Zodwa Dlamini *Editor*

Artificial Intelligence and Precision Oncology

Bridging Cancer Research and Clinical Decision Support

Springer

Artificial Intelligence and Precision Oncology

Zodwa Dlamini
Editor

Artificial Intelligence and Precision Oncology

Bridging Cancer Research and Clinical Decision Support

Springer

Editor
Zodwa Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI)
University of Pretoria
Pretoria, South Africa

ISBN 978-3-031-21505-6 ISBN 978-3-031-21506-3 (eBook)
https://doi.org/10.1007/978-3-031-21506-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Precision oncology is a novel approach for cancer care where diagnosis, prognosis, treatment and disease monitoring are tailored to an individual patient or group of patients, ensuring higher rates of patient survival and treatments with less side effects and more success in controlling or eliminating the disease. The advent of novel high-throughput technologies in recent years such as next generation sequencing (NGS) has led to an "omics" revolution, generation of massive amount of complex patient data, that has prompted the development of novel computational infrastructures, platforms and tools to store, retrieve and analyse this data efficiently. Artificial Intelligence (AI) is ideal for interpreting patterns in large datasets and offers unique opportunities for advancing precision oncology. AI can accurately interpret not only omics data, but it can also integrate data from other sources such as patient reports and medical imaging to give a more precise view of the individual or population, allowing for better clinical decision making.

In this book, we provide an overview of AI in precision oncology and it is divided into 3 parts. First part: Artificial Intelligence for Screening, Diagnosis and Monitoring in Precision Oncology. This section includes the use of AI and novel biomarkers, including circulating cell free nucleic acids (ccfNAs), in the diagnosis, prognosis and monitoring of cancer. It also focuses on the use of AI-enhanced digital pathology and radiogenomics in precision oncology. Second part: Artificial Intelligence and Omics in Precision Oncology. It highlights the use of AI and epigenetics, metabolomics and microbiomics in precision oncology. These sources of omics data are relatively recent sources of data and are highlighted here as they represent a departure from the more often discussed genomic, transcriptomic and proteomic data. Third part: Artificial Intelligence in Cancer Therapy and Its Clinical Applications. It highlights the use of AI-based medical devices, AI-guided drug design to target alternative splicing in cancer, AI prediction tools in maximising therapeutic efficacy and AI-empowered decision support systems such as AI-Pathway companion in recommending the most effective therapeutic approaches. It also highlights the use of AI tools for risk prediction, early detection, diagnosis and accurate prognosis.

The authors are basic scientists and clinical experts working in the field of cancer research and have forged a collaborative effort and writing on this important transdisciplinary subject of AI and Precision Oncology in cancer care and clinical decision making.

The editor has chosen a unique opportunity to capture the most up-to-date perspectives in Artificial Intelligence and Precision Oncology, which is part of advancing Society 5.0 and Healthcare by using digital technologies and providing opportunities for improving cancer care delivery and outcomes.

Pretoria, South Africa Zodwa Dlamini

Contents

Editor and Contributors

About the Editor

Zodwa Dlamini is a Professor of Molecular Oncology and the founding Director of the Pan African Cancer Research Institute (PACRI) in South Africa. She is the Director of South African Medical Research Council (SAMRC) Precision Oncology Research Unit (PORU). She is the DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP). She is a member of the American Association for Cancer Research (AACR) Advisory Group on Sub-Saharan Africa. She is a member of the AACR Pathology Resources in Africa Advisory Group. Professor Dlamini is a member of the African Organisation for Research and Training in Cancer (AORTIC) Research Committee Scientific Advisory Board. She is an elected member of the Academy of Science of South Africa. Professor Dlamini is an overseas Fellow of the Royal Society of Medicine (London). She is the Associate Editor for Frontiers in Oncology: Cancer Genetics.

Contributors

Afra Basera Department of Medical Oncology, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Ahmed Bhayat Department of Community Dentistry, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

Meshack Bida Department of Anatomical Pathology, National Health Laboratory Service (NHLS), University of Pretoria, Pretoria, South Africa

Nkhensani Chauke-Malinga Department of Plastic and Reconstructive Surgery, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Botle Precious Damane Department of Surgery, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Demetra Demetriou SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Zodwa Dlamini SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Bahoueli Gaudji Department of Urology, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Rodney Hull SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Rosemary I. Kabahuma Department of Otorhinolaryngology, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Mmamoletla Kgoebane-Maseko Department of Anatomical Pathology, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

Mahlatse Cordelia Kgokolo Department of Dermatology, Bridge A, Level 9, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Mpho Kgomo Department of Gastroenterology, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

Richard Khanyile Department of Medical Oncology, Steve Biko Academic Hospital and University of Pretoria, Pretoria, South Africa

Zarina Lockhat Department of Radiology, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Thifhelimbilu Luvhengo Department of Surgery, Charlotte Maxeke Johannesburg Academic Hospital, Parktown, Johannesburg, South Africa

Mzubanzi Mabongo Department of Maxillofacial and Oral Surgery, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa

Malose Makgoka Department of Surgery, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Mosebo A. Manabile SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Rahaba Marima SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Kgomotso Mathabe Department of Urology, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Precious Mathebela Department of Surgery, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Sikhumbuzo Z. Mbatha Department of Surgery, University of Pretoria, Pretoria, South Africa

Mandisa Mbeje SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
Department of Medical Oncology, Steve Biko Academic Hospital and University of Pretoria, Pretoria, South Africa

Zukile Mbita Department of Biochemistry, Microbiology and Biotechnology, University of Limpopo, Polokwane, South Africa

Langanani Mbodi Department of Obstetrics and Gynaecology, Gynaecologic Oncology Unit, Charlotte Maxeke Johannesburg Academic Hospital, University of the Witwatersrand, Johannesburg, South Africa

Mahlori Mkhabele SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Zilungile Lynette Mkhize-Kwitshana Department of Medical Microbiology, School of Laboratory Medicine & Medical Sciences, University of KwaZulu-Natal, Durban, South Africa

Thulo Molefi Department of Medical Oncology, Steve Biko Academic Hospital and University of Pretoria, Pretoria, South Africa

Daniel Montwedi Department of General Surgery, Steve Biko Academic Hospital and University of Pretoria, Pretoria, South Africa

Lesetja Motadi Department of Biochemistry, University of Johannesburg, Johannesburg, South Africa

Nompumelelo Mtshali Department of Anatomical Pathology, Faculty of Health Sciences, University of the Witwatersrand, Parktown, South Africa

Thanyani Victor Mulaudzi SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Shingai Mutambirwa Department of Urology, Sefako Makgatho Health Sciences University (SMU), Ga-Rankuwa, South Africa

Nomsa Nkosi Department of Colorectal Surgery, University of the Witwatersrand, Parktown, South Africa

Boitumelo Phakathi Department of Surgery, Faculty of Health Sciences, University of KwaZulu Natal, Durban, South Africa

Serwalo Ramagaga Department of Head and Neck, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa

Amanda Skepu Council for Scientific and Industrial Research (CSIR), Next Generation Health Cluster, Pretoria, South Africa

Chapter 1
The Application of AI in Precision Oncology: Tailoring Diagnosis, Treatment, and the Monitoring of Disease Progression to the Patient

Zodwa Dlamini and Rodney Hull

Abstract Personalised oncology has long been the ideal when it comes to the management of cancer. The ability to tailor screening, diagnosis, therapy and monitoring to an individual patient or group of patients would vastly decrease the burden of cancer while ensuring higher rates of patient survival and treatments with less side effects and more success in controlling or eliminating the disease. Precision oncology requires that as much information regarding the patient or population group be known. In terms of the underlying molecular basis of the disease, this is now being realised further to the advent of high throughput technologies such as next-generation sequencing (NGS) and advances in mass spectrophotometry. This has led to an "omics" revolution, with large datasets of information regarding the molecular basis of cancer in individuals being generated. Artificial intelligence (AI) is the ideal technology to manage and interpret these large datasets. In conjunction with machine learning (ML) and deep learning (DL), AI can more accurately interpret not only omics data, but it can also integrate data from other sources such as patient reports and medical imaging to give a more precise view of the individual or population, allowing for better clinical decision-making.

Keywords Precision Oncology · Artificial Intelligence · Screening · Diagnosis · Treatment · Disease Monitoring · Prognosis · Omics · NGS · Digital Twinning

Z. Dlamini (✉) · R. Hull
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: zodwa.dlamini@up.ac.za; rodney.hull@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_1

1.1 Introduction

The underlying molecular basis of cancer is complex and deciphering it has been the basis for many decades of research. The revolution in the available techniques that occurred in the 1970s led to the first in-depth studies concerning the molecular basis of cancer. Gaining an understanding of the changes that give rise to cancer at the molecular level allowed to not only understand how events in the body give rise to the disease but also how it progresses and also how these events could be targeted for the development of therapies. It was these initial studies that allowed for the development of the first drugs that could target molecules and signalling pathways to treat oncogenic processes, such as uncontrolled proliferation and resistance to apoptosis. When microarrays were first developed, their ability to create a profile of gene transcription in cancer (reviewed in (Govindarajan et al., 2012)) led to the first discussions of precision oncology. Precision oncology, a type of precision medicine, involves the tailoring of screening or treatment to an individual or specific population group based on the molecular profiles specific to that individual or group of individuals (Batch et al., 2022). The understanding of the molecular biology underlying cancer has been advanced in recent decades by the development of high throughput techniques such as next-generation sequencing (NGS) and advanced proteomics techniques such as SWATH. The data generated by these techniques has been used to decipher the molecular mechanisms of tumour initiation and progression. This data has also been used to construct database resources to integrate and analyse molecular mechanisms underlying cancer.

The ability of scientists to use these large datasets and databases to make useful observations and predictions concerning cancer is due to the advent and application of artificial intelligence. Artificial intelligence (AI) is an analytical or predictive operation performed by computers to emulate the decision-making processes of human beings. It has intensive problem-solving capabilities and can be used to perform tasks such as making predictions, scaling data, integrating different datasets, and reducing the dimensionality of data. Most importantly precision oncology can associate different patterns within data with real-world diagnoses, prognoses, or disease monitoring capabilities. The ability of AI to analyse large sets of data and transform this data into clinically actionable knowledge relies on the ability of AI to learn from either previous data or model teaching datasets. This learning ability is based on machine learning (ML) and deep learning (DL)-based approaches (Jiang et al., 2017) (Saltz et al., 2018) (Huang et al., 2020) (Ibrahim et al., 2020). The increase in interest in AI, precision oncology and precision medicine can be seen in the number of entries these topics find when used as search terms in PubMed. Standalone terms that carry entries for AI or precision medicine go back to the 1950s while the earliest entries for precision oncology date to the 1970s. There has also been a lot of interest in AI since the 1990s while interest in the other two increased rapidly from 2010 onwards. A combination of terms involving AI both precision medicine and precision oncology AND AI have only been topics of interest since 2017 (Fig. 1.1).

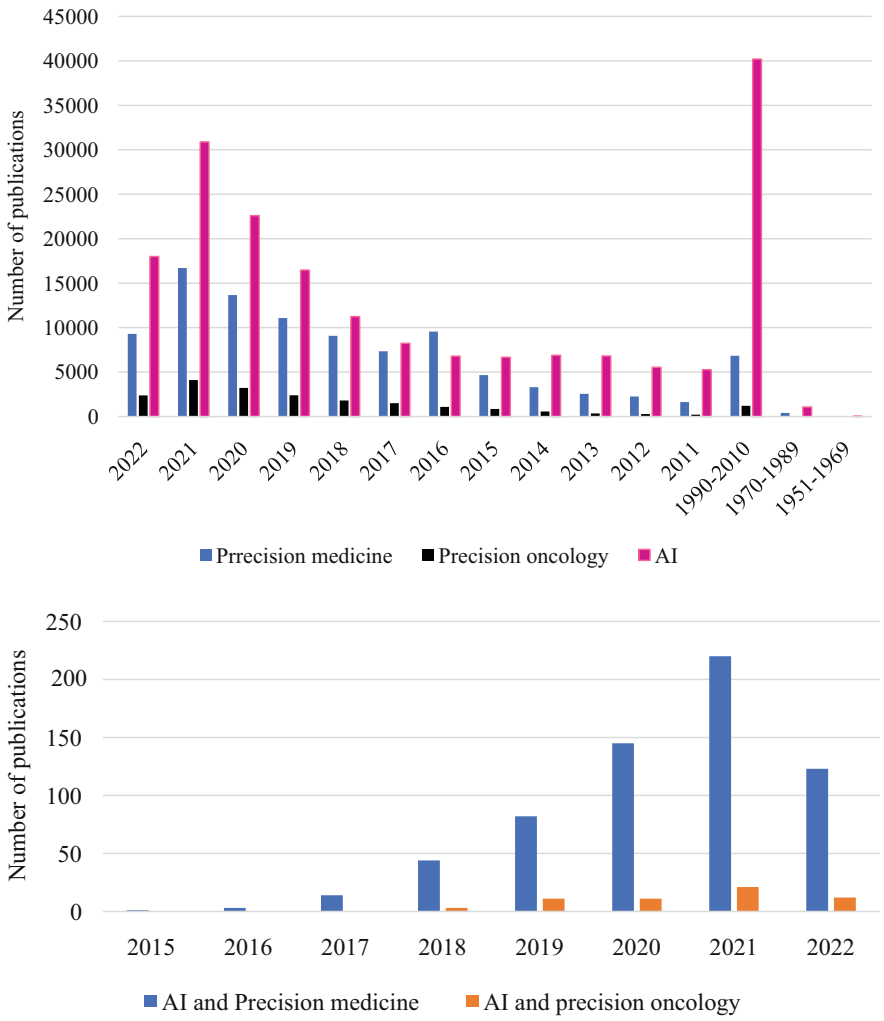


Fig. 1.1 PubMed entries on AI and precision medicine/oncology (**a**) using the terms independently the earliest references to AI or precision medicine come from the 1950s. While the earliest reference to precision oncology comes from 1977. All terms show an increase in the number of entries in PubMed. (**b**) The number of entries in PubMed for AI AND precision medicine and AI AND precision oncology since 2015/ This is a depiction of the number of entries identified in PubMed when the search terms AI AND precision medicine and AI AND precision oncology are used. For both terms, there are only regular entries after 2015 and the number of papers increases dramatically as time goes on. This demonstrates that these are topics of growing interest to researchers

1.2 AI in Medicine

In order for AI to accurately make predictions regarding a patient's health and treatment requirements, it must be able to learn from previous data and analyses. In this way, it emulates the human clinician learning from past experiences. The ability of computing algorithms to learn and adjust their performance to better recognise patterns in data is known as machine learning (ML). Initially, an AI does this using training data to create or fine tune mathematical models (Hakenberg et al., 2012). Deep learning (DL) is a specific type of ML which uses data that is labelled (supervised) and data that is unlabeled (unsupervised) in the training process. It integrates these different types of data by using multi-layer non-linear analysis and classification. One of the applications of DL is in a process known as natural language processing, and reinforcement learning (Falk et al., 2019) (Kaelbling et al., 1996).

Natural language processing (NLP) algorithms use two terms and establish if they are linked by counting the number of times they occur together. If they occur together more frequently than they are associated (Cheng et al., 2008) (Santus et al., 2019). This technique is used to search large amounts of literature or databases of information for articles or cases of interest. This is important because there are vast amounts of literature relating to cancer research and studies. One of these algorithms, known as MEDscape uses NLP to search and organise medical patient notes. The useful data retrieved from these notes is used to automatically update patient records (Morin et al., 2021). AI using NLP algorithms have been used to accurately predict patient outcomes using a variety of data including imaging reports and oncologist notes from thousands of patients with multiple different tumour types. The predictions the AI was able to make included cancer progression, treatment response and the likelihood and speed of metastasis (Kehl et al., 2021).

AI makes use of neural networks to copy the way humans think and interpret data but without user bias and human error. These neural networks allow AI to make logical conclusions similar to those that could be reached by humans (Joshi et al., 2021). These networks use multiple fundamental computing units (neurons) to convert raw input data into classified, annotated and analysed output data. The nodes are connected to form a network that contains multiple layers including an input layer, multiple functional or hidden layers, and an output layer (Kuwahara et al., 2021). There are multiple types of neural networks. Artificial Neural Networks (ANNs) use multiple interconnected computational neurons that distribute data analysis tasks. These networks are useful for analysing multidimensional complex data. The distribution and initial decisions the network make regarding this data are based on the learning by these algorithms. This algorithm also analyses the data sorting decisions by analysing if these decisions make the outcome worsens or improves the output (Baskin et al., 2016). Convolutional neural networks (CNNs), a type of ANN, contain neurons that are self-optimised through learning. (O'shea et al., 2021). They are classed as Deep Neural Networks, because CNNs have multiple layers (Alquraishi & Sorger, 2021). Recurrent neural network (RNN)

Table 1.1 Decision tree techniques

Decision tree technique	Method	Reference
Random decision forests	Construct multiple decision trees at the training stage. The final decision is the most common output.	(Ho, 1995)
Neighbour joining	Bottom-up method where outputs that are most similar are grouped together.	(Saitou & Nei, 1987)
Regression analysis	Nodes represent the mean of the results of the preceding nodes.	(Kamiński et al., 2018).
Binary Decision Tree	*Sequential decision process with features evaluated with one of two outcomes.*	(García Márquez et al., 2019)

remembers previous analyses both the inputs and resulting outcomes and then treats all future inputs and outputs as related (Dupond, 2019).

AI must be able to make decisions to perform its analysis and useful feature selection. The decision tools used are generally decision trees. These decision tools are named trees as the graphical representation of the decision-making process resembles a flowchart. The AI performs a test or analysis of each piece of data, and this gives rise to separate results. Each decision is represented as a node and each result represented as a branch. The final results then lead to a further analysis of each branch. This gives rise to the branched tree structure with some decisions (results proving to be dead ends). The final terminal nodes are known as classification or label (Kamiński et al., 2018). There are different types of trees as shown in Table 1.1 and Fig. 1.2.

AI has used decision trees to improve diagnosis. One study used lung cancer samples from the Lung Image Database Consortium (LIDC) dataset. This data was split 90% for training and 10% for testing. A labelled subset of the training set was used to train a CNN-based ransom decision tree. Once the CNN random decision tree was trained it was tested on the test data. This tree was able to accurately assign labels to the unlabeled data (Zheng et al., 2019). The origin of cancer tissue gas has been predicted based on miRNA profiling using AI based on two types of decision tree. Firstly, with neighbour joining methods and secondly with binary decision tree analyses. The neighbour joining method with an accuracy of 93.9%. The prediction accuracy of the binary decision tree method was 84.8% (Park et al., 2021).

Guidelines have been established in order to assist in the validation of the analysis provided by AI. These guidelines are known as the critical assessment of genome interpretation (CAGI) and were formulated using variants that were experimentally validated to cause disease and assessing if those predictions obeying the guidelines match these validated results (Andreoletti et al., 2019). The fifth edition of CAGI created in 2021 consists of 14 questions or criteria known as challenges (Andreoletti et al., 2019).

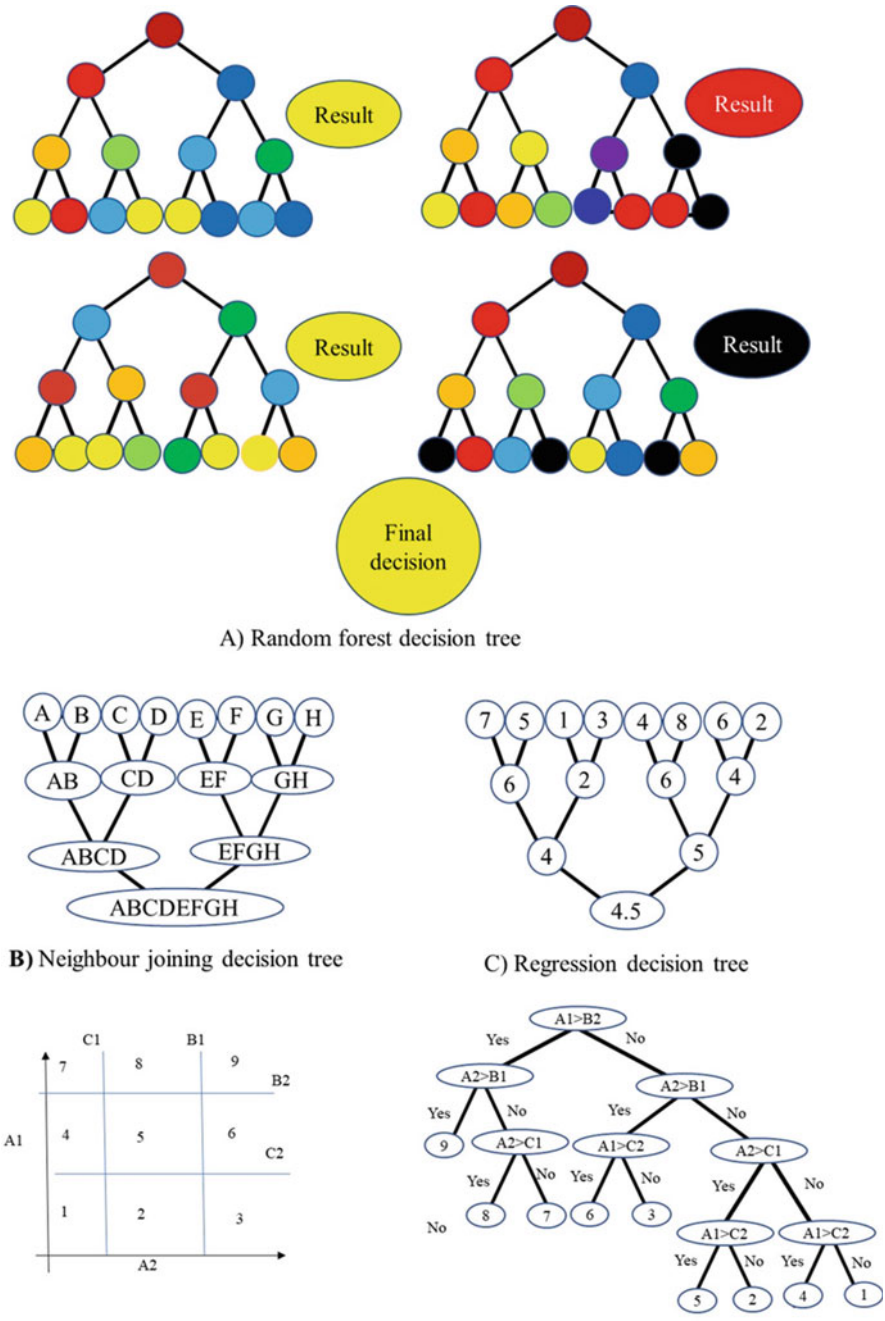


Fig. 1.2 Depictions of common decision tree methods (**a**) Random Forest trees use multiple trees and then select the most common outcome (**b**) Neighbour joining tree group nodes by similarity and select between these similar nodes (**c**) Regression trees the nodes are the mean of the previous nodes (**d**) Binary trees with sequential decisions based on one of two possible outcomes

1.3 Biomarker Discovery and Application

An ideal strategy to improve the screening, diagnosis, classification, staging and treatment of various cancers, is the identification of various molecules or molecular patterns or profiles that can serve as biomarkers. These biomarkers can be genomic mutations, transcripts, non-coding RNAs, proteins, metabolites, or even epigenetic markers. When patients present with symptoms indicating that they may have cancer, the standard procedure is a physical examination and radiographic imaging, this may be followed by biopsy examination. Many cancers screening procedures require invasive or expensive procedures. Biomarkers are normally classified as prognostic or predictive. Prognostic biomarkers are used to categorise patients by their risk of developing disease (screening), diagnosis of the disease, risk of disease progression, severity of disease and risk of death from the disease (Echle et al., 2021). Predictive biomarkers can be used to select a targeted treatment. These predictive biomarkers can also be used for drug discovery or in clinical trials for new treatments (Echle et al., 2021). The discovery of these biomarkers relies on the use of large omics datasets and the identification of patterns of the presence or absence of molecules in these large datasets that can be associated with disease. AI is a vital tool in this discovery process as it allows these large datasets to be rapidly and accurately analysed and associations with diseases to be identified. This is made possible through the use of machine and deep learning algorithms. Indeed, DL-based image analysis has broad applications in multiple fields of modern medicine that involve image data: in radiology, DL performs (Echle et al., 2021).

Liquid biopsies involve the identification of biomarkers in various body fluids. This can be blood, urine, saliva, or even cerebral spinal fluid. This is a more ideal diagnostic or prognostic technique than normal biopsies as they are less invasive and traumatising to a patient. This is also an important consideration for precision medicine as these samples can be obtained and analysed rapidly to give a current view of the patients' health and status (Kaur et al., 2017). These biomarkers can be transcripts, genomic markers in the form of DNA, proteins, or metabolites. In the case of RNA and DNA the transcripts would appear in biological fluids in the form of circulating cell free nucleic acids (ccfNAs). These ccfNAs have already been used as biomarkers in cancer diagnosis, prognosis, and monitoring (reviewed in (Pös et al., 2018)). It has also been established that these ccfNA appear in higher amounts in disorders such as cancer (Pös et al., 2018).

These nucleic acids can be in the form of cell free DNA which is fragmented DNA usually no longer than 450 bp in size. This DNA can be either of genomic or mitochondrial origin (Thierry et al., 2016). Circulating cell free RNAs include mRNA transcripts, non-coding RNAs, such as microRNA (miRNA), long non-coding RNA (lncRNA) and circular RNAs, transfer RNAs and ribosomal RNAs (reviewed in (Pös et al., 2018)). These nucleic acids are normally released into the body fluids as a result of cell death or in the case of many of the RNA molecules through active secretion (Vita et al., 2022).

1.4 Multi-omics Data

High throughput techniques like NGS allow for in-depth analysis of the mutational landscapes, gene expression patterns and epigenetic modifications for a large number of samples. Integration of "multi-omics" (genomics, epi-genomics, transcriptomics, proteomics, and metabolomics), and "non-omics" (medical/mass-spectrometry imaging, patient clinical history, treatments, and disease endemicity) data could help overcome the challenges in the accurate detection, characterisation, and monitoring of cancers. The complex analysis, annotation and combination of various omics data is sometimes only possible following data simplification. When these simplification processes are performed it is important to note that it may lead to the loss of information. The complexity of data is normally measured by the number of dimensions (variables) it has (Pezoulas et al., 2021). This reduction allows for increased ease and speed of analysis as well as a reduction in the space needed to store the data (Meng et al., 2016).

1.4.1 Genomics

The generation of large genomic datasets is due to advances in next-generation sequencer (NGS) (Paolillo et al., 2016)) and in silico computational algorithms. Whole genome sequencing allows for the analysis of all genomic alterations in cancer. It gives information regarding the number and identity of driver mutations and allows the mutational signature of the tumour to be identified. WGS has led to multiple sequencing projects and the establishment of databases containing the DNA sequence profiles of many cancers. These databases are listed in Table 1.2. To be truly useful genomic data must be integrated with clinical data, patient demographics, survival data, treatment status (Robinson et al., 2017). This is needed to link genomic events to specific cancers prognoses, and treatment responses (Robinson et al., 2017). AI has immense potential to contribute significantly at every stage of cancer management ranging from reliable early detection, stratification, determination of infiltrative tumour margins during surgical treatment, response to drugs/therapy, tracking tumour evolution and potential acquired resistance to treatments over time, prediction of tumour aggressiveness, metastasis pattern and recurrence (Bi et al., 2019).

1.4.2 Transcriptomics

Transcriptome includes the transcribed mRNAs, the alternately spliced isoforms of those mRNAs as well as non-coding RNAs such as miRNA. Any study looking at all these transcripts will aim to identify all the transcripts involved in metabolic

Table 1.2 DNA sequence databases and their applications

Database	Application	Reference
The Cancer Genome Atlas (TCGA)	Understand the molecular basis of cancer through the application of genomics.	(Wang, Jensen, & Zenklusen, 2016)
International Cancer Genome Consortium (ICGC)	Voluntary collaborative forum.	(Zhang et al., 2019)
CatLog of Somatic Mutations in Cancer (COSMIC)	Catalogue of somatic mutations in human cancer showing the impact of these mutations.	(Forbes et al., 2015)
The NCI's Genomic Data Commons (GDC)	A unified repository for cancer knowledge enabling data sharing across cancer genomic studies in support of precision medicine.	(Gao et al., 2013)
cBioPortal	Provides visualisation, analysis and download of large-scale cancer genomics data sets.	(Gao et al., 2013)
Methyl-Cancer	Database for human DNA Methylation and Cancer	(He et al., 2008)
UCSC Cancer Genomics Browser	Displays whole-*genome* and pathway-oriented views of genome-wide experimental measurements for individual and sets of samples.	(Goldman et al., 2013)
Moonshot project	Aims to address inequalities in access to cancer screening in the USA.	(Hsu et al., 2017)

processes and how they interact to result in gene expression. Studies that only examine specific sets of transcripts, i.e., mRNAs or miRNAs will provide answers to more specific questions. The result of epigenetic changes that occur in cancer can and have been studied by examining the transcriptome of cancers where these epigenetic changes have occurred. These studies have been undertaken in breast cancer (Robinson et al., 2015), prostate (Varambally et al., 2002) (Bhasin et al., 2015), head and neck squamous cell carcinoma (HNSCC) (Kelley et al., 2017).

1.4.3 Proteomics

Proteomic profiles reveal the actual cellular response to the conditions a cell is faced with. The change in protein expression also provides information regarding processes that affect protein modification, transport, and stability. Datasets of protein expression profiles are created using mass spectrometry and have been used to profile protein expression changes in response to therapy, monitor drug toxicity, and for diagnosis using specific biomarkers. These biomarker profiles, which are identified through protein expression signatures can also be to monitor disease progression, establish metastatic risk, do treatment follow-up to check for recurrence and stratify patients according to subtype (Keyl et al., 2022). Once again, these large data sets require AI to interpret them accurately, reliably, rapidly and consistently.

Many AI algorithms have been used to infer protein–protein interaction networks from proteomic datasets (Keyl et al., 2022). Another significant role for AI in proteomics is predicting docking capabilities between drugs and their target compounds.

AI can also be used to combine and integrate proteomic and genomic data to identify DNA mutations related to protein signalling. These genetic changes can then be said to be genetic drives of cancer. This has been performed in breast cancer, where the identification of signalling pathways specifically altered in different breast cancer subtypes was achieved. It also identified SKP1 and CETN3 as two new markers for basal-like breast cancer (Mertins et al., 2016). Proteomic and transcriptomic data can be integrated to identify changes in the splicing of mRNA and the generation of different protein isoforms that may be characteristic of different cancers (Liu et al., 2017). Proteomic data can show a much stronger association to the clinical characteristics of a patient, and this is reflected by the close association of integrated proteomics data with the clinical outcomes, for example MS analysis integrated with histopathological diagnosis (Huber et al., 2014). This can be done with very small amounts of extracted proteins, for example a study was performed where very small amounts of protein were analysed using LC-MS which led to deep coverage of entire proteomes of specific cell types (Kulak et al., 2017). A recent development has been the use of single-cell proteomics which has gained importance since it is able to give insights into cancer heterogeneity and the metastatic ability of single cells compared to colonies. It is also able to provide information concerning rare/mutated cells (Doerr, 2019). This has been successfully used to grade and rank acute myeloid leukaemia hierarchy (Schoof et al., 2021b).

1.4.4 Metabolomics

Metabolomics is the analysis of small molecules, such as amino acids, lipids, nucleotides, carbohydrates and organic acids, which are produced because of primary or secondary metabolic processes. The populations of these molecules changes during, growth, in response to stress and consequently during the development and progression of cancer (Bertini et al., 2009) (Lin et al., 2011) (Veselkov et al., 2011). Therefore, metabolomics can be used as an indicator of the molecular mechanisms underlying tumorigenesis.

It can also be used to monitor disease progression, the response of the tumour to drugs and other treatments. As with proteomics, the profiling of metabolites relies on mass spectrometry but with the additional use of nuclear magnetic resonance (NMR) spectroscopy (Merz & Serkova, 2009). Traditionally the sample had to be separated or fractionated to achieve the best results, but separation-free MS techniques have been developed which reduce the volume of sample required and reduce variation in the data generated through the analysis. These include direct infusion-MS, MALDI-MS, mass spectrometry imaging (MSI), and direct analysis in real-time mass spectrometry (Dettmer et al., 2007). The Global Natural Product Social Molecular

Networking (GNPS) is a *small-molecule mass spectrometry networking hub. Researchers can deposit their own MS data for small molecules and this repository is available for other uses to search and use*. GNPS has been shown to be very useful for cataloguing and organising MS/MS data using AI in the form of correlation and visualisation approaches. These can be used to identify spectra from related molecules (Wang, Carver, et al., 2016). Techniques such as principal component analysis or hierarchical clustering can be used in conjunction with ML to data mine these repositories to enhance the identification of spectra (Bertini et al., 2009) (Duan et al., 2005). These techniques have been used to identify metabolic biomarkers for multiple cancers including colorectal (Yamazaki, 2015), pancreatic (Zhang et al., 2012), lung (Zhuang et al., 2016), breast (Li et al., 2020), gastric (Ikeda et al., 2012), ovarian (Zhang et al., 2013) and prostate (Kelly et al., 2016).

1.4.5 Microbiomics

It has been estimated that the microbiota of the average human contains 40 trillion microbial cells (Sender et al., 2016). This microbiota is now known to play a role in the development and progression of cancer, especially through interactions with the nervous system and what is known as the gut–brain axis (reviewed in (Hull et al., 2021)). The profiling of all the microbial genes, metabolites, proteins and transcripts within a single patient is known as the patient's microbiome (Sepich-Poore et al., 2021). This can partly be explained by the interaction between the microbiome and the immune system as this may favour the development of cancer (Mangani et al., 2017). Microbiomes have been so closely associated with cancer, that it is now known that specific populations of microorganisms and microbial metabolites are associated with specific cancers. Therefore, different microbial signatures can be used as biomarkers to diagnose or monitor cancer, and affect the safety, tolerability and efficacy of specific treatments. Microbiomics are studies using the same high throughput techniques such as NGS and mass spectrometry. Once again this gives rise to large databases, which require the use of AI and machine or deep learning to analyse and interpret this data. Any attempt to integrate this microbiomic data with other "omics" data would require the use of AI (reviewed in (Cammarota et al., 2020)). AI can also be used to identify and evaluate microbiome community interactions with other microbes or the host. This is done using Network analysis and is useful for the identification of changes in these interactions may be caused by microbial community structure, environmental factors, metabolites, clinical. These networks can be constructed based on similarity or correlation coefficients between pairwise variables. Extended relationships can then be inferred based on these pairwise interactions. This is done using algorithms such as SparCC (Sparse Correlations for Compositional data) (Friedman & Alm, 2012) and Compositionally Robust Inference of Microbial Ecological Networks) (Faust et al., 2012).

1.5 Imaging

Medical imaging techniques such as Magnetic Resonance Imaging (MRI), CT scans, and Positron emission tomography (PET), are commonly used in the diagnosis of cancer. This is because these techniques are good at soft tissue contrast. This allows them to be good at locating tumours and monitoring tumour progression. They are also non-invasive and have a high resolution (Magadza & Viriri, 2021) (Menze et al., 2014). The aim of imaging cancer or suspected cancer tissue is known as tumour segmentation. This is the act of distinguishing between normal and cancerous tissue. This is a vital procedure for the use of imaging techniques in diagnosis and treatment planning, monitoring treatment response and disease progression (Bousselham et al., 2019). AI has been successfully used to automate the interpretation of medical imaging. It has been shown to be able to analyse stained sections of temper tissue and segment these images allowing for the identification and quantification of various parameters. These include the rate and amount of mitosis (Romo-Bucheli et al., 2017), the presence and abundance of mutations (Coudray et al., 2018), the differentiation between nuclei from benign cells versus those from cancer cells (Sirinukunwattana et al., 2016) (Xu et al., 2016), spatial localisation of proteins (Saltz et al., 2018). AI-based image analysis is more reproducible, objective and is quantitative compared to manual assessment. Convolutional neural networks (CNNs) are most commonly used for image analysis (Muhammad et al., 2020). There are two types of automated segmentation, generative and discriminating methods (Magadza & Viriri, 2021). Both methods use the same seven stages of analysis image acquisition, image preprocessing (deionising/enhancement/restoration), image segmentation/feature extraction and object recognition (Pan, 2007).

Image Segmentation techniques are all based on pixel-based selection to discern a Region of Interest (ROI). However, there are different methods that are used to achieve this, In the region-based method, a pixel in the ROI is selected as the reference or seed pixel. Neighbouring pixels are then compared to this pixel in order to establish if they are similar enough to be included (Punitha et al., 2018). In the edge-based method, the image is reduced to only its important structural characteristics. This decreases the image size. It also allows for the image's background to be separated from the object (Farag, 1992). The fuzzy theory-based method is an amalgamation of the region and the edge methods. (Basir et al., 2003). The partial differential equation (PDF) method calculates an energy of the image function. It then uses a partial differential equation (PDE) to describe the parametric curve evolution based on the energy of the image. It then uses this equation to find similar pixels (Sliž & Mikulka, 2016). In the threshold-based method, a grayscale binary image is created to reduce image complexity. This makes it easier to classify pixels (Bhargavi & Jyothi, 2014). Finally, the semantic segmentation network method classifies every individual pixel as either tumour or normal (Chen et al., 2017). When it comes to performing Whole Slide Image (WSI) segmentation some of these methods are more time and computing power consuming than others. The semantic method is the slowest and requires the most computing power (Guo et al., 2019).

1.5.1 Radiogenomics

Histopathological images have been integrated with genomics data in order to enhance feature selection based on cancer tissue architecture (López de Maturana et al., 2019). In a similar way, multi-omics data have been associated with features in medical images to develop predictive models using AI algorithms. This has been successfully performed for prostate cancer (Robinson et al., 2015), renal cell carcinoma (Schoof et al., 2021a), low-grade glioma (Brat et al., 2015), non-small cell lung cancer (Yu et al., 2016) and breast cancer (Yuan et al., 2012). This technique was initially given the name imaging genomics since it associated image features with genomic data. However, another term, radiomics or radiogenomics has been used to cover all the different omics data that can be associated with image features (Bodalal et al., 2019). Image features that can be associated with this omics data include structures, shapes, lines, points, colours or boundaries. It can even be extended to regions of the image associated with these features (Bi et al., 2019). In order to carry out a radiogenomics analysis, the AI must extract features identified on an image and link these features with phenotypes which is due to protein expression which can then be associated with genomic, transcriptomic and epigenomic or other omics changes (Rutman & Kuo, 2009). The appearance of these features on an image can then be an indication that these omics changes are present in the patient and the tumour. In the same way, these omics profiles can be used as indicators of for instance patient survival or disease progression, these associated image features can now be used to do the same (Berger & Mardis, 2018). AI is also necessary in radiogenomics as some of the features or changes in the cancer tissue may be so subtle that they may be missed by the human eye. Computer-assisted image analysis will accurately and consistently detect these changes based on what the algorithm has learned from previous data thanks to the application of machine and deep learning. These changes can then be associated accurately and without bias to any genomic, proteomic, transcriptomic, epigenomic, metabolomic or feature within the patient records. This is due to the analysis the AI can conduct on this data to extract unique features and then associate them with the unique image features. As previously stated, this integration would be too complex for a human being to complete accurately and timeously (Hussein et al., 2017). This end-to-end, automated data analysis or pipeline is able to compute and discriminate a vast number of features in both the image analysis and patient records or omics data to achieve the most accurate selection of features that are associated. And these models' ability to learn means that they are optimising their analytical ability and performance while integrating these data sets and looking for associations (Jansen et al., 2018).

Ai and radiogenomics have been shown to be able to predict the neoadjuvant therapy response in esophageal cancer using a convolutional neural network to analyse fluorodeoxyglucose positron emission tomography (18F-FDG PET) images. It was able to associate features from these images with transcriptomic data and make highly specific and accurate predictions (Ypsilantis et al., 2015). In another study, AI algorithms were used to identify image features within PET images in

breast cancer patients and associate these features with a genetic biomarker (Fujishima et al., 2017). Studies have also shown that features in images could be associated with tumour mutational burden, the average number of genetic mutations per megabase (Angus et al., 2019) and with the metastatic ability of the tumour (Trivizakis et al., 2019).

1.6 Drugs, AI and Precision Oncology

1.6.1 Drug Discovery and Re-purposing

The design or discovery of new drugs is a time-consuming and expensive undertaking with many potential compounds that have already had large amounts of money, $314 million to $2.8 billion, spent on them failing in the final stages. This means that all the time and money spent on them was essentially wasted (Waring et al., 2015). It is estimated that 90% of drugs fail to enter clinical trials for regulatory approval in (Fleming, 2018). AI can be used to remove those compounds most likely to fail from further development and prevent resources being wasted on them (Gawehn et al., 2016). This can be achieved using modelling to design better drugs by assessing a compound's binding abilities, identifying their binding partners that may be biologically significant and establishing if there are any toxic interactions they may have. Some of these modelling algorithms that have been developed and that are already in use include the quantitative structure-activity relationship (QSAR) models. These models still face problems since they need to learn from experimental data sets. If these datasets are small, it may decrease the accuracy of the model. If the data is not validated there may be errors that would lead to errors in the final model due to the algorithm learning from incorrect data (Roy & Pratim Roy, 2009) (Zhao et al., 2017). AI can also be used to search through chemical databases to identify compounds with a structure that may indicate their ability to bind to a specific target. The searching of these large libraries is known as high-throughput screening techniques (HTS) (Inglese et al., 2006) (Zhu et al., 2016).

AI can also be used to predict how a drug will behave with respect to its physicochemical properties, bioactivity and toxicity. Physiochemical properties can be predicted using AI-based tools such as using Quantitative Structure Property Relationship (QSPR) workflow. This algorithm was originally designed to predict the physiochemical properties of environmental toxins (Zang et al., 2017). Other algorithms have also been developed that are able to perform function such as predicting the solubility of a drug, these include undirected graph recursive neural networks and graph-based convolutional neural networks (CVNN) (Kumar et al., 2017). The efficacy of drugs can be predicted by establishing their affinity for their target molecule, and toxicity and side effects may be predicted by identifying any unintended interactions it may have. AI is able to accomplish these actions by calculating the binding affinities for the drug on a large number of molecules. It can do this by identifying any similar features or structures the drug has with similar

molecules or targets similar to the intended target (Öztürk et al., 2018). Screening for the most effective treatment for a specific patient is also possible using AI. One way this can be done is through the use of a digital twin.

1.6.2 Digital Twins

An important concept in the use of AI in medicine is the creation of a digital twin. This digital twin is the use of patient-specific data to create a virtual copy of the patient. An accurate digital twin requires accurate, detailed and up-to-date information about the patient (Batch et al., 2022). Deciding on the best treatment for an individual patient is one of the primary uses of the digital twin. This process is demonstrated in Fig. 1.3. As much data concerning the patient is gathered. This includes various omics data, patient records, medical imaging and imaging reports and any data concerning demographic or risk factors. AI then creates the digital twin. These twins are then duplicated, and each twin is given a virtual treatment. Using information regarding the molecular basis of these treatments, their side effects and case reports and studies of these treatments and AI algorithm can then run simulations for each individual treatment on the digital twin. The results can be used to select the best treatment option (Björnsson et al., 2019).

There are many ways these drugs can be tested in these simulations. One example is the use of protein–protein interaction (PPI) networks, constructed using a patient's proteomic or transcriptomic data as a map. Changes in protein expression caused by

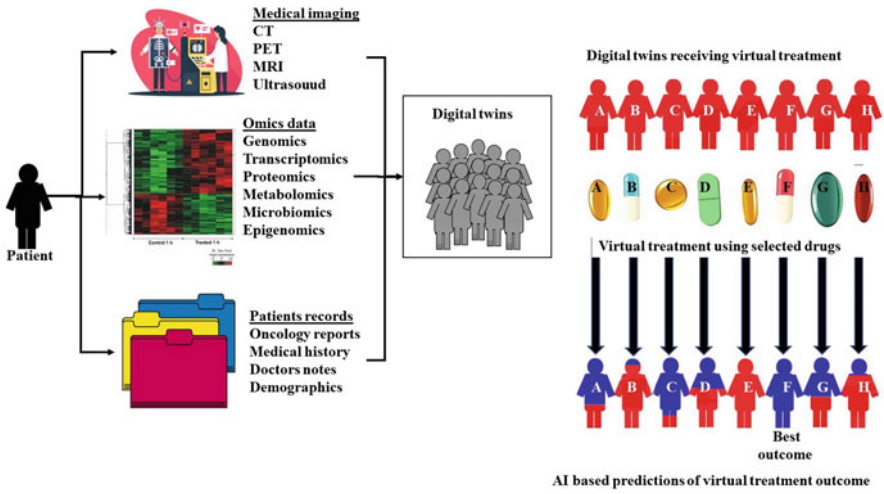


Fig. 1.3 The use of digital twins in drug discovery. Various types of data from a patient are used to create the most accurate digital twin possible. This twin is then duplicated and treated virtually with all available drugs. Artificial intelligence then calculates treatment outcomes based on drug molecular interactions and the molecular environment of the digital twin

a treatment can then be mapped to the patients PPI to identify changes in the pathways the drug could cause when used to treat the patient (Barabási et al., 2011) (Zhou et al., 2014). Another example could involve genetic changes detected in a patient. These alterations that lead to transcript and protein changes can be used to create a twin with the altered protein and protein expression patterns. A treatment targeting this protein can be used to treat the digital twin. The resulting effects on PPI and pathways can then be simulated in the twin.

1.7 Conclusion

The integration and analysis of data from various sources such as different "omics", medical images and medical imaging reports, electronic medical records, or handwritten doctor's notes, is only possible in a practical manner using AI and machine learning. The requirement for the use of AI has been necessitated due to the advancements in multidimensional "omics" technologies. The application of AI to biological data enables the understanding of complex biological systems. AI is already used in the automated extraction of information as well as the automated integration of health records. It is also currently used to organise, annotate and store data in big data storage systems such as cloud scaling. AI can outperform human clinicians and pathologists in all these tasks and it has enabled us to develop new techniques to study cancer, detect cancer at an early stage, more accurately predict patient outcomes. Decide on the correct treatment, monitor disease progression and treatment effectiveness, design new drugs and therapies, and stratify and classify tumours (Fig. 1.4).

This chapter has served as a brief introduction to the various topics that will be covered in the following chapters of this book. The initial chapters will examine the use of AI in the identification and application of novel biomarkers for precision oncology. This involves the use of these biomarkers in diagnosis, screening, monitoring drug resistance and in the choice of the most appropriate treatment regimen. They will also discuss the novel use of ccfNAs as biomarkers in precision medicine. The last of these initial chapters will discuss the use of digital pathology in accomplishing these tasks and how the new field of radiogenomics will allow image features to be associated with molecular signatures. The book will then discuss some of the less discussed "omics" that are studied to obtain data that can be used to identify biomarkers for use in precision oncology. These include epigenomics, metabolomics and microbiomics. The final chapters of the book will discuss the practical and clinical application of AI to precision oncology in detail. The first of these applications the book will discuss is the use of nanotechnology in AI-based precision oncology. It will then focus on the use of AI-based devices in cancer screening. This will be followed by a chapter describing the use of AI to design new drugs and then a chapter describing the application of AI to increase the efficiency of immunotherapy. For the final applications, the book will discuss the role AI can play in helping clinicians and oncologists choose the correct treatment

Fig. 1.4 A summary of the applications of AI in precision oncology

for an individual patient using various AI tools and techniques. The concluding chapter will summarise the topics covered and offer insights into the future of AI in precision oncology.

References

Alquraishi, M., & Sorger, P. K. (2021). Differentiable biology: Using deep learning for biophysics-based and data-driven modeling of molecular mechanisms. *Nature Methods, 18*, 1169–1180.

Andreoletti, G., Pal, L. R., Moult, J., & Brenner, S. E. (2019). Reports from the fifth edition of CAGI: The critical assessment of genome interpretation. *Human Mutation, 40*, 1197–1201.

Angus, L., Smid, M., Wilting, S. M., Van Riet, J., Van Hoeck, A., Nguyen, L., Nik-Zainal, S., Steenbruggen, T. G., Tjan-Heijnen, V. C. G., Labots, M., Van Riel, J., Bloemendal, H. J., Steeghs, N., Lolkema, M. P., Voest, E. E., Van de Werken, H. J. G., Jager, A., Cuppen, E., Sleijfer, S., & Martens, J. W. M. (2019). The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. *Nature Genetics, 51*, 1450–1458.

Barabási, A. L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: A network-based approach to human disease. *Nature Reviews. Genetics, 12*, 56–68.

Basir, O., Zhu, H., & Karray, F. (2003). Fuzzy based image segmentation. In M. Nachtegael (Ed.), *Fuzzy filters for image processing*. Springer.

Baskin, I. I., Winkler, D., & Tetko, I. V. (2016). A renaissance of neural networks in drug discovery. *Expert Opinion on Drug Discovery, 11*, 785–795.

Batch, K. E., Yue, J., Darcovich, A., Lupton, K., Liu, C. C., Woodlock, D. P., El Amine, M. A. K., Causa-Andrieu, P. I., Gazit, L., Nguyen, G. H., Zulkernine, F., Do, R. K. G., & Simpson, A. L. (2022). Developing a cancer digital twin: Supervised metastases detection from consecutive structured radiology reports. *Frontiers in Artificial Intelligence, 5*, 826402–826402.

