bayesian approach to in silico blood-brain barrier penetration modeling. *J Chem Inf Model.* (2012) 52:1686–97. doi: 10.1021/ci300124c

195. Shen J, Cheng F, Xu Y, Li W, Tang Y. Estimation of ADME properties with substructure pattern recognition. *J Chem Inf Model.* (2010) 50:1034–41. doi: 10.1021/ci100104j

196. Shao D, Dai Y, Li N, Cao X, Zhao W, Cheng L, et al. Artificial intelligence in clinical research of cancers. *Brief Bioinform.* (2022) 23:bbab523. doi: 10.1093/bib/bbab523

197. Mokou M, Lygirou V, Angeloudaki I, Paschalidis N, Stroggilos R, Frantzi M, et al. A novel pipeline for drug repurposing for bladder cancer based on patients' omics signatures. *Cancers.* (2020) 12:1–16. doi: 10.3390/cancers12123519

198. Askin S, Burkhalter D, Calado G, El Dakrouni S. Artificial intelligence applied to clinical trials: opportunities and challenges. *Health Technol.* (2023) 13:203–13. doi: 10.1007/s12553-023-00738-2

199. Hassanzadeh H, Karimi S, Nguyen A. Matching patients to clinical trials using semantically enriched document representation. *J Biomed Inform.* (2020) 105:103406. doi: 10.1016/j.jbi.2020.103406

200. Gligorijevic J, Gligorijevic D, Pavlovski M, Milkovits E, Glass L, Grier K, et al. Optimizing clinical trials recruitment via deep learning. *J Am Med Inform Assoc.* (2019) 26:1195–202. doi: 10.1093/jamia/ocz064

201. Li Y, Li Z, Zhang K, Dan R, Jiang S, Zhang Y. Chatdoctor: a medical chat model fine-tuned on a large language model meta-AI (LLaMA) using medical domain knowledge. *Cureus.* (2023) 15(6):e40895. doi: 10.7759/cureus.40895

202. Tu T, Azizi S, Driess D, Schackermann M, Amin M, Chang P, et al. Towards generalist biomedical AI. *NEJM AI.* (2024) 1(3). doi: 10.1056/aioa2300138

203. Ji Z, Lee N, Frieske R, Yu T, Su D, Xu Y, et al. Survey of hallucination in natural language generation. *ACM Comput Surv.* (2022) 55(12):1–38. doi: 10.1145/3571730

204. Lee H, Phatale S, Mansoor H, Mesnard T, Ferret J, Lu K, et al. RLAIIF vs RLHF: scaling reinforcement learning from Human feedback with AI feedback. *arXiv.org* (2023).

205. Razuvayevskaya O, Wu B, Leite JA, Heppell F, Srba I, Scarton C, et al. Comparison between parameter-efficient techniques and full fine-tuning: a case study on multilingual news article classification. *PLoS One.* (2024) 19(5):e0301738. doi: 10.1371/journal.pone.0301738

206. Tobin J. Artificial intelligence: development, risks and regulation. House of lords library, UK parliament (2023). Available online at: <https://lordslibrary.parliament.uk/artificial-intelligence-development-risks-and-regulation/> (accessed May 10, 2024).

207. Jarrahi MH, Askay D, Eshraghi A, Smith P. Artificial intelligence and knowledge management: a partnership between human and AI. *Bus Horiz.* (2023) 66(1): 87–99. doi: 10.1016/j.bushor.2022.03.002

208. Khan B, Fatima H, Qureshi A, Kumar S, Hanan A, Hussain J, et al. Drawbacks of artificial intelligence and their potential solutions in the healthcare sector. *Biomed Mater Devices.* (2023) 1:731–8. doi: 10.1007/s44174-023-00063-2

209. Endsley MR. Ironies of artificial intelligence. *Ergonomics.* (2023) 66(11):1656–68. doi: 10.1080/00140139.2023.2243404

210. World Health Organisation, Digital Health and Innovation. *Global Strategy on Digital Health 2020–2025*. Geneva: World Health Organization (2021).

211. Ministry of Health and Family Welfare. National digital health blueprint (2019). Available online at: [https://abdm.gov.in:8081/uploads/ndhb\\_1\\_56ec695bc8.pdf](https://abdm.gov.in:8081/uploads/ndhb_1_56ec695bc8.pdf) (accessed January 23, 2025).

212. Indian Council for Medical Research (ICMR). Ethical guidelines for application of artificial intelligence in biomedical research and healthcare (2023). Available online at: <https://www.icmr.gov.in/ethical-guidelines-for-application-of-artificial-intelligence-in-biomedical-research-and-healthcare> (Accessed May 10, 2024).

213. International Telecommunication Union. Artificial intelligence for health (2023). Available online at: <https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/default.aspx> (accessed January 23, 2025).

214. Reddy S. Global harmonisation of AI-enabled software as a medical device regulation: addressing challenges and unifying standards. *Mayo Clin Proc Digit Health.* (2024) 3(1):100191. doi: 10.1016/j.mcpdig.2024.100191

215. Zhang J, Bajari R, Andric D, Gerthoffert F, Lepsa A, Nahal-Bose H, et al. The international cancer genome consortium data portal. *Nat Biotechnol.* (2019) 37:367–9. doi: 10.1038/s41587-019-0055-9

216. Rehm HL, Page AJH, Smith L, Adams JB, Alterovitz G, Babb LJ, et al. GA4GH: international policies and standards for data sharing across genomic research and healthcare. *Cell Genom.* (2021) 1(2):100029. doi: 10.1016/j.xgen.2021.100029

217. Dhup S, Ramamurthy R. A genome atlas mapping cancers in India. *Nat India.* (2024). doi: 10.1038/d44151-024-00019-5

218. Jain A, Bhojar RC, Pandhare K, Mishra A, Sharma D, Imran M, et al. Indigenomes: a comprehensive resource of genetic variants from over 1000 Indian genomes. *Nucleic Acids Res.* (2021) 49(D1):D1225–32. doi: 10.1093/nar/gkaa923

219. GenomeIndia. (2024). Available online at: <https://genomeindia.in/index.php> (accessed January 23, 2025).

220. Deo SVS. Computerized clinical database development in oncology. *Indian J Palliat Care.* (2011) 17:S2–3. doi: 10.4103/0973-1075.76229

221. Liao J, Li X, Gan Y, Han S, Rong P, Wang W, et al. Artificial intelligence assists precision medicine in cancer treatment. *Front Oncol.* (2023) 12:998222. doi: 10.3389/fonc.2022.998222

222. Mishra US, Yadav S, Joe W. The Ayushman Bharat digital mission of India: an assessment. *Health Syst Reform.* (2024) 10(2):2392290. doi: 10.1080/23288604.2024.2392290

223. National Cancer Grid. NCG guidelines manual (2021). Available online at: <https://www.ncgindia.org/assets/ncg-guidelines-2021/ncg-guidelines-manual-2021.pdf> (accessed January 23, 2025).

224. Horgan D, Hajdich M, Vrana M, Soderberg J, Hughes N, Omar MI, et al. European health data space - an opportunity now to grasp the future of data-driven healthcare. *Healthcare.* (2022) 10(9):1629. doi: 10.3390/healthcare10091629

225. Badawi O, Brennan T, Celi LA, Feng M, Ghassemi M, Ippolito A, et al. Making big data useful for health care: a summary of the inaugural mit critical data conference. *JMIR Med Inform.* (2014) 2(2):e22. doi: 10.2196/medinform.3447

226. The Hindu. IndiaAI mission to boost artificial intelligence (2024) Available online at: <https://www.thehindu.com/news/national/karnataka/indiaai-mission-to-boost-artificial-intelligence/article68927972.ece> (accessed January 23, 2025).

227. Press Information Bureau. Cabinet approves over Rs 10,300 crore for IndiaAI mission, will empower AI startups and expand compute infrastructure access. (2024). Available online at: <https://pib.gov.in/PressReleasePage.aspx?PRID=2012375> (accessed January 23, 2025).

228. Ministry of Health & Family Welfare, India. Status of the action plan for cancer screening in rural areas. (2024). Available online at: <https://pib.gov.in/PressReleaseFramePage.aspx?PRID=2040933> (accessed January 23, 2025).

229. Ministry of Electronics and Information Technology. Responsible AI for social empowerment (RAISE). (2020). Available online at: <https://indiaai.gov.in/raise> (accessed January 23, 2025).

230. Cascini F, Gentili A, Causio FA, Altamura G, Melnyk A, Beccia F, et al. Strengthening and promoting digital health practice: results from a global digital health partnership's survey. *Front Public Health.* (2023) 11:1147210. doi: 10.3389/fpubh.2023.1147210

231. Rani P, Dutta K, Kumar V. Artificial intelligence techniques for prediction of drug synergy in malignant diseases: past, present, and future. *Comput Biol Med.* (2022) 144:105334. doi: 10.1016/j.combiomed.2022.105334

232. Rodrigues R. Legal and human rights issues of AI: gaps, challenges and vulnerabilities. *J Responsible Technol.* (2020) 4:100005. doi: 10.1016/j.jrt.2020.100005

233. Bartlett M. Solving the AI accountability gap - towards data science. Medium (2021). Available online at: <https://towardsdatascience.com/solving-the-ai-accountability-gap-dd35698249fe> (accessed May 10, 2024).

234. U.S. Food and Drug Administration. Transparency for machine learning-enabled medical devices: guiding principles (2024). Available online at: <https://www.fda.gov/medical-devices/software-medical-device-samd/transparency-machine-learning-enabled-medical-devices-guiding-principles> (accessed January 23, 2025).

235. U.S. Food and Drug Administration. Good machine learning practice for medical device development: guiding principles (2021). Available online at: <https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles> (accessed January 23, 2025).

236. U.S. Food and Drug Administration. Artificial intelligence and machine learning in software (2024). Available online at: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device> (accessed January 23, 2025).

237. Ministry of Law and Justice, Government of India. The digital personal data protection act (2023). Available online at: <https://www.meity.gov.in/writeradddata/files/Digital%20Personal%20Data%20Protection%20Act%202023.pdf> (accessed January 23, 2025).

238. Price WN, Cohen IG. Privacy in the age of medical big data. *Nat Med.* (2019) 25:37–43. doi: 10.1038/s41591-018-0272-7

239. Sathishkumar K, Chaturvedi M, Das P, Stephen S, Mathur P. Cancer incidence estimates for 2022 & projection for 2025: result from national cancer registry programme, India. *Indian J Med Res.* (2022) 156:598–607. doi: 10.4103/ijmr.ijmr\_1821\_22
240. Basurto-Hurtado JA, Cruz-Albarran IA, Toledano-Ayala M, Ibarra-Manzano MA, Morales-Hernandez LA, Perez-Ramirez CA. Diagnostic strategies for breast cancer detection: from image generation to classification strategies using artificial intelligence algorithms. *Cancers.* (2022) 14(14):3442. doi: 10.3390/cancers14143442