Title: Current State of Artificial Intelligence (AI) in Oncology: A Review

Author(s): Aqib Ali¹, Samreen Naeem², Sania Anam³, Muhammad Munawar Ahmed⁴

Affiliation(s): ¹Concordia College Bahawalpur, Pakistan.
²Superior College Bahawalpur, Pakistan.
³Govt Associate College for Women Ahmedpur East, Bahawalpur, Pakistan.
⁴Islamia University Bahawalpur, Bahawalpur, Pakistan.

DOI: https://doi.org/10.32350/cto.31.01

History: Received: November 16, 2022, Revised: May 04, 2023, Accepted: May 23, 2023, Published: June 15, 2023


Copyright: © The Authors

Licensing: This article is open access and is distributed under the terms of Creative Commons Attribution 4.0 International License

Conflict of Interest: Author(s) declared no conflict of interest
Current State of Artificial Intelligence (AI) in Oncology: A Review

Aqib Ali1*, Samreen Naeem2, Sania Anam3, Muhammad Munawar Ahmed4

1Department of Computer Science, Concordia College Bahawalpur, Pakistan.
2Department of Computer Science, Superior College Bahawalpur, Pakistan
3Department of Computer Science, Govt. Associate College for Women Ahmedpur East, Bahawalpur, Pakistan.
4Department of Information Technology, Islamia University Bahawalpur, Bahawalpur, Pakistan.

ABSTRACT

Artificial intelligence (AI) has made significant strides in recent years toward resolving a wide range of biological issues, including a number of occurrences connected to cancer. Deep learning is an adaptable sector of AI that enables the automatic extraction of features. It is increasingly being used in various fields of cancer research, both scientific and clinical. In this study, several current applications of AI in oncology, including the situations in which deep learning has effectively addressed intractable disputes, are discussed. This study also discusses the challenges that need to be surmounted before such an application of AI may be implemented more broadly. Similarly, it brings to light resources and data sets that can assist in maximizing the potential of AI. Essential insights in oncology may be generated based on the development of novel methodologies and uses of AI, making it possible for substantial changes to be made to clinical practice.

Keywords: artificial intelligence (AI), deep learning, machine learning (ML), oncology

1. INTRODUCTION

Molecular ‘omics’ (genomics, proteomics, and epigenomics) and ‘big data’ (which analyzes large databases while combining biomarkers and population trends to generate knowledge in real-time) are two of the most significant technological developments to have occurred in the field of oncology in recent times [1]. Before X-ray imaging and histopathology testing were available, cancer diagnosis was made using a combination of these two methods. In contrast, molecular tests may now detect changes in hundreds of genes and proteins, which can be used to diagnose cancer in an

* Corresponding Author: aqibcsit@gmail.com
individual, indicate their prognosis, and determine how they should be treated [2]. Targeted medicines for all tumor types have advanced cancer therapy. At present, cancer treatment relies on cytotoxic chemotherapy drugs. These medications attack fast-dividing cells, especially cancer cells. However, they can also destroy healthy cells, causing adverse effects. On the other hand, targeted medicines that target molecular and genetic defects which cause cancer have been developed, recently. These medicines may be safer and more effective than chemotherapy. Targeted medicines target cancer cell growth and survival proteins or signaling pathways. Immunotherapies use the immune system to fight cancer. Precision medicine uses genomic and molecular profiling to discover a patient’s cancer-causing genetic mutations. This data may be utilized to create customized cancer treatment for each patient. Targeted treatment and therapy frequently focus on targeting biomarkers already present in the tumor, as well as the signaling pathways involved in the progression of cancer, and modulate the patient’s immune system to fight the disease (for instance, checkpoint inhibitors) [3].

As a matter of fact, due to these advancements, the timespan patients may survive and the quality of their lives have improved. However, the application of precision medicine presents significant obstacles for medical personnel, notably due to the exponential increase of medical knowledge and the requirement to deliver highly customized cancer care [4]. This discussion is not limited to highly developed countries. Providing state-of-the-art, all-encompassing cancer care to millions of patients still poses a great challenge. This is especially true for populations living in suburban and rural areas. Regarding the application of AI, biomedical data has various characteristics that make it challenging to categorize the data. These characteristics include high dimensionality, temporal dependence, parity, and irregularity [5]. There is no optimized method available that may arrange and standardize the many pieces of patient-specific information (for example, narrative text in patient records and clinical notes, radiological scans, laboratory data, genetic information, pharmacogenomics, and prescription lists) [6].

This issue is further complicated by the utilization of several medical ontologies to generalize the data (for example, SNOMED-CT, UMLS, ICD-9, and ICD-10). These ontologies introduce conflicts and inconsistencies, further complicating the situation [7]. In addition, instructional and case
management support tools need to be built to ensure that complete and evidence-based information provided by machine learning technology is actionable for every patient. The efficient use of all-encompassing electronic health information systems, which should include data from the actual world, to direct the clinical decision-making process is a potential solution to the said problem.

Since 1956, when the term ‘artificial intelligence’ was first used, AI has made significant strides. The earliest developments in AI concentrated on building neural networks (sometimes written as ‘neural networks’), which were modeled after the capacity of the human brain to draw conclusions based on the information presented to it [8]. At the same time, the term ‘machine learning’ or ML started gaining traction due to advancements in artificial neural networks. The term refers to the capability of a computer program to analyze the data and recognize patterns within the data. This allows the program to gain knowledge from the data, which it may use to solve problems and make intelligent choices [9].

Figure 1. An Overview of Artificial Intelligence in Oncology [10]
Then came the trend of deep learning, also known as ‘deep learning,’ which is a more complex subset of ML. Unlike traditional ML methods, deep learning does not require human intervention for the machine to progress. Instead, it determines whether or not it has made accurate predictions on its own and then continues the process of learning. The AI powered computers of today make use of a combination of ML and deep learning and have the potential to be used in a wide variety of fields, including cancer [11]. According to a recent GE Healthcare and MIT Technology Review, 79% of doctors feel that AI technologies have helped expedite operations, enabling practitioners to provide better services and patient-centric care.

The key potential benefits of utilizing AI-enabled technology and supporting clinical decisions in oncology include

- Automated cancer diagnosis
- Enhanced prediction capabilities
- Real-time data updates
- Personalized attention
- Better results
- Increased efficiency, improved results, and reduced costs

AI and ML can help doctors to enhance cancer care. These technologies can improve treatment accuracy and personalization by analyzing complex data and discovering underlying patterns. This may improve patient outcomes and reduce adverse effects. AI and ML prediction models can detect high-risk individuals and enable early cancer treatment. Clinicians would have greater freedom and effectiveness in putting outcomes-based medicine into practice as a direct result of the ongoing development of scalable deep learning systems [12].

2. ROLE OF AI IN CANCER DIAGNOSIS

AI based algorithms provide a potential means of improving the accuracy of diagnostic imaging, thus freeing up the doctor’s time to devote to patient care. This is a win-win situation [13]. According a research recently published in Academic Radiology, in order for radiologists to keep up with their daily workload, they must analyze one image every three to four seconds, on average. Since they have access to more information than
their human counterparts, AI components in radiology and image processing would promote higher efficiency in these disciplines. Breast, lung, colorectal, prostate, and head and neck tumors, to mention a few, are some cancers that are now the focus of research efforts by scientists, worldwide. It appears that there is a massive opportunity to use AI in cancer, considering the rapid growth of research in this area [14].

More than 200 million women worldwide undergo mammograms yearly, making it the most widely used tool for detecting breast cancer at an early stage. Despite being a commonly used diagnostic tool, mammography faces challenges in accurately interpreting images. This is evident from the high rates of false-positive and false-negative outcomes, which require about 10% of the 40 million American women who undergo mammography each year to undergo further diagnostic imaging. However, only 4.5% of these women are eventually diagnosed with breast cancer. The use of AI would prevent an unnecessary diagnostic procedure for nearly half a million people [15].
For women with a biopsy-proven breast cancer diagnosis and average follow-up imaging results at least 365 days later, Google and two screening centers in the United Kingdom (25,856 women) and the United States (3,097 women) collaborated to develop a system to detect the presence of breast cancer using mammograms. Women who already had a breast cancer diagnosis were employed for this purpose. Six radiologists obtained data via clinical practice to compare these predictions as part of an independent study. The scientists provided evidence to support their assertion that the AI system outperformed both historical judgments made by radiologists and the findings reached by the six professional radiologists after evaluating 500 randomly selected instances [17]. Overall, the number of false-positives was reduced by 5.7% in the US and 1.2% in the UK, while the number of false-negatives decreased by 9.4% and 2.7%, respectively. The results of this pilot investigation paved the way for more extensive clinical trials that may hopefully improve the accuracy and efficiency with which breast cancer may be identified using digital technology [18].

As part of an effort to illustrate the advantages of employing the new AI in combination with radiologists, the MammoScreen system was examined. This tool is meant to locate areas in 2D mammograms where breast cancer is suspected and estimates the likelihood of malignancy. The scientists examined several metrics, including the reading time, the sensitivity, the specificity, and the area under the ROC curve (AUC). The study found that without AI, the average AUC among readers was 0.769 (95% CI, 0.724-0.814). With AI, the AUC was 0.797 (95% CI, 0.754-0.840). The significance threshold was 0.035, while the confidence range for the AUC difference was 0.002-0.055. Eleven of the fourteen radiologists who utilized MammoScreen reported an average of 18% reduction in the false-negative rate [19]. There was an improvement in screening uniformity amounting to a 25% reduction in false-positives among eight radiologists. The United States Food and Drug Administration (FDA) granted MammoScreen marketing clearance in March 2022.

As a further illustration, a computer vision deep learning algorithm used one thousand AI generated CT scans to teach itself how to scan lung tissue for abnormalities. The algorithm discovered that AI could identify lung cancer with a level of accuracy that was 30% higher than that of human beings [20]. These findings illustrate the potential for the AI technology to
identify cancer screening tests and increase their accuracy, when supplemented with the evaluations of human radiologists. While AI systems could identify pixel-level changes in tissues which remain imperceptible to human eye, people may employ the kind of reasoning that AI has no access to. The final objective would be to identify the most effective approach to integrate the two in order to revolutionize the field of radiology [21].

2.1. AI And Digital Pathology

On July 21, 2022, FDA approved the commercialization of software to aid pathologists in detecting suspected cancer spots in digitally scanned histology plates or slides of prostate biopsies [22]. Prostate biopsies would be facilitated by this program. In case of prostate, Paige Prostate is the first artificial intelligence-based program which was developed specifically to locate a region of interest within the picture with the highest likelihood of harboring cancer. This allows the pathologist to investigate the region further if it has not been adequately recognized [23].

FDA analyzed the results of a clinical investigation in which 16 pathologists examined 527 prostate biopsy photos digitized through a scanner. The images ranged from 171 cancerous to 356 benign images. The pathologist performed one reading without the aid of Paige Prostate (the process is also known as an ‘unassisted reading’). The other reading was performed with the assistance of Paige Prostate (also known as assisted reading) [24]. According to the findings, using Paige Prostate enhanced cancer detection on single-slide photos by an average of 7.3%, as compared to the unaided readings performed by pathologists on full-slide images of single biopsies. At the same time, the readability of the images’ soft slides remained unaffected. To conclude, both false-negative and false-positive results are possible; however, these are less likely when the device is used as a supplement and when the patient’s case is reviewed by a trained pathologist who takes into account the patient’s medical history and any other relevant clinical information before making a diagnosis. Indeed, material analysis is required in the lab before a diagnosis can be determined [25].

As part of the AI for the health challenge conducted by the Ile-de-France Region in 2019, the RACE AI study’s findings were presented at the European Society for Medical Oncology (ESMO) in 2021. These results confirmed the deep learning analysis applied in digital pathology to
categorize patients with localized breast cancer into high-risk and low-risk of metastatic relapse at five years. Additionally, the study reliably assessed the risk of relapse with an area under the curve AUC. This AI might become a valuable tool for therapeutic decision-making in adjuvant treatment [26].

2.2. Applying AI to Identify Molecular Shifts in Cancerous Tumors

Researchers in the United Kingdom found that an algorithm was able to distinguish between cancerous and non-cancerous tissues. Moreover, it was able to identify aberrant genomic patterns in a total of 28 different types of cancer. This is the result of presumably the most comprehensive pan-cancer analysis ever conducted in training computer vision combined with digital pathology. It was determined that in addition to arising via chromosomal rearrangements and tandem duplications, genetic variations may also arise from point mutations in classical genes [27].

The TCGA’s fresh-frozen tissue slides from 10,452 cancer patients were examined. A set of 14 standard tissue samples devoid of any genetic mutations served as control. In 2016, Google created an algorithm to recognize everyday online items and the same algorithm was used to evaluate presentations. More than 160 DNA patterns, hundreds of RNA changes in tumors, and a number of tumor-infiltrating lymphocytes (TILs) and their location were accurately detected across all 28 cancer types. The discovery of TILs can enhance our comprehension of the tumor microenvironment as a predictor of immunotherapy and targeted treatments. Since fresh-frozen tissue cannot be obtained from 99% of cancer patients, researchers would have to switch to working with paraffin-fixed tissues to use AI, even though the technology shows promise [28].

AI research can drive and accelerate new developments in the fields of structural biology, physics, and machine learning, as shown by Google DeepMind’s AlphaFold system. This system demonstrated how AI research could accelerate new developments in these fields by predicting the 3D structure of a protein based solely on its genetic sequence. The 3D models of proteins generated by AlphaFold are very accurate. This is an excellent example of AI’s use to discover new therapeutics. Proteins fold themselves naturally in a matter of milliseconds and the knowledge of their structure cannot be deduced from their genetic sequence [29]. Their functions might differ depending on the three-dimensional structure they take on. Due to the interactions between amino acids, modeling a protein with a higher
molecular weight is inherently more challenging. This so-called ‘protein folding issue’ has provided the impetus for a vast array of innovations, ranging from IBM’s efforts in supercomputing (BlueGene) to new initiatives (Folding@Home and FoldIt) and engineering disciplines including rational protein design. These techniques, which are based on deep neural networks, have the potential to contribute to drug discovery, while also lowering the cost of experimentation [30].

Bypassing preclinical inbred models, which tend to forecast the outcome of clinical research with only a moderate degree of accuracy, for instance, an AI system may search for biomarkers and determine the likelihood that a medicine would be approved by FDA [31]. According to the iNetMed model, AI might replace the first and last steps in the preclinical drug discovery process. iNetMed’s computational arm models the disease by creating a map of successive changes in gene expression. This map recognizes the gene expression patterns with mathematical precision, which the computational arm of iNetMed then uses in a Phase ‘0’ clinical trial based on a living biobank of organoids. The translational arm of iNetMed is responsible for developing new treatments for diseases. This technique might significantly alter the way academics examine large amounts of data to discover relevant insights that are of enormous utility to patients, the pharmaceutical industry, and the healthcare systems [32].

2.3. Utilizing AI to Get Optimal Therapy

A team from the University of Toronto developed an AI tool that showed considerable promise in lowering the amount of time needed to adapt radiation therapy regimens to the specific needs of individual patients. This specific AI tool made therapy recommendations based on past radiation data and its success rate was equivalent to that of radiation oncology doctors. It was able to recreate the detailed treatment plans that the top doctors arrived at after several days of hard work in only twenty minutes, which optimized the planning of radiation treatment [33].

In order to differentiate and assign importance to specific areas of an image, neural ‘jumping’ is used in AI and ML algorithms based on convolutional neural networks (CNN). These are meant to mimic neural networks in the cerebral cortex that allow us to take ‘shortcuts’ in our thinking and to conjure rational ideas without ‘convoluted’ nonlinear efforts. An analysis of 18F-FDG PET/CT images was used to test the
hypothesis that the protein expression status of programmed death ligand-1 (PD-L1) and the presence of epidermal growth factor (EGFR) mutations in 837 patients with NSCLC may be captured using these methods [34]. These algorithms may one day advocate using tyrosine kinase inhibitors (TKIs) or checkpoint inhibitors (ICIs) in real-time. The values under the curve for predicting EGFR mutation status were 0.86, 0.83, and 0.81, for training, validation, and external test cohorts, respectively. While, the values for discriminating positive PD-L1 status were 0.89, 0.84, and 0.82 for the same metrics. Furthermore, this held across all three cohorts—the internal, the external, and the validation. Moreover, in 67 patients treated with TKIs and 149 patients treated with ICIs, excellent patient ratings were associated with progression-free survival (PFS) [35].

At this point, the most practical use of AI in the clinical practice of cancer consists of combining advice from deep learning with human supervision, as well as the operationalization of suggestions made by AI with the help of medical staff. The feasibility, repeatability, scalability, and advantages of a deep learning-based virtual tumor board (DLVTB) were established in research carried out by Massive Bio (New York) on a cohort of 35 patients with advanced colorectal adenocarcinoma (CRC) [36]. Compared to the data from the past, the DLVTB cohort’s use of deep learning based on natural language exhibited an increase of 12 months in the median survival time for each patient. In addition to suggesting biomarker testing for 71% of patients, it indicated that 80% of patients should undergo precision oncology treatment. Moreover, 63% of patients met the requirements for participation in at least one clinical trial, which is much higher than the national average of 3%. SYNERGY-AI, an open, decentralized registry that can accelerate these efforts globally and at scale to cancer patients and access clinical trials using AI, has been developed since then. At cancer centers authorized by the National Cancer Institute (NCI) in the United States, this Massive Bio AI system is currently implemented [37].

**Table 1.** A dictionary of Terms Related to Artificial Intelligence (AI) and Precision Oncology

<table>
<thead>
<tr>
<th>Terms</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm</td>
<td>A predetermined course of action to take in order to resolve an issue or carry out an activity</td>
</tr>
</tbody>
</table>
### Terms

<table>
<thead>
<tr>
<th>Terms</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area under curve</td>
<td>A measure of the accuracy of a classifier when applied to a binary classification.</td>
</tr>
<tr>
<td>Artificial intelligence</td>
<td>Intelligent systems exhibit intelligent behavior by analyzing their surroundings and taking actions to varying degrees of independence in order to accomplish their objectives.</td>
</tr>
<tr>
<td>Artificial neural network</td>
<td>A computer model in machine learning takes its cues from the biological architecture and processes of the human brain.</td>
</tr>
<tr>
<td>Computer-aided detection/diagnosis</td>
<td>Techniques that make use of computer science to aid in the interpretation of medical pictures by medical professionals.</td>
</tr>
<tr>
<td>Deep learning</td>
<td>An area of machine learning that replicates the capacity of the human brain to execute unsupervised learning tasks by employing numerous layers of neural networks in its models.</td>
</tr>
<tr>
<td>Machine learning</td>
<td>A subfield of computer science that focuses on the construction of computational models that are capable of ‘learning’ from their input data and making accurate predictions.</td>
</tr>
<tr>
<td>Radiomics</td>
<td>A process that takes radiological pictures and extracts and examines significant quantities of more advanced quantitative image attributes to produce databases that can be mined.</td>
</tr>
<tr>
<td>Radio genomics</td>
<td>The study of how cancer imaging characteristics and gene expression are correlated with one another.</td>
</tr>
</tbody>
</table>

**3. CONCLUSION**

Opportunities to develop precision oncology beyond genomes and traditional molecular approaches can be found through the multimodal integration of sophisticated molecular diagnostics, radiologic and histologic
imaging, and encoded clinical data. However, the vast majority of medical data sets are not adequate to effectively train contemporary ML strategies. Hence, a lot of work needs to be done before this issue is resolved. In order to achieve success in biomedical research, combined data engineering efforts, computational approaches for analyzing heterogeneous data, and instantiation of synergistic data models are necessary. The future looks very bright and developed countries, which have a unified health care system, have a unique opportunity to take advantage of AI and so-called ‘big data’ in order to attract investment in clinical research, bioinformatics development, and the creation of new therapeutics.

REFERENCES


7. Pfaff ER. Cohort identification using semantic web technologies: Ontologies and triplestores as engines for complex computable


32. Volansky M. *Electronic phenotyping via the anchor and learn framework with physical therapy emphasis* [Doctoral dissertation]. Rutgers: The State University of New Jersey; 2020.