Berger, M. F., & Mardis, E. R. (2018). The emerging clinical relevance of genomics in cancer medicine. *Nature Reviews. Clinical Oncology, 15*, 353–365.

Bertini, I., Calabrò, A., De Carli, V., Luchinat, C., Nepi, S., Porfirio, B., Renzi, D., Saccenti, E., & Tenori, L. (2009). The metabonomic signature of celiac disease. *Journal of Proteome Research, 8*, 170–177.

Bhargavi, K., & Jyothi, S. (2014). A survey on threshold based segmentation technique in image processing. *International Journal of Innovative Research & Development, 3*, 234–239.

Bhasin, J. M., Lee, B. H., Matkin, L., Taylor, M. G., Hu, B., Xu, Y., Magi-Galluzzi, C., Klein, E. A., & Ting, A. H. (2015). Methylome-wide sequencing detects DNA hypermethylation distinguishing indolent from aggressive prostate cancer. *Cell Reports, 13*, 2135–2146.

Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., Allison, T., Arnaout, O., Abbosh, C., Dunn, I. F., Mak, R. H., Tamimi, R. M., Tempany, C. M., Swanton, C., Hoffmann, U., Schwartz, L. H., Gillies, R. J., Huang, R. Y., & Aerts, H. (2019). Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA: a Cancer Journal for Clinicians, 69*, 127–157.

Björnsson, B., Borrebaeck, C., Elander, N., Gasslander, T., Gawel, D. R., Gustafsson, M., Jörnsten, R., Lee, E. J., Li, X., Lilja, S., Martínez-Enguita, D., Matussek, A., Sandström, P., Schäfer, S., Stenmarker, M., Sun, X. F., Sysoev, O., Zhang, H., & Benson, M. (2019). Digital twins to personalize medicine. *Genome Medicine, 12*, 4.

Bodalal, Z., Trebeschi, S., Nguyen-Kim, T. D. L., Schats, W., & Beets-Tan, R. (2019). Radiogenomics: Bridging imaging and genomics. *Abdominal Radiology (New York), 44*, 1960–1984.

Bousselham, A., Bouattane, O., Youssfi, M., & Raihani, A. (2019). Towards reinforced brain tumor segmentation on MRI images based on temperature changes on pathologic area. *International Journal of Biomedical Imaging, 3*, 1758948.

Brat, D. J., Verhaak, R. G., Aldape, K. D., Yung, W. K., Salama, S. R., Cooper, L. A., Rheinbay, E., Miller, C. R., Vitucci, M., Morozova, O., Robertson, A. G., Noushmehr, H., Laird, P. W., Cherniack, A. D., Akbani, R., Huse, J. T., Ciriello, G., Poisson, L. M., Barnholtz-Sloan, J. S., Berger, M. S., Brennan, C., Colen, R. R., Colman, H., Flanders, A. E., Giannini, C., Grifford, M., Iavarone, A., Jain, R., Joseph, I., Kim, J., Kasaian, K., Mikkelsen, T., Murray, B. A., O'Neill, B. P., Pachter, L., Parsons, D. W., Sougnez, C., Sulman, E. P., Vandenberg, S. R., Van Meir, E. G., Von Deimling, A., Zhang, H., Crain, D., Lau, K., Mallery, D., Morris, S., Paulauskis, J., Penny, R., Shelton, T., Sherman, M., Yena, P., Black, A., Bowen, J., Dicostanzo, K., Gastier-Foster, J., Leraas, K. M., Lichtenberg, T. M., Pierson, C. R., Ramirez, N. C., Taylor, C., Weaver, S., Wise, L., Zmuda, E., Davidsen, T., Demchok, J. A., Eley, G., Ferguson, M. L., Hutter, C. M., Mills Shaw, K. R., Ozenberger, B. A., Sheth, M., Sofia, H. J., Tarnuzzer, R., Wang, Z., Yang, L., Zenklusen, J. C., Ayala, B., Baboud, J., Chudamani, S., Jensen, M. A., Liu, J., Pihl, T., Raman, R., Wan, Y., Wu, Y., Ally, A., Auman, J. T., Balasundaram, M., Balu, S., Baylin, S. B., Beroukhim, R., Bootwalla, M. S., Bowlby, R., Bristow, C. A., Brooks, D., Butterfield, Y., Carlsen, R., Carter, S., Chin, L., Chu, A., et al. (2015). Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *The New England Journal of Medicine, 372*, 2481–2498.

Cammarota, G., Ianiro, G., Ahern, A., Carbone, C., Temko, A., Claesson, M. J., & gasbarrini, A. & Tortora, G. (2020). Gut microbiome, big data and machine learning to promote precision medicine for cancer. *Nature Reviews Gastroenterology & Hepatology, 17*, 635–648.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 40*, 834–848.

Cheng, D., Knox, C., Young, N., Stothard, P., Damaraju, S., & Wishart, D. S. (2008). PolySearch: A web-based text mining system for extracting relationships between human diseases, genes, mutations, drugs and metabolites. *Nucleic Acids Research, 36*, W399–W405.

Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A. L., Razavian, N., & Tsirigos, A. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature Medicine, 24*, 1559–1567.

Dettmer, K., Aronov, P. A., & Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews, 26*, 51–78.

Doerr, A. (2019). Single-cell proteomics. *Nature Methods, 16*, 20.

Duan, K. B., Rajapakse, J. C., Wang, H., & Azuaje, F. (2005). Multiple SVM-RFE for gene selection in cancer classification with expression data. *IEEE Transactions on Nanobioscience, 4*, 228–234.

Dupond, S. (2019). A thorough review on the current advance of neural network structures. *Annual Reviews in Control, 14*, 200–230.

Echle, A., Rindtorff, N. T., Brinker, T. J., Luedde, T., Pearson, A. T., & Kather, J. N. (2021). Deep learning in cancer pathology: A new generation of clinical biomarkers. *British Journal of Cancer, 124*, 686–696.

Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., Böhm, A., Deubner, J., Jäckel, Z., Seiwald, K., Dovzhenko, A., Tietz, O., Dal Bosco, C., Walsh, S., Saltukoglu, D., Tay, T. L., Prinz, M., Palme, K., Simons, M., Diester, I., Brox, T., & Ronneberger, O. (2019). U-Net: Deep learning for cell counting, detection, and morphometry. *Nature Methods, 16*, 67–70.

Farag, A. A. (1992). Edge-based image segmentation. *Remote Sensing Reviews, 6*, 95–121.

Faust, K., Sathirapongsasuti, J. F., Izard, J., Segata, N., Gevers, D., Raes, J., & Huttenhower, C. (2012). Microbial co-occurrence relationships in the human microbiome. *PLoS Computational Biology, 8*, e1002606.

Fleming, N. (2018). How artificial intelligence is changing drug discovery. *Nature, 557*, S55–s57.

Forbes, S. A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., Ward, S., Kok, C. Y., Jia, M., De, T., Teague, J. W., Stratton, M. R., Mcdermott, U., & Campbell, P. J. (2015). COSMIC: Exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Research, 43*, D805–D811.

Friedman, J., & Alm, E. J. (2012). Inferring correlation networks from genomic survey data. *PLoS Computational Biology, 8*, e1002687.

Fujishima, H., Fumoto, S., Shibata, T., Nishiki, K., Tsukamoto, Y., Etoh, T., Moriyama, M., Shiraishi, N., & Inomata, M. (2017). A 17-molecule set as a predictor of complete response to neoadjuvant chemotherapy with docetaxel, cisplatin, and 5-fluorouracil in esophageal cancer. *PLoS One, 12*, e0188098.

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., Cerami, E., Sander, C., & Schultz, N. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science Signaling, 6*, 11.

García Márquez, F. P., Segovia Ramírez, I., & Pliego Marugán, A. (2019). Decision making using logical decision tree and binary decision diagrams: A real case study of wind turbine manufacturing. *Energies, 12*, 1753.

Gawehn, E., Hiss, J. A., & Schneider, G. (2016). Deep learning in drug discovery. *Molecular Informatics, 35*, 3–14.

Goldman, M., Craft, B., Swatloski, T., Ellrott, K., Cline, M., Diekhans, M., Ma, S., Wilks, C., Stuart, J., Haussler, D., & Zhu, J. (2013). The UCSC cancer genomics browser: Update 2013. *Nucleic Acids Research, 41*, D949–D954.

Govindarajan, R., Duraiyan, J., Kaliyappan, K., & Palanisamy, M. (2012). Microarray and its applications. *Journal of Pharmacy & Bioallied Sciences, 4*, S310–S312.

Guo, Z., Liu, H., Ni, H., Wang, X., Su, M., Guo, W., Wang, K., Jiang, T., & Qian, Y. J. S. R. (2019). A fast and refined cancer regions segmentation framework in whole-slide breast pathological images. *Scientific Reports, 9*, 1–10.

Hakenberg, J., Voronov, D., Nguyên, V. H., Liang, S., Anwar, S., Lumpkin, B., Leaman, R., Tari, L., & Baral, C. (2012). A SNPshot of PubMed to associate genetic variants with drugs, diseases, and adverse reactions. *Journal of Biomedical Informatics, 45*, 842–850.

He, X., Chang, S., Zhang, J., Zhao, Q., Xiang, H., Kusonmano, K., Yang, L., Sun, Z. S., Yang, H., & Wang, J. (2008). MethyCancer: The database of human DNA methylation and cancer. *Nucleic Acids Research, 36*, D836–D841.

Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition* (pp. 278–282). IEEE.

Hsu, E. R., Klemm, J. D., Kerlavage, A. R., Kusnezov, D., & Kibbe, W. A. (2017). Cancer Moonshot data and technology team: enabling a national learning healthcare system for cancer to unleash the power of data. *Clinical Pharmacology and Therapeutics, 101*, 613–615.

Huang, S., Yang, J., Fong, S., & Zhao, Q. (2020). Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. *Cancer Letters, 471*, 61–71.

Huber, K., Feuchtinger, A., Borgmann, D. M., Li, Z., Aichler, M., Hauck, S. M., Zitzelsberger, H., Schwaiger, M., Keller, U., & Walch, A. (2014). Novel approach of MALDI drug imaging, immunohistochemistry, and digital image analysis for drug distribution studies in tissues. *Analytical Chemistry, 86*, 10568–10575.

Hull, R., Lolas, G., Makrogkikas, S., Jensen, L. D., Syrigos, K. N., Evangelou, G., Padayachy, L., Egbor, C., Mehrotra, R., Makhafola, T. J., Oyomno, M., & Dlamini, Z. (2021). Microbiomics in collusion with the nervous system in carcinogenesis: Diagnosis, pathogenesis and treatment. *Microorganisms, 9*, 2129.

Hussein, S., Green, A., Watane, A., Reiter, D., Chen, X., Papadakis, G. Z., Wood, B., Cypess, A., Osman, M., & Bagci, U. (2017). Automatic segmentation and quantification of white and Brown adipose tissues from PET/CT scans. *IEEE Transactions on Medical Imaging, 36*, 734–744.

Ibrahim, A., Gamble, P., Jaroensri, R., Abdelsamea, M. M., Mermel, C. H., Chen, P. C., & Rakha, E. A. (2020). Artificial intelligence in digital breast pathology: Techniques and applications. *Breast, 49*, 267–273.

Ikeda, A., Nishiumi, S., Shinohara, M., Yoshie, T., Hatano, N., Okuno, T., Bamba, T., Fukusaki, E., Takenawa, T., Azuma, T., & Yoshida, M. (2012). Serum metabolomics as a novel diagnostic approach for gastrointestinal cancer. *Biomedical Chromatography, 26*, 548–558.

Inglese, J., Auld, D. S., Jadhav, A., Johnson, R. L., Simeonov, A., Yasgar, A., Zheng, W., & Austin, C. P. (2006). Quantitative high-throughput screening: A titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 11473–11478.

Jansen, R. W., Van Amstel, P., Martens, R. M., Kooi, I. E., Wesseling, P., De Langen, A. J., Der Houven, M.-V., Van Oordt, C. W., Jansen, B. H. E., Moll, A. C., Dorsman, J. C., Castelijns, J. A., De Graaf, P., & De Jong, M. C. (2018). Non-invasive tumor genotyping using radiogenomic biomarkers, a systematic review and oncology-wide pathway analysis. *Oncotarget, 9*, 20134–20155.

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology, 2*, 230–243.

Joshi, S., Vibhute, G., Ayachit, A., & Ayachit, G. (2021). Big data and artificial intelligence - tools to be future ready? *Indian Journal of Ophthalmology, 69*, 1652–1653.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research, 4*, 237–285.

Kamiński, B., Jakubczyk, M., & Szufel, P. (2018). A framework for sensitivity analysis of decision trees. *Central European Journal of Operations Research, 26*, 135–159.

Kaur, J., Rahat, B., Thakur, S., Verma, M., & Barh, D. (2017). *Progress and challenges in precision medicine*. Academic Press.

Kehl, K. L., Xu, W., Gusev, A., Bakouny, Z., Choueiri, T. K., Riaz, I. B., Elmarakeby, H., Van Allen, E. M., & Schrag, D. (2021). Artificial intelligence-aided clinical annotation of a large multi-cancer genomic dataset. *Nature Communications, 12*, 7304.

Kelley, D. Z., Flam, E. L., Izumchenko, E., Danilova, L. V., Wulf, H. A., Guo, T., Singman, D. A., Afsari, B., Skaist, A. M., Considine, M., Welch, J. A., Stavrovskaya, E., Bishop, J. A., Westra, W. H., Khan, Z., Koch, W. M., Sidransky, D., Wheelan, S. J., Califano, J. A., Favorov, A. V., Fertig, E. J., & Gaykalova, D. A. (2017). Integrated analysis of whole-genome ChIP-Seq and RNA-Seq data of primary head and neck tumor samples associates HPV integration sites with open chromatin marks. *Cancer Research, 77*, 6538–6550.

Kelly, R. S., Vander Heiden, M. G., Giovannucci, E., & Mucci, L. A. (2016). Metabolomic biomarkers of prostate cancer: Prediction, diagnosis, progression, prognosis, and recurrence. *Cancer Epidemiology, Biomarkers & Prevention, 25*, 887–906.

Keyl, P., Bockmayr, M., Heim, D., Dernbach, G., Montavon, G., Müller, K. R., & Klauschen, F. (2022). Patient-level proteomic network prediction by explainable artificial intelligence. *NPJ Precision Oncology, 6*, 35.

Kulak, N. A., Geyer, P. E., & Mann, M. (2017). Loss-less Nano-fractionator for high sensitivity, high coverage proteomics. *Molecular & Cellular Proteomics, 16*, 694–705.

Kumar, R., Sharma, A., Siddiqui, M. H., & Tiwari, R. K. (2017). Prediction of human intestinal absorption of compounds using artificial intelligence techniques. *Current Drug Discovery Technologies, 14*, 244–254.

Kuwahara, T., Hara, K., Mizuno, N., Haba, S., Okuno, N., Koda, H., Miyano, A., & Fumihara, D. (2021). Current status of artificial intelligence analysis for endoscopic ultrasonography. *Digestive Endoscopy, 33*, 298–305.

Li, L., Zheng, X., Zhou, Q., Villanueva, N., Nian, W., Liu, X., & Huan, T. (2020). Metabolomics-based discovery of molecular signatures for triple negative breast cancer in Asian female population. *Scientific Reports, 10*, 370.

Lin, X., Wang, Q., Yin, P., Tang, L., Tan, Y., Li, H., Yan, K., & Xu, G. J. M. (2011). A method for handling metabonomics data from liquid chromatography/mass spectrometry: Combinational use of support vector machine recursive feature elimination, genetic algorithm and random forest for feature selection. *Metabolomics, 7*, 549–558.

Liu, Y., Gonzàlez-Porta, M., Santos, S., Brazma, A., Marioni, J. C., Aebersold, R., Venkitaraman, A. R., & Wickramasinghe, V. O. (2017). Impact of alternative splicing on the human proteome. *Cell Reports, 20*, 1229–1241.

López de Maturana, E., Alonso, L., Alarcón, P., Martín-Antoniano, I. A., Pineda, S., Piorno, L., Calle, M. L., & Malats, N. (2019). Challenges in the integration of omics and non-omics data. *Genes (Basel), 10*, 238.

Magadza, T., & Viriri, S. (2021). Deep learning for brain tumor segmentation: a survey of state-of-the-art. *Journal of Imaging, 7*, 19.

Mangani, D., Weller, M., & Roth, P. (2017). The network of immunosuppressive pathways in glioblastoma. *Biochemical Pharmacology, 130*, 1–9.

Meng, C., Zeleznik, O. A., Thallinger, G. G., Kuster, B., Gholami, A. M., & Culhane, A. C. (2016). Dimension reduction techniques for the integrative analysis of multi-omics data. *Briefings in Bioinformatics, 17*, 628–641.

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., & Wiest, R. (2014). The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Transactions on Medical Imaging, 34*, 1993–2024.

Mertins, P., Mani, D. R., Ruggles, K. V., Gillette, M. A., Clauser, K. R., Wang, P., Wang, X., Qiao, J. W., Cao, S., Petralia, F., Kawaler, E., Mundt, F., Krug, K., Tu, Z., Lei, J. T., Gatza, M. L., Wilkerson, M., Perou, C. M., Yellapantula, V., Huang, K. L., Lin, C., Mclellan, M. D., Yan, P., Davies, S. R., Townsend, R. R., Skates, S. J., Wang, J., Zhang, B., Kinsinger, C. R., Mesri, M.,

Rodriguez, H., Ding, L., Paulovich, A. G., Fenyö, D., Ellis, M. J., & Carr, S. A. (2016). Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature, 534*, 55–62.

Merz, A. L., & Serkova, N. J. (2009). Use of nuclear magnetic resonance-based metabolomics in detecting drug resistance in cancer. *Biomarkers in Medicine, 3*, 289–306.

Morin, O., Vallières, M., Braunstein, S., Ginart, J. B., Upadhaya, T., Woodruff, H. C., Zwanenburg, A., Chatterjee, A., Villanueva-Meyer, J. E., Valdes, G., Chen, W., Hong, J. C., Yom, S. S., Solberg, T. D., Löck, S., Seuntjens, J., Park, C., & Lambin, P. (2021). An artificial intelligence framework integrating longitudinal electronic health records with real-world data enables continuous pan-cancer prognostication. *Nature Cancer, 2*, 709–722.

Muhammad, K., Khan, S., Del Ser, J., & De Albuquerque, V. H. C. (2020). Deep learning for multigrade brain tumor classification in smart healthcare systems: A prospective survey. *IEEE Transactions on Neural Networks and Learning Systems, 32*, 507–522.

O'shea, R. J., Sharkey, A. R., Cook, G. J. R., & Goh, V. (2021). Systematic review of research design and reporting of imaging studies applying convolutional neural networks for radiological cancer diagnosis. *European Radiology, 31*, 7969–7983.

Öztürk, H., Özgür, A., & Ozkirimli, E. (2018). DeepDTA: Deep drug-target binding affinity prediction. *Bioinformatics, 34*, i821–i829.

Pan, Y. (2007). *Image segmentation using PDE, variational, morphological and probabilistic methods*. University of Tennessee.

Paolillo, C., Londin, E., & Fortina, P. (2016). Next generation sequencing in cancer: Opportunities and challenges for precision cancer medicine. *Scandinavian Journal of Clinical and Laboratory Investigation. Supplementum, 245*, S84–S91.

Park, J. W., Jeong, J. M., Cho, K. S., Cho, S. Y., Cheon, J. H., Choi, D. H., Park, S. J., & Kim, H. K. (2021). MiR-30a and miR-200c differentiate cholangiocarcinomas from gastrointestinal cancer liver metastases. *PLoS One, 16*, e0250083.

Pezoulas, V. C., Hazapis, O., Lagopati, N., Exarchos, T. P., Goules, A. V., Tzioufas, A. G., Fotiadis, D. I., Stratis, I. G., Yannacopoulos, A. N., & Gorgoulis, V. G. (2021). Machine learning approaches on high throughput NGS data to unveil mechanisms of function in biology and disease. *Cancer Genomics Proteomics, 18*, 605–626.

Pös, O., Biró, O., Szemes, T., & Nagy, B. (2018). Circulating cell-free nucleic acids: Characteristics and applications. *European Journal of Human Genetics, 26*, 937–945.

Punitha, S., Amuthan, A., & Joseph, K. S. (2018). Benign and malignant breast cancer segmentation using optimized region growing technique. *Future Computing and Informatics Journal, 3*, 348–358.

Robinson, D., Van Allen, E. M., Wu, Y. M., Schultz, N., Lonigro, R. J., Mosquera, J. M., Montgomery, B., Taplin, M. E., Pritchard, C. C., Attard, G., Beltran, H., Abida, W., Bradley, R. K., Vinson, J., Cao, X., Vats, P., Kunju, L. P., Hussain, M., Feng, F. Y., Tomlins, S. A., Cooney, K. A., Smith, D. C., Brennan, C., Siddiqui, J., Mehra, R., Chen, Y., Rathkopf, D. E., Morris, M. J., Solomon, S. B., Durack, J. C., Reuter, V. E., Gopalan, A., Gao, J., Loda, M., Lis, R. T., Bowden, M., Balk, S. P., Gaviola, G., Sougnez, C., Gupta, M., Yu, E. Y., Mostaghel, E. A., Cheng, H. H., Mulcahy, H., True, L. D., Plymate, S. R., Dvinge, H., Ferraldeschi, R., Flohr, P., Miranda, S., Zafeiriou, Z., Tunariu, N., Mateo, J., Perez-Lopez, R., Demichelis, F., Robinson, B. D., Schiffman, M., Nanus, D. M., Tagawa, S. T., Sigaras, A., Eng, K. W., Elemento, O., Sboner, A., Heath, E. I., Scher, H. I., Pienta, K. J., Kantoff, P., De Bono, J. S., Rubin, M. A., Nelson, P. S., Garraway, L. A., Sawyers, C. L., & Chinnaiyan, A. M. (2015). Integrative clinical genomics of advanced prostate cancer. *Cell, 161*, 1215–1228.

Robinson, D. R., Wu, Y. M., Lonigro, R. J., Vats, P., Cobain, E., Everett, J., Cao, X., Rabban, E., Kumar-Sinha, C., Raymond, V., Schuetze, S., Alva, A., Siddiqui, J., Chugh, R., Worden, F., Zalupski, M. M., Innis, J., Mody, R. J., Tomlins, S. A., Lucas, D., Baker, L. H., Ramnath, N., Schott, A. F., Hayes, D. F., Vijai, J., Offit, K., Stoffel, E. M., Roberts, J. S., Smith, D. C., Kunju, L. P., Talpaz, M., Cieślik, M., & Chinnaiyan, A. M. (2017). Integrative clinical genomics of metastatic cancer. *Nature, 548*, 297–303.

Romo-Bucheli, D., Janowczyk, A., Gilmore, H., Romero, E., & Madabhushi, A. (2017). A deep learning based strategy for identifying and associating mitotic activity with gene expression derived risk categories in estrogen receptor positive breast cancers. *Cytometry. Part A, 91*, 566–573.

Roy, K., & Pratim Roy, P. (2009). Comparative chemometric modeling of cytochrome 3A4 inhibitory activity of structurally diverse compounds using stepwise MLR, FA-MLR, PLS, GFA, G/PLS and ANN techniques. *European Journal of Medicinal Chemistry, 44*, 2913–2922.

Rutman, A. M., & Kuo, M. D. (2009). Radiogenomics: Creating a link between molecular diagnostics and diagnostic imaging. *European Journal of Radiology, 70*, 232–241.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution, 4*, 406–425.

Saltz, J., Gupta, R., Hou, L., Kurc, T., Singh, P., Nguyen, V., Samaras, D., Shroyer, K. R., Zhao, T., Batiste, R., Van Arnam, J., Shmulevich, I., Rao, A. U. K., Lazar, A. J., Sharma, A., & Thorsson, V. (2018). Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Reports, 23*, 181–193.e7.

Santus, E., Li, C., Yala, A., Peck, D., Soomro, R., Faridi, N., Mamshad, I., Tang, R., Lanahan, C. R., Barzilay, R., & Hughes, K. (2019). Do neural information extraction algorithms generalize across institutions? *JCO Clinical Cancer Informatics, 3*, 1–8.

Schoof, E. M., Furtwängler, B., Üresin, N., Rapin, N., Savickas, S., Gentil, C., Lechman, E., Auf Dem Keller, U., Dick, J. E., & Porse, B. T. (2021a). Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nature Communications, 12*, 745679.

Schoof, E. M., Furtwängler, B., Üresin, N., Rapin, N., Savickas, S., Gentil, C., Lechman, E., Auf Dem Keller, U., Dick, J. E., & Porse, B. T. (2021b). Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *bioRxiv, 12*, 745679.

Sender, R., Fuchs, S., & Milo, R. (2016). Revised estimates for the number of human and bacteria cells in the body. *PLoS Biology, 14*, e1002533.

Sepich-Poore, G. D., Zitvogel, L., Straussman, R., Hasty, J., Wargo, J. A., & Knight, R. (2021). The microbiome and human cancer. *Science, 371*, eabc4552.

Sirinukunwattana, K., Ahmed Raza, S. E., Yee-Wah, T., Snead, D. R., Cree, I. A., & Rajpoot, N. M. (2016). Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Transactions on Medical Imaging, 35*, 1196–1206.

Sliž, J. & Mikulka, J. (2016). Advanced image segmentation methods using partial differential equations: A concise comparison. In *2016 Progress in Electromagnetic Research Symposium (PIERS)* (pp. 1809–1812). IEEE.

Thierry, A. R., El Messaoudi, S., Gahan, P. B., Anker, P., & Stroun, M. (2016). Origins, structures, and functions of circulating DNA in oncology. *Cancer Metastasis Reviews, 35*, 347–376.

Trivizakis, E., Manikis, G. C., Nikiforaki, K., Drevelegas, K., Constantinides, M., Drevelegas, A., & Marias, K. (2019). Extending 2-D convolutional neural networks to 3-D for advancing deep learning cancer classification with application to MRI liver tumor differentiation. *IEEE Journal of Biomedical and Health Informatics, 23*, 923–930.

Varambally, S., Dhanasekaran, S. M., Zhou, M., Barrette, T. R., Kumar-Sinha, C., Sanda, M. G., Ghosh, D., Pienta, K. J., Sewalt, R. G., Otte, A. P., Rubin, M. A., & Chinnaiyan, A. M. (2002). The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature, 419*, 624–629.

Veselkov, K. A., Vingara, L. K., Masson, P., Robinette, S. L., Want, E., Li, J. V., Barton, R. H., Boursier-Neyret, C., Walther, B., Ebbels, T. M., Pelczer, I., Holmes, E., Lindon, J. C., & Nicholson, J. K. (2011). Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery. *Analytical Chemistry, 83*, 5864–5872.

Vita, G. M., De Simone, G., De Marinis, E., Nervi, C., Ascenzi, P., & Di Masi, A. (2022). *Serum albumin and nucleic acids biodistribution: From molecular aspects to biotechnological applications*. IUBMB Life.

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen, D. D., Watrous, J., Kapono, C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W. T., Crüsemann, M., Boudreau, P. D., Esquenazi, E., Sandoval-Calderón, M., Kersten, R. D., Pace, L. A., Quinn, R. A., Duncan, K. R., Hsu, C. C., Floros, D. J., Gavilan, R. G., Kleigrewe, K., Northen, T., Dutton, R. J., Parrot, D., Carlson, E. E., Aigle, B., Michelsen, C. F., Jelsbak, L., Sohlenkamp, C., Pevzner, P., Edlund, A., Mclean, J., Piel, J., Murphy, B. T., Gerwick, L., Liaw, C. C., Yang, Y. L., Humpf, H. U., Maansson, M., Keyzers, R. A., Sims, A. C., Johnson, A. R., Sidebottom, A. M., Sedio, B. E., Klitgaard, A., Larson, C. B., Torres-Mendoza, D., Gonzalez, D. J., Silva, D. B., Marques, L. M., Demarque, D. P., Pociute, E., O'Neill, E. C., Briand, E., Helfrich, E. J. N., Granatosky, E. A., Glukhov, E., Ryffel, F., Houson, H., Mohimani, H., Kharbush, J. J., Zeng, Y., Vorholt, J. A., Kurita, K. L., Charusanti, P., Mcphail, K. L., Nielsen, K. F., Vuong, L., Elfeki, M., Traxler, M. F., Engene, N., Koyama, N., Vining, O. B., Baric, R., Silva, R. R., Mascuch, S. J., Tomasi, S., Jenkins, S., Macherla, V., Hoffman, T., Agarwal, V., Williams, P. G., Dai, J., Neupane, R., Gurr, J., Rodríguez, A. M. C., Lamsa, A., Zhang, C., Dorrestein, K., Duggan, B. M., Almaliti, J., Allard, P. M., Phapale, P., et al. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nature Biotechnology, 34*, 828–837.

Wang, Z., Jensen, M. A., & Zenklusen, J. C. (2016). A practical guide to the cancer genome atlas (TCGA). *Methods in Molecular Biology, 1418*, 111–141.

Waring, M. J., Arrowsmith, J., Leach, A. R., Leeson, P. D., Mandrell, S., Owen, R. M., Pairaudeau, G., Pennie, W. D., Pickett, S. D., Wang, J., Wallace, O., & Weir, A. (2015). An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nature Reviews. Drug Discovery, 14*, 475–486.

Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., & Madabhushi, A. (2016). Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. *IEEE Transactions on Medical Imaging, 35*, 119–130.

Yamazaki, Y. (2015). Metabolome analysis of human serum: Implications for early detection of colorectal cancer. *Rinsho Byori, 63*, 328–335.

Ypsilantis, P. P., Siddique, M., Sohn, H. M., Davies, A., Cook, G., Goh, V., & Montana, G. (2015). Predicting response to neoadjuvant chemotherapy with PET imaging using convolutional neural networks. *PLoS One, 10*, e0137036.

Yu, K. H., Zhang, C., Berry, G. J., Altman, R. B., Ré, C., Rubin, D. L., & Snyder, M. (2016). Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nature Communications, 7*, 12474.

Yuan, Y., Failmezger, H., Rueda, O. M., Ali, H. R., Gräf, S., Chin, S. F., Schwarz, R. F., Curtis, C., Dunning, M. J., Bardwell, H., Johnson, N., Doyle, S., Turashvili, G., Provenzano, E., Aparicio, S., Caldas, C., & Markowetz, F. (2012). Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Science Translational Medicine, 4*, 157ra143.

Zang, Q., Mansouri, K., Williams, A. J., Judson, R. S., Allen, D. G., Casey, W. M., & Kleinstreuer, N. C. (2017). In silico prediction of physicochemical properties of environmental chemicals using molecular fingerprints and machine learning. *Journal of Chemical Information and Modeling, 57*, 36–49.

Zhang, J., Bajari, R., Andric, D., Gerthoffert, F., Lepsa, A., Nahal-Bose, H., Stein, L. D., & Ferretti, V. (2019). The international cancer genome consortium data portal. *Nature Biotechnology, 37*, 367–369.

Zhang, L., Jin, H., Guo, X., Yang, Z., Zhao, L., Tang, S., Mo, P., Wu, K., Nie, Y., Pan, Y., & Fan, D. (2012). Distinguishing pancreatic cancer from chronic pancreatitis and healthy individuals by (1)H nuclear magnetic resonance-based metabonomic profiles. *Clinical Biochemistry, 45*, 1064–1069.

Zhang, T., Wu, X., Ke, C., Yin, M., Li, Z., Fan, L., Zhang, W., Zhang, H., Zhao, F., Zhou, X., Lou, G., & Li, K. (2013). Identification of potential biomarkers for ovarian cancer by urinary metabolomic profiling. *Journal of Proteome Research, 12*, 505–512.

Zhao, L., Wang, W., Sedykh, A., & ZHU, H. (2017). Experimental errors in QSAR Modeling sets: What we can do and what we cannot do. *ACS Omega, 2*, 2805–2812.

Zheng, G., Han, G., Soomro, N. Q., Ma, L., Zhang, F., Zhao, Y., Zhao, X., & Zhou, C. (2019). A novel computer-aided diagnosis scheme on small annotated set: G2C-CAD. *BioMed Research International, 2019*, 6425963.

Zhou, X., Menche, J., Barabási, A. L., & Sharma, A. (2014). Human symptoms-disease network. *Nature Communications, 5*, 4212.

Zhu, X. W., Xin, Y. J., & Chen, Q. H. (2016). Chemical and in vitro biological information to predict mouse liver toxicity using recursive random forests. *SAR and QSAR in Environmental Research, 27*, 559–572.

Zhuang, J., Tang, X., Du, Z., Yang, M., & Zhou, Y. (2016). Prediction of biomarkers of therapeutic effects of patients with lung adenocarcinoma treated with gefitinib based on progression-free-survival by metabolomic fingerprinting. *Talanta, 160*, 636–644.

Part I
Artificial Intelligence for Screening, Diagnosis, Monitoring in Precision Oncology

Chapter 2
Application of AI in Novel Biomarkers Detection that Induces Drug Resistance, Enhance Treatment Regimens, and Advancing Precision Oncology

Rahaba Marima, Nompumelelo Mtshali, Kgomotso Mathabe, Afra Basera, Mahlori Mkhabele, Meshack Bida, and Zodwa Dlamini

R. Marima (✉) · M. Mkhabele · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: rahaba.marima@up.ac.za; u13029003@tuks.co.za; zodwa.dlamini@up.ac.za

N. Mtshali
Department of Anatomical Pathology, Faculty of Health Sciences, University of the Witwatersrand, Parktown, South Africa
e-mail: nompumelelo.mtshali@wits.ac.za

K. Mathabe
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Department of Urology, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: kgomotso.mathabe@up.ac.za

A. Basera
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Department of Medical Oncology, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: u22800370@tuks.co.za

M. Bida
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Department of Anatomical Pathology, National Health Laboratory Service (NHLS), University of Pretoria, Pretoria, South Africa
e-mail: meshack.bida@nhls.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_2

Abstract Artificial intelligence (AI) is changing the medical research and patient care field by showing data patterns that allow predicting disease, disease progress, and treatment outcomes for individual patients. Big-data sets from these fields require advanced technology for analysis. High cancer mortality negates advances in oncology research. Traditional approaches are becoming inadequate to efficiently combat cancer due to cancer's heterogenous nature. Accurate risk assessment, prevention, detection, segmentation, and cancer treatment present major challenges for successful patient outcomes. AI-based tool advancement presents a potent weapon for improved cancer care by advancing personalized patient care. These tools have promise for improved therapeutic potential and identifying novel biomarkers and drug targets. Effective implementation of precision oncology needs a positive impact on patient outcome, provides decision support in real time, and discovery of unique patient patterns of disease progression. Emerging technologies present with new challenges; the benefits of AI technology in precision oncology outweigh the challenges. AI-based precision oncology provides augmented intelligence to aid clinician decision-making. Advancement of wet-lab-based assays, high throughput NGS data, bioinformatics tools, and strategies to detect novel biomarkers that accurately predict prognosis and enhance treatment regimens are urgently warranted. This review will focus on AI-based tools in the detection and identification of cancer biomarkers for accurate prognosis with the overall aim of enhancing treatment regimens, advancing precision oncology, and improving patient outcomes.

Keywords Artificial Intelligence (AI) · Biomarkers · Precision oncology · Cancer treatment · Tissue biopsies · Liquid biopsies · DNA methylation

2.1 Introduction

Artificial intelligence (AI) uses in oncology include cancer research, prognosis indication, and treatment response. These applications are corroborated by AI-based tools in understanding the molecular biology of tumors (Farina et al., 2022). With mounting data from cancer research such as the OMICS data, it has become imperative to couple these advances with high technology tools such as AI tools. AI relies on computers following algorithms learned by computer methods or even established by humans to execute certain tasks or support decision-making (Hosny et al., 2018). Machine learning (ML) is a branch of AI (Rajkomar et al., 2019). Deep learning (DL) is a subfield of ML, where mathematical algorithms are arrayed using computational units that are multi-layered, resembling human cognizance. These include differentiated neuronal networks (DNN), recurrent neuronal networks (RNN), convolutional neuronal networks (CNNs), and artificial neuronal networks (ANNs).

In unstructured medical data, artificial neuronal networks (ANNs) can be used to analyze this type of data. Unstructured data is common in health systems and is used to record qualitative and personal patient information, which may be obtained by imaging acquisition or patient–health care provider interactions (Wang et al., 2019).

On the other hand, CNNs are DL algorithms suitably used in imaging files. In addition, DL can be used in the prediction of future health risks. For example, in cardiology, when patients do routine CT scans and MRI, DL models have already been trained to predict cardiovascular scores from CT scans (Elton et al., 2021; Pickhardt et al., 2020). Furthermore, it has been shown that DL CNNs can predict a five-year future breast cancer risk, from a normal mammogram. The ability to predict future cancer development from normal scans is a potent intervention tool. Additionally, CNNs have been demonstrated to show over 90% precision in discriminating between benign from malignant hematoxylin and eosin (H and E) stained breast biopsies (Cai et al., 2010). Similarly, malignant or benign skin lesions can be accurately classified by a dermoscope, maintaining a trained dermatologist standard (Rezvantalab et al., 2018). The identification of change in disease patterns by AI tools holds great promise in advancing precision oncology. Biomarkers are measurable biological indicators used to predict disease initiation and progression (Chen et al., 2015; Lin et al., 2019; Strimbu & Tavel, 2010). Some of the molecules such as nucleic acids, proteins, lipids, and other metabolites are shed into peripheral blood, and these can be used as biomarkers for cancer screenings (Fang, 2020). Although DP has been the cornerstone of cancer biomarker discovery, the use of liquid biopsies in cancer biomarker discovery is emerging as a potent tool to accurately predict prognosis and response to treatment. This review will discuss the AI-based tools in the detection and identification of cancer biomarkers for accurate prognosis with the overall aim of enhancing treatment regimens, advancing precision oncology, and improving patient outcomes.

2.2 AI Advances in Healthcare and Precision Oncology

The ability to analyze complex and comprehensive patient information for the monitoring and differentiating between healthy and sick people relies on the identification of population and personalized biomarkers. This will aid in comprehending biological signals that can indicate any health shifts. It has also been documented that harnessing the use of electronic health records by the integration of distinct data sources and by discovering patient-specific disease progression patterns will facilitate the implementation of effective personalized health care and positively enhance patient outcome. Advanced technological tools are needed to enhance intraoperability, interoperability and networking of laboratory, clinical and public health systems. These tools will also aid to address valid social and ethical issues associated with the protection and privacy of healthcare data. However, data hacking and breaching, uncertainty of black boxes use of algorithms to resolve output pose as limitations and obstacles to efficient AI implementation in health (Ahmed et al., 2020).

AI tools also have the potential to redress the burdening effects of medical errors, which has been reported to be the third leading cause of death. It has been reported that in the USA, approximately 200,000 people are dying every year due to medical

errors (Makary & Daniel, 2016). Unfortunately, a decline in this number has not been reported due to poorly coordinated healthcare and costs, communication breakdown, and misdiagnosis. The concept of precision medicine primarily relies on the 4Ps. These are Predictive, Preventative, Personalized, and Participatory. This is aimed at enabling clinicians to effectively understand how variations in personalized clinical data can positively contribute to improved patient outcomes through precise predictive parameters, accurate diagnosis, and prognosis (Mayekar & Bivona, 2017). The significance of healthcare data mining cannot be ignored.

AI tools have been reported to play a significant role in precision oncology. This is evidenced by the facilitation of early diagnosis of various cancer types which usually present at advanced stages such as epithelial ovarian cancer (EOC). Due to the lack of clinical symptoms in EOC early stages and deficiency of effective screening tests, about 70% of EOC patients are diagnosed at late stages (Jacobs et al., 2016). ML has been demonstrated to hold promise to alleviate this burden (Ma et al., 2021). Ma et al. (2021) used 8 ML techniques to derive predicting information from 11 peripheral blood parameters from EOC patients (Ma et al., 2021). These ML methods included random Forest (RF), Logistic Regression, Gradient Boosting Machine, Naïve Bayes, Conditional Random Forest, Elastic Net, Neural Network and Support Vector Machine. They demonstrated that ML techniques, Random Forest was superior to conventional regression-based classifiers in the prediction of various clinical parameters associated with EOC. This study concluded that ML techniques may provide risk stratification for EOC patients, and this could be done prior to initial intervention. This commendation was particularly referring to the use of blood variables, specifically the circulating tumor cells (CTCs). The predictive algorithms through ML hold a promise to pre-treatment stratification of EOC patients and therefore could facilitate customized treatment.

Risk stratification is essential to improve long-term outcomes of cancer patients, especially cancers with a lack of early detective measures. This AI approach could precisely delineate cancer characteristics and facilitate outcome prediction prior to initial intervention (Narod, 2016). Conventionally, clinical factors such as tumor grade, age, and history may be used to aid in prognosis assessment. However, these parameters have been shown to have inadequate predictive value (Chen et al., 2007; Kang et al., 2012). Contrarily, emerging evidence demonstrates that CTCs in EOC patients hold potential as prognosticators, and these may be applicable to the survival outcomes of various tumors (Aktas et al., 2011). It has been demonstrated that ML can predict early disease and disease progression in EOC patients through liquid biopsies constituents such as the CTCs as biomarkers. Distinct classes of biomarkers are discussed below.

2.3 Classification of Biomarkers

Biomarkers can be classified as type 0, I, and II, according to functions and characteristics (Banerjee et al., 2017; Heckman-Stoddard, 2012; Sahutoglu et al., 2017). Type 0 biomarkers are linked to known clinical indicators and are used to measure the diseases' natural history. On the other hand, type I biomarkers are associated with the efficiency of pharmacological agents. Type II biomarkers are alternate endpoint biomarkers intended for the substitution of clinical endpoints (Sahutoglu et al., 2017; Waseem et al., 2017). Additional to these, oncology biomarkers are further classified into molecular categories which are: genetic, epigenetic, proteins, glycoproteins, receptors, hormones, etc. (Verma, 2012). Tumor biomarkers are also known to be categorized into predictive, diagnostic, prognostic, and pharmacodynamics (Alizadeh et al., 2000; Cai et al., 2010; Chyla et al., 2018; Lesko & Atkinson Jr, 2001; Maisel et al., 2011). Diagnostic biomarkers are used to detect or confirm the presence of a disease. Diagnostic biomarkers may be present at any stage of cancer development. Pharmacodynamics biomarkers are used for the selection of doses of chemotherapeutic agents in a specific set of tumor-patient conditions. Pharmacodynamic biomarkers are also used to assess imminent drug treatment effects. Predictive biomarkers are used to identify subsets of patients who are more likely to respond to a particular treatment, while prognostic biomarkers are used to measure disease progression.

Oncological clinical workflows are dependent on predictive and prognostic molecular biomarkers. However, one of the obstacles with this growing trend of such biomarkers is complexity, cost, and time for clinical decision-making (Echle et al., 2021). Predictive and prognostic molecular biomarkers are currently used in oncology workflows. Predictive biomarkers enable oncologists to choose a specific targeted treatment for a particular patient group. For example, in breast cancer, detecting HER2 positivity in patients makes them eligible for anti-HER2 treatment, thus considering HER2 a strong biomarker in this setting (Le et al., 2017; Naito & Urasaki, 2018). Similarly, with in treatment-refractory stage IV colorectal cancer (CRC), microsatellite instability (MSI) is an FDA-approved biomarker for immunotherapy-based treatment (Le et al., 2017). Here, MSI detection is associated with good therapeutic response prospects, thus rendering MSI a good predictive biomarker. The choice of standard treatment for non-small cell lung cancer (NSCLC) is also driven by the various number of molecular biomarkers such as ALK, PDL1, EGFR, and other genes (Lim et al., 2020). Contrary to predictive biomarkers, prognostic biomarkers enable oncologists to categorize patients based on their disease progression risk and thus can be used for treatment intensity adjustment for individual patients. In clinical trials, the design of new therapeutic agents for solid tumors is progressively coupled with the discovery and identification of predictive biomarkers. MSI is used as a predictive biomarker for immunotherapy, while fusion-driven tumors have a good response to molecularly targeted therapy. Furthermore, homologous repair deficiency tumors such as prostate cancer respond well to inhibitors of poly ADP-ribose polymerase (PARP). With advances in

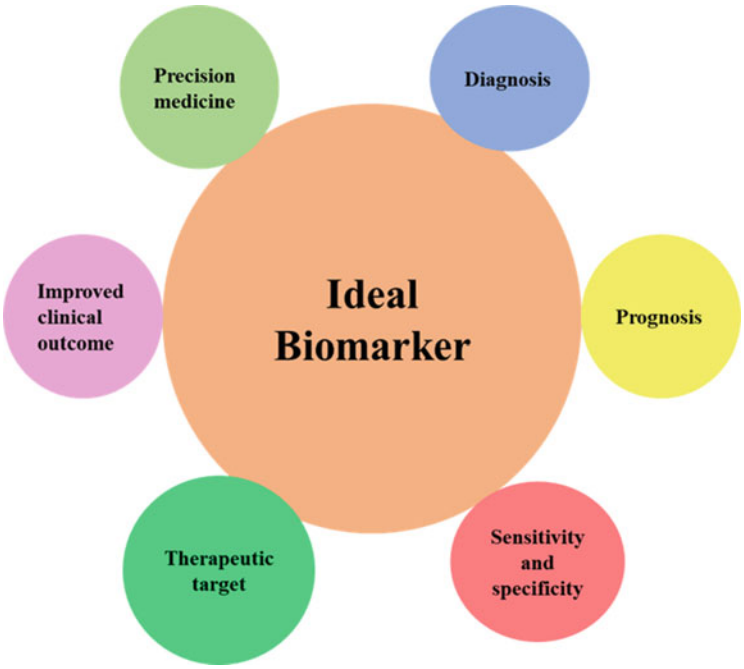


Fig. 2.1 The discovery of ideal biomarkers is key to improved patient outcomes. Biomarkers can be targeted for improved diagnosis, prognosis, and therapeutics. The identification of ideal biomarkers holds the key potential to personalized medicine and overall improved clinical outcome

biomarker detection and discovery, it is imperative to couple these efforts with technology advances such as those of AI, to effectively implement precision oncology. Figure 2.1 illustrates key features of discovering an ideal biomarker, for efficient use in a clinical setting.

2.4 Oncology Biomarkers: Solid Biomarkers vs Liquid Biomarkers

Tissue biopsies have been the gold standard of biomarker discovery. However, some inherent limitations associated with tissue biopsies have been identified. Tissue biopsies can be invasive and clinically risky, usually requiring surgical resection (Ilié & Hofman, 2016). Additionally, tissue biopsies provide spatially limited information obtained from a specific tumor tissue region due to sampling limitations and may therefore not reflect the broad intra-tumor heterogeneity (Gerlinger et al., 2014; Sabaawy, 2013). This may limit accurate diagnosis and prognosis (Hoffman et al., 2002). Alternatively, multi-region tissue biopsies can be used. These biopsies can capture intra-tumor heterogeneity (Gerlinger et al., 2014; Zhang et al., 2014).

However, the clinical application of multi-region tissue biopsies is limited by accessibility and volume of tumor tissues (Ilié & Hofman, 2016; Khan et al., 2018).

Contrarily, liquid biopsies are emerging as minimally invasive, compared to tissue biopsies. Tissue biopsies use circulating blood materials as the CTCs, cell-free nucleic acid (cfNA) (cfDNA-cell-free DNA, cfRNA-cell-free RNA) and exosomes detect and identify molecular modifications that can indicate cancer progression. cfDNA/cfRNA is fragmented cellular DNA/RNA that is released into the bloodstream by dying cells either by programmed cell death (apoptosis) or by necrosis, or even by active secretion (Jahr et al., 2001; Wan et al., 2017). This cellular nucleic fragmentation is a normal physiological process occurring in healthy individuals, where usually cfDNA is mainly derived from apoptotic hematopoietic cells (Lehmann-Werman et al., 2016). Contrarily, in cancer patients, cfDNA/cfRNA can be derived from tumor tissue and tumor microenvironment, usually reflecting the epigenetic and genetic modifications of the tumor microenvironment and tissue (Wyatt et al., 2017). Due to the ability to better reflect the tumor microenvironment than tissue biopsies and even enhanced stability than cfRNA, extensive studies are emerging on cfDNA as a biomarker in early cancer detection, cancer stratification, and cancer surveillance. However, CTCs and cfDNAs have been reported to have short life span (Mondelo-Marcia et al., 2021).

The tumor tissues available from biopsies have an enormous amount of information that is clinically relevant and that remains to be completely exploited. Advances in DL have allowed hidden information from histology cancer images to be mined, thus providing clinically useful information. On the other hand, the evidence of clinical validity and utility for ctDNAs-based assays is currently inadequate for advanced cancers, while lacking for the application of early cancer detection (Merker et al., 2018; Ossandon et al., 2018; Pantel, 2016). It has been proposed that epigenetic modifications unlike mutations, can be targeted as potent biomarker discovery tools with various applications in risk assessment, therapy response prediction, early cancer detection and prognosis (Gordevičius et al., 2018; Leygo et al., 2017; Liu et al., 2018). In the early stages of tumourigenesis, epigenetic changes mostly precede the somatic mutations and the associated histopathological alterations can be detected (Peltomäki, 2012). An FDA-approved epigenetic-based assay, Epi pro Colon is used for the detection of colon cancer (Koch et al., 2018). Like traditional biomarkers, the genetic and epigenetic biomarkers may be limited by low sensitivity and specificity (Fiala & Diamandis, 2018).

2.5 Advances in Biomarker Discovery: Liquid Biopsies

In liquid biopsies, DNA methylation is the most studied epigenetic feature in cancer biomarker discovery, particularly in cfDNA. Five-methylcytosine (5mC) modification at the 5'-C-phosphate-G-3′ (CpG) dinucleotides is one of the most studied DNA methylation modifications (Feng et al., 2019; Guo et al., 2017; Kang et al., 2017; Leygo et al., 2017; Xu et al., 2017; Zeng et al., 2018, 2019). Xu et al. (2017)

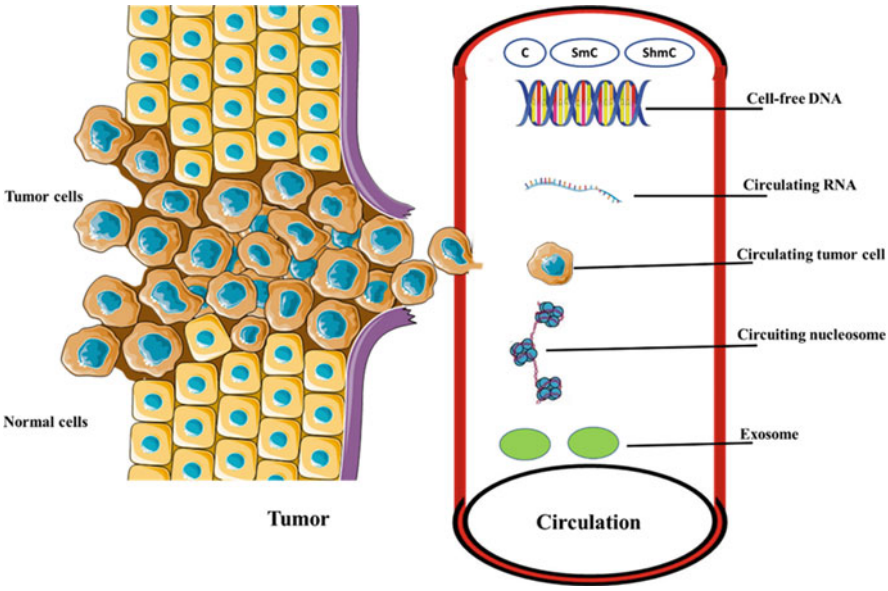


Fig. 2.2 Liquid biopsy biomarkers: Liquid biopsy biomarkers such as CTCs, cfNA, cfDNA, cNucleosomes and exosomes can be used in cancer diagnosis and prognosis. These biomarkers can be used to infer response to treatment also. DNA methylation patterns are key in detecting and identifying molecular alterations

demonstrated that 5mC DNA methylation biomarkers derived from ctDNA showed better prognosis indication than other biomarkers such as serum-based alpha-feto-protein (AFP) and TNM staging in hepatocellular carcinoma (Xu et al., 2017). Additionally, Wedge et al. (2017) showed that long intersperse nucleotide element 1 (LINE-1) cfDNA methylation alterations have a strong association with clinical outcomes in diffuse large B cell lymphoma, illustrating prognostic biomarker potential (Wedge et al., 2017).

In addition, identifying tissue-specific methylation haplotypes as biomarkers to elucidate tumor burden and cfDNA tissue of origin can be explored in 5mC biomarker discovery. Unlike the conventional single-CpG methylation biomarker in cancer stratification, the use of multi-CpG haplotypes shows potent promise in clinical applications (Guo et al., 2017). Furthermore, additional epigenetic features such as nucleosome positioning and occupancy on cfDNA, 5-hydroxymethylcytosine (5hmC) have been used to deduce cfDNA tissue of origin and cancer progression (Ivanov et al., 2015; Li et al., 2017; Snyder et al., 2016; Song et al., 2017; Tian et al., 2018). Notably, the clinical application of genome-wide nucleosome distribution of cfDNAs has not yet been broadly studied but may provide valuable information in distinguishing pooled cfDNA to construe tissue of origin (Lehmann-Werman et al., 2016; Snyder et al., 2016). Figure 2.2 demonstrates how tumor cells shed their cellular contents such as nucleic acids, into the

bloodstream. CTCs can also leave primary tumor site and infiltrate the bloodstream. CTCs, cfNAs especially cfDNA can be used as liquid biomarkers.

2.6 AI in Cancer Biomarker Discovery

Tissue biomarkers are unique information sets written in the tissue. These unique molecules are used in Pathology for the recognition of particular patient subsets with predictive, diagnostic, or prognostic purposes. Thus, biomarkers are key molecules in precision oncology. While Pathology may be considered as a subjective discipline due to personal natural differences in visual abilities, data integration and overall judgment, AI-based tools may aid in bridging this gap (Lancellotti et al., 2021). Whole slide imaging (WSI) applications in the 1990s led to growing interest in AI in Pathology, as it previously happened in Radiology field (Colling et al., 2019; Niazi et al., 2019; Parwani & Amin, 2020). As ML was the first AI method to be applied in Pathology, ML algorithms that were routinely used to discriminate benign against malignant tumors had limitations. ML algorithms were used to annotate unique morphological features such as the size of the cell, cytoplasmic texture, and nuclear shape. This unfortunately was identified to be time-consuming and fixed around anchored feature to the problem. Transformation of the AI-based tools in histopathology was experienced with the introduction of DL. DL methods were demonstrated to be able to learn directly from WSI/raw data and by not relying on engineering an anchored feature to the problem. However, raw data still need a threshold control, even though DL does not require pre-existing standards. For this reason, DL techniques are divided into unsupervised, weakly supervised, and strongly supervised. AI-based biomarkers are aimed at predicting response to treatment and patient survival and identifying somatic mutations (Coudray et al., 2018; Kather, Krisam, et al., 2019; Saltz et al., 2018; Skrede et al., 2020; Wulczyn et al., 2020). For example, the expression of Ki67, HER2/neu, estrogen, and progesterone tissue biomarkers allow for the selection of suitable treatment and outcome prediction of breast cancer patients. To date, various AI software are available either openly or commercially, as illustrated in Table 2.1. These can be used to address AI-based biomarker aims.

DL is an AI method that makes use of ANNs in the identification of recurring patterns in complex datasets such as in medical records. DL is a powerful tool used in modern day to directly extract hidden information from routinely available data, advancing the traditional molecular biology biomarker discovery. Imaging datasets in particular, have been reported to have high-density information which can be used with DL techniques. For instance, in radiology, DL has been reported to perform human tasks such as organ segmentation or tumor detection on CT images. DL use in radiology is mounting with currently approved FDA DL techniques such as DL-based CT data analysis that was carried out in lung cancer screening trial in 2019 (Ardila et al., 2019). MRI datasets have also been reported to work well with DL-based mining (Lundervold & Lundervold, 2019). Furthermore, DL has also been

Table 2.1 List of public and commercially available software

AI-based software	Source
1. HistoQC	https://github.com/choosehappy/HistoQC
2. Isola	https://histolab.readthedocs.io/en/latest/
3. HistomicsTK	https://digitalslidearchive.github.io/HistomicsTK/
4. QuPath	https://qupath.github.io
5. PyHIST	https://github.com/manuel-munoz-aguirre/PyHIST
6. PytorchDigitalPathology	https://github.com/CielAl/PytorchUnet
7. ASAP	https://computationalpathologygroup.github.io/ASAP
8. Ibex	https://ibex-ai.com/
9. Visiopharm	https://visiopharm.com/
10. Aiforia	https://www.aiforia.com/
11. Paige	https://paige.ai/
12. Proscia	https://proscia.com/

reported to demonstrate strong results for non-radiology tasks such as in dermoscopy for skin cancer detection and in colonoscopy (Haenssle et al., 2018; Luo et al., 2019; Yamada et al., 2019; Yap et al., 2018). However, due to the larger size of histology images, this imaging modality may contain more abundant information than is found in radiological images. With histology images, millions of different cells can be visualized on a slide. The cell's spatial arrangement and morphology have been reported to carry rich information compared to other medical imaging techniques. In comparing datasets from a whole histology slide and whole chest CT dataset of the same tumor from the same patient, the radiology imaging provided datasets fewer than the histology imaging. Even with the proposed use of DL methods in radiology, histological images are a rich source of DL-based biomarker mining compared to radiological images. Figure 2.3 shows a multistep process toward AI-based biomarker development.

Big data and AI cannot be separated in this modern-day precision oncology era. In genomics, "big data" can be referred to as large amounts of data generated by high throughput sequencing such as NGS. Big data is characterized by the 5 Vs, velocity, volume, value, variety, and veracity (Fountzilas & Tsimberidou, 2018). Traditional methods are unable to process big data, and thus advanced technology tools such as AI are required (Ioannidis & Khoury, 2018; Jain, 2016). Computational technologies are being continually developed for the identification of diagnostic and prognostic algorithms using the available clinical data (Mesko, 2017). For example, the IBM's (International Business Machines) Watson for Oncology is an AI program developed to analyze data from clinical notes, scientific reports and research, including the National Comprehensive Cancer Network (NCCN) Clinical Practice in Oncology guidelines (Kohn et al., 2014). This AI system can then suggest evidence-based personalized treatment by combining these data with patients' records. It has been reported that combining physician's notes and AI-obtained diagnostic algorithms is key to diagnostic accuracy of 99.5% (Wang et al., 2016).

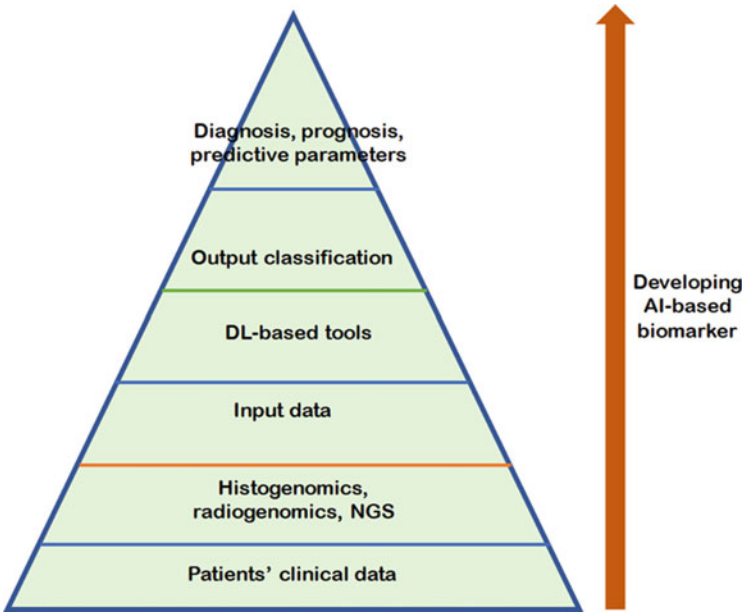


Fig. 2.3 Developing the AI-based biomarker. The AI program is fed by big data that may include NGS data, digital images, and clinical information. These programs learn to separate the classifications of interest without pre-existing assumptions. The output returns as categorized information of significant clinical value in the prediction, diagnosis, prognosis, and response to treatment of cancer

2.7 AI in the Detection of Novel Biomarkers for Accurate Prognostication and Prediction of Drug Resistance to Enhance Treatment

AI-based algorithms have demonstrated effectiveness in the prediction of microsatellite instability by analyzing general H and E-stained tissue slides (Hildebrand et al., 2021; Kather, Pearson, et al., 2019). Low-cost integrated use of this biomarker can be beneficial to immunotherapy application and aid to identify high-risk families. Furthermore, AI-based DNA methylation patterns in cancers can be useful in early detection and intervention (Dlamini et al., 2020; Wrzeszczynski et al., 2017). AI has the potential to transform the healthcare system, reducing costs and disparities. The workflow in the diagnosis of oncology patients with solid tumors involves obtaining tissue samples by either biopsy or by surgical resection. This is then followed by the preparation of pathology tissue slides for histological staining such as H and E or immunohistochemistry (IHC). H and E slides are the most routinely available for the majority of cancer patients. This makes it convenient to access these histology slides which contain abundant information that can be analyzed by DL methods. Although DL methods can be used in tumor detection and segmentation, these AI tools can be

used in histology images to advance personalized cancer patient care through predictive biomarker discovery. In addition to histopathology, it has been demonstrated that advanced methods such as DL are required to comprehend high-level labels directly from histology that cannot be comprehended by traditional pathology. This assumption involves three major areas which are interpreting genetic alterations, prognostic indicators, and the prediction of treatment response. These renders DL methods as valuable in clinical decision-making.

Contrary to DL methods in the clinical workflow, wet-lab-based assays such as polymerase chain reaction (PCR), in situ hybridization (ISH), IHC, NGS that are usually performed in parallel to histopathology, have been illustrated to have limitations. These limitations such as high costs, time consuming and complexity of datasets unfortunately pose as barriers to novel biomarker detection and to effective treatment regimens (Kather et al., 2020). However, challenges with current DL-based applications in oncology workflow include smaller training sample size and phenotypic strength of genetic targets. These limitations can be overcome by large datasets using a specific technical solution like the transfer learning approach (Costa & Czerniecki, 2020; Lancellotti et al., 2021; Mayekar & Bivona, 2017). Molecular testing in oncology clinical workflows form the basis to customize treatment according to the molecular makeup of the tumor tissue of advanced cancers. DL-based genotype biomarkers can be used to screen patients prior to genetic testing, while the solid use of these biomarkers is still in developmental stages, has not yet superseded the wet-lab assays. Notably, it has been reported that most of DL-based genotyping methods have lower reported AUROC values between 0.70 and 0.90, translating to 50% and 90% specificity and sensitivity. Although this is a low performance of what would usually be required of definitive tests, DL-based methods can be used to pre-screen cancer patients for rare traits. This may have significant alleviating effects on molecular testing loads. In DP, advances in DL biomarkers could meet the AUROC 0.90 threshold, translating to specificity and sensitivity equivalent to molecular assays. In such instances, DL methods could be considered as primary definitive testing or perhaps even be used parallel to web-lab assays to detect genotypic variations directly from histology slides, which may aid in treatment tailoring and predicting response.

2.8 Challenges, Limitations, and Opportunities

Oncology clinical workflows depend on predictive and prognostic biomarkers. However, it has been noted that due to the growing number of biomarker discovery and identification in clinical workflows, the delay in clinical decision-making and associated cost in this rising era cannot be ignored (Fitzgerald et al., 2021). In real clinical settings, the AI application is still limited by various problems, which include the low digitization level. For example, a survey in England showed that only 30% of Institutions had access to a full DP workstation. Of these, DP applications were revealed to be more common in research, teaching, and quality assurance

compared to clinical use. Even so, DP applications were found to be less likely to be used in primary diagnosis compared to consultations (Williams et al., 2018). Increased development of open source datasets such as the Cancer Genome Atlas, the Cancer Imaging Archives and Grand Challenges, a recent H2020 program IMI2–2019–2018 for a central repository can help alleviate the impediments of small training datasets for AI tools.

Although limitations in precision oncology involve ethical, legal, financial, and technological, data sharing is key for the success of this phenomenon. The American Association for Cancer Research (AACR, n.d.) has invested in an international data-sharing project, Genomics, Evidence, Neoplasia, Information, Exchange (GENIE). GENIE is a data-sharing registry that collects and combines Clinical Laboratory Improvement Amendments of 1988 (CLIA)-certified genomic data. Here, data is acquired from routine practice, from cancer treated patients at different institutions (AACR, n.d.). Additionally, the American Society of Clinical Oncology (ASCO's) Cancer Learning Intelligence Network for Quality (CancerLinQ) is also a data sharing and learning health program for the analysis of large aggregated de-identified electronic health records of similar patients' cases. This data-sharing program is reported to represent real-world evidence to help inform personalized patient care and treatment (Miller & Wong, 2018). Substantial resources have been dedicated to these initiatives, which are key to a paradigm shift in precision oncology, by identifying unique tumor patterns in predicting treatment response and overall patient outcome.

2.9 Conclusions and Perspectives

It is important to detect and identify biomarkers that induce drug resistance and also to identify novel therapeutic targets to enhance regimes of treatment (Dlamini et al., 2020). Precision oncology for populations or individuals could allow the use of specific diagnostic or prognostic biomarkers. This data can be used to monitor disease progression and response to treatment. Thus, this information can be used to decipher molecular alterations that can be targeted for the prediction of drug resistance and improve patient outcomes (Dlamini et al., 2022). The significance of molecular biomarkers is founded on sensitivity, specificity, and predictability for reduced therapeutic instability and improved patient outcomes (Barreto et al., 2012; Cheng et al., 2014; Dlamini et al., 2021). While tissue biomarkers have been used as gold standards in disease diagnosis and progression, AI tools have great potential in enriching human capabilities to optimize use of these in precision oncology. Additionally, liquid biopsies also hold great promise toward accurate prognostic and predictive biomarker use, despite their limitations in a clinical setting. Furthermore, findings generated from 'big data' studies can be overwhelming and therefore should be accompanied by advanced technological tools. It is undeniable that AI-based tools can enhance human potential in identifying changes in disease patterns. This will improve the prediction of unique patients' response to therapy and thus forge

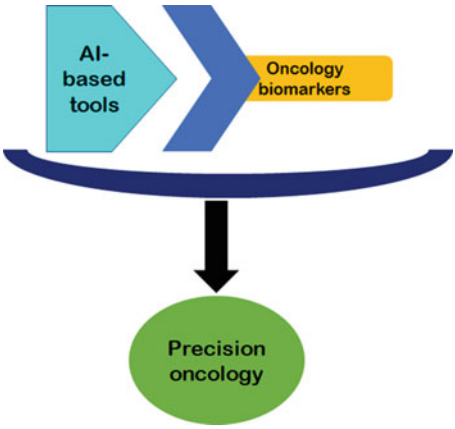


Fig. 2.4 AI-based tools in novel oncology biomarker discovery. AI-based tools have the potential to advance precision oncology through the detection and discovery of novel prognostic and predictive oncology biomarkers, thus enhancing treatment regimens. These parameters function in synergy such as lock (biomarkers) and key (AI-based tools) to transform oncology clinical workflows

personalized patient care. Both tissue and liquid biopsies can be leveraged by AI-based tools to identify novel predictive and prognostic biomarkers in advancing precision oncology, Fig. 2.4.

References

AACR. (n.d.) *Project GENIE* [Online]. Available: http://www.aacr.org/Research/Research/Pages/aacr-project-genie.aspx [Accessed].

Ahmed, Z., Mohamed, K., Zeeshan, S., & Dong, X. (2020). *Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine*. Database (Oxford).

Aktas, B., Kasimir-Bauer, S., Heubner, M., Kimmig, R., & Wimberger, P. (2011). Molecular profiling and prognostic relevance of circulating tumor cells in the blood of ovarian cancer patients at primary diagnosis and after platinum-based chemotherapy. *International Journal of Gynecological Cancer, 21*, 822–830.

Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., & Yu, X. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature, 403*, 503–511.

Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., Tse, D., Etemadi, M., Ye, W., & Corrado, G. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature Medicine, 25*, 954–961.

Banerjee, J., Pradhan, R., Gupta, A., Kumar, R., Sahu, V., Upadhyay, A. D., Chaterjee, P., Dwivedi, S., Dey, S., & Dey, A. B. (2017). CDK4 in lung, and head and neck cancers in old age: Evaluation as a biomarker. *Clinical and Translational Oncology, 19*, 571–578.

Barreto, L., Azambuja, D., & Morais, J. C. D. (2012). Expression of immunohistochemical markers in patients with AIDS-related lymphoma. *Brazilian Journal of Infectious Diseases, 16*, 74–77.

Cai, Q., Verma, S. C., Choi, J.-Y., Ma, M., & Robertson, E. S. (2010). Kaposi's sarcoma-associated herpesvirus inhibits interleukin-4-mediated STAT6 phosphorylation to regulate apoptosis and maintain latency. *Journal of Virology, 84*, 11134–11144.

Chen, J., Sun, M., & Shen, B. (2015). Deciphering oncogenic drivers: From single genes to integrated pathways. *Briefings in Bioinformatics, 16*, 413–428.

Chen, H. Y., Yu, S. L., Chen, C. H., Chang, G. C., Chen, C. Y., Yuan, A., Cheng, C. L., Wang, C. H., Terng, H. J., Kao, S. F., Chan, W. K., Li, H. N., Liu, C. C., Singh, S., Chen, W. J., Chen, J. J., & Yang, P. C. (2007). A five-gene signature and clinical outcome in non-small-cell lung cancer. *The New England Journal of Medicine, 356*, 11–20.

Cheng, L., Davison, D. D., Adams, J., Lopez-Beltran, A., Wang, L., Montironi, R., & Zhang, S. (2014). Biomarkers in bladder cancer: Translational and clinical implications. *Critical Reviews in Oncology/Hematology, 89*, 73–111.

Chyla, B., Daver, N., Doyle, K., Mckeegan, E., Huang, X., Ruvolo, V., Wang, Z., Chen, K., Souers, A., & Leverson, J. (2018). Genetic biomarkers of sensitivity and resistance to venetoclax monotherapy in patients with relapsed acute myeloid leukemia. *American Journal of Hematology, 93*, E202.

Colling, R., Pitman, H., Oien, K., Rajpoot, N., Macklin, P., Snead, D., Sackville, T., & Verrill, C. (2019). Artificial intelligence in digital pathology: A roadmap to routine use in clinical practice. *The Journal of Pathology, 249*, 143–150.

Costa, R. L. B., & Czerniecki, B. J. (2020). Clinical development of immunotherapies for HER2+ breast cancer: A review of HER2-directed monoclonal antibodies and beyond. *NPJ Breast Cancer, 6*, 1–11.

Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A. L., Razavian, N., & Tsirigos, A. (2018). Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature Medicine, 24*, 1559–1567.

Dlamini, Z., Francies, F. Z., Hull, R., & Marima, R. (2020). Artificial intelligence (AI) and big data in cancer and precision oncology. *Computational and Structural Biotechnology Journal, 18*, 2300–2311.

Dlamini, Z., Mbele, M., Makhafola, T. J., Hull, R., & Marima, R. (2021). HIV-associated cancer biomarkers: A requirement for early diagnosis. *International Journal of Molecular Sciences, 22*, 8127.

Dlamini, Z., Skepu, A., Kim, N., Mkhabele, M., Khanyile, R., Molefi, T., Mbatha, S., Setlai, B., Mulaudzi, T., Mabongo, M., Bida, M., Kgoebane-Maseko, M., Mathabe, K., Lockhat, Z., Kgokolo, M., Chauke-Malinga, N., Ramagaga, S., & Hull, R. (2022). AI and precision oncology in clinical cancer genomics: From prevention to targeted cancer therapies-an outcomes based patient care. *Informatics in Medicine Unlocked, 31*, 100965.

Echle, A., Rindtorff, N. T., Brinker, T. J., Luedde, T., Pearson, A. T., & Kather, J. N. (2021). Deep learning in cancer pathology: A new generation of clinical biomarkers. *British Journal of Cancer, 124*, 686–696.

Elton, D. C., Chen, A., Pickhardt, P. J., & Summers, R. M. (2021). *Cardiovascular disease and all-cause mortality risk prediction from abdominal CT using deep learning* (pp. 694–701). SPIE.

Fang, B. (2020). Introduction to this special issue: "Biomarker discovery and precision medicine". *Journal of Cancer Metastasis and Treatment, 6*, 1.

Farina, E., Nabhen, J. J., Dacoregio, M. I., Batalini, F., & Moraes, F. Y. (2022). An overview of artificial intelligence in oncology. *Future Science Open Access, 8*, Fso787.

Feng, H., Jin, P., & Wu, H. (2019). Disease prediction by cell-free DNA methylation. *Briefings in Bioinformatics, 20*, 585–597.

Fiala, C., & Diamandis, E. P. (2018). Utility of circulating tumor DNA in cancer diagnostics with emphasis on early detection. *BMC Medicine, 16*, 166.

Fitzgerald, J., Higgins, D., Mazo Vargas, C., Watson, W., Mooney, C., Rahman, A., Aspell, N., Connolly, A., Aura Gonzalez, C., & Gallagher, W. (2021). Future of biomarker evaluation in the realm of artificial intelligence algorithms: Application in improved therapeutic stratification of patients with breast and prostate cancer. *Journal of Clinical Pathology, 74*, 429.

Fountzilas, E., & Tsimberidou, A. M. (2018). Overview of precision oncology trials: Challenges and opportunities. *Expert Review of Clinical Pharmacology, 11*, 797–804.

Gerlinger, M., Horswell, S., Larkin, J., Rowan, A. J., Salm, M. P., Varela, I., Fisher, R., Mcgranahan, N., Matthews, N., Santos, C. R., Martinez, P., Phillimore, B., Begum, S., Rabinowitz, A., Spencer-Dene, B., Gulati, S., Bates, P. A., Stamp, G., Pickering, L., Gore, M., Nicol, D. L., Hazell, S., Futreal, P. A., Stewart, A., & Swanton, C. (2014). Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nature Genetics, 46*, 225–233.

Gordevičius, J., Kriščiūnas, A., Groot, D. E., Yip, S. M., Susic, M., Kwan, A., Kustra, R., Joshua, A. M., Chi, K. N., Petronis, A., & Oh, G. (2018). Cell-free DNA modification dynamics in abiraterone acetate-treated prostate cancer patients. *Clinical Cancer Research, 24*, 3317–3324.

Guo, S., Diep, D., Plongthongkum, N., Fung, H. L., Zhang, K., & Zhang, K. (2017). Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. *Nature Genetics, 49*, 635–642.

Haenssle, H. A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., Kalloo, A., Hassen, A. B. H., Thomas, L., & Enk, A. (2018). Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of Oncology, 29*, 1836–1842.

Heckman-Stoddard, B. M. (2012). *Oncology biomarkers: Discovery, validation, and clinical use* (pp. 93–98). Elsevier.

Hildebrand, L. A., Pierce, C. J., Dennis, M., Paracha, M., & Maoz, A. (2021). Artificial intelligence for histology-based detection of microsatellite instability and prediction of response to immunotherapy in colorectal cancer. *Cancers, 13*, 391.

Hoffman, R. M., Gilliland, F. D., Adams-Cameron, M., Hunt, W. C., & Key, C. R. (2002). Prostate-specific antigen testing accuracy in community practice. *BMC Family Practice, 3*, 19.

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer, 18*, 500–510.

Ilié, M., & Hofman, P. (2016). Pros: Can tissue biopsy be replaced by liquid biopsy? *Translational Lung Cancer Research, 5*, 420–423.

Ioannidis, J. P. A., & Khoury, M. J. (2018). Evidence-based medicine and big genomic data. *Human Molecular Genetics, 27*, R2–r7.

Ivanov, M., Baranova, A., Butler, T., Spellman, P., & Mileyko, V. (2015). Non-random fragmentation patterns in circulating cell-free DNA reflect epigenetic regulation. *BMC Genomics, 16*(Suppl 13), S1.

Jacobs, I. J., Menon, U., Ryan, A., Gentry-Maharaj, A., Burnell, M., Kalsi, J. K., Amso, N. N., Apostolidou, S., Benjamin, E., Cruickshank, D., Crump, D. N., Davies, S. K., Dawnay, A., Dobbs, S., Fletcher, G., Ford, J., Godfrey, K., Gunu, R., Habib, M., Hallett, R., Herod, J., Jenkins, H., Karpinskyj, C., Leeson, S., Lewis, S. J., Liston, W. R., Lopes, A., Mould, T., Murdoch, J., Oram, D., Rabideau, D. J., Reynolds, K., Scott, I., Seif, M. W., Sharma, A., Singh, N., Taylor, J., Warburton, F., Widschwendter, M., Williamson, K., Woolas, R., Fallowfield, L., Mcguire, A. J., Campbell, S., Parmar, M., & Skates, S. J. (2016). Ovarian cancer screening and mortality in the UK collaborative trial of ovarian cancer screening (UKCTOCS): A randomised controlled trial. *Lancet, 387*, 945–956.

Jahr, S., Hentze, H., Englisch, S., Hardt, D., Fackelmayer, F. O., Hesch, R. D., & Knippers, R. (2001). DNA fragments in the blood plasma of cancer patients: Quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Research, 61*, 1659–1665.

Jain, A. (2016). *The 5 Vs of big data*. Available from https://www.ibm.com/blogs/watson-health/the-5-vs-of-big-data

Kang, J., D'Andrea, A. D., & Kozono, D. (2012). A DNA repair pathway-focused score for prediction of outcomes in ovarian cancer treated with platinum-based chemotherapy. *Journal of the National Cancer Institute, 104*, 670–681.

Kang, S., Li, Q., Chen, Q., Zhou, Y., Park, S., Lee, G., Grimes, B., Krysan, K., Yu, M., Wang, W., Alber, F., Sun, F., Dubinett, S. M., Li, W., & Zhou, X. J. (2017). CancerLocator: Non-invasive cancer diagnosis and tissue-of-origin prediction using methylation profiles of cell-free DNA. *Genome Biology, 18*, 53.

Kather, J. N., Heij, L. R., Grabsch, H. I., Loeffler, C., Echle, A., Muti, H. S., Krause, J., Niehues, J. M., Sommer, K. A. J., & Bankhead, P. (2020). Pan-cancer image-based detection of clinically actionable genetic alterations. *Nature Cancer, 1*, 789–799.

Kather, J. N., Krisam, J., Charoentong, P., Luedde, T., Herpel, E., Weis, C.-A., Gaiser, T., Marx, A., Valous, N. A., Ferber, D., Jansen, L., Reyes-Aldasoro, C. C., Zörnig, I., Jäger, D., Brenner, H., Chang-Claude, J., Hoffmeister, M., & Halama, N. (2019). Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS Medicine, 16*, e1002730.

Kather, J. N., Pearson, A. T., Halama, N., Jäger, D., Krause, J., Loosen, S. H., Marx, A., Boor, P., Tacke, F., & Neumann, U. P. (2019). Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nature Medicine, 25*, 1054–1056.

Khan, K. H., Cunningham, D., Werner, B., Vlachogiannis, G., Spiteri, I., Heide, T., Mateos, J. F., Vatsiou, A., Lampis, A., Damavandi, M. D., Lote, H., Huntingford, I. S., Hedayat, S., Chau, I., Tunariu, N., Mentrasti, G., Trevisani, F., Rao, S., Anandappa, G., Watkins, D., Starling, N., Thomas, J., Peckitt, C., Khan, N., Rugge, M., Begum, R., Hezelova, B., Bryant, A., Jones, T., Proszek, P., Fassan, M., Hahne, J. C., Hubank, M., Braconi, C., Sottoriva, A., & Valeri, N. (2018). Longitudinal liquid biopsy and mathematical Modeling of clonal evolution forecast time to treatment failure in the PROSPECT-C phase II colorectal cancer clinical trial. *Cancer Discovery, 8*, 1270–1285.

Koch, A., Joosten, S. C., Feng, Z., De Ruijter, T. C., Draht, M. X., Melotte, V., Smits, K. M., Veeck, J., Herman, J. G., Van Neste, L., Van Criekinge, W., De Meyer, T., & Van Engeland, M. (2018). Analysis of DNA methylation in cancer: Location revisited. *Nature Reviews. Clinical Oncology, 15*, 459–466.

Kohn, M. S., Sun, J., Knoop, S., Shabo, A., Carmeli, B., Sow, D., Syed-Mahmood, T., & Rapp, W. (2014). IBM's health analytics and clinical decision support. *Yearbook of Medical Informatics, 9*, 154–162.

Lancellotti, C., Cancian, P., Savevski, V., Kotha, S. R. R., Fraggetta, F., Graziano, P., & Di Tommaso, L. (2021). Artificial Intelligence & tissue biomarkers: advantages, risks and perspectives for pathology. *Cells, 10*, 787.

Le, D. T., Durham, J. N., Smith, K. N., Wang, H., Bartlett, B. R., Aulakh, L. K., Lu, S., Kemberling, H., Wilt, C., & Luber, B. S. (2017). Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science, 357*, 409–413.

Lehmann-Werman, R., Neiman, D., Zemmour, H., Moss, J., Magenheim, J., Vaknin-Dembinsky, A., Rubertsson, S., Nellgård, B., Blennow, K., Zetterberg, H., Spalding, K., Haller, M. J., Wasserfall, C. H., Schatz, D. A., Greenbaum, C. J., Dorrell, C., Grompe, M., Zick, A., Hubert, A., Maoz, M., Fendrich, V., Bartsch, D. K., Golan, T., Ben Sasson, S. A., Zamir, G., Razin, A., Cedar, H., Shapiro, A. M., Glaser, B., Shemer, R., & Dor, Y. (2016). Identification of tissue-specific cell death using methylation patterns of circulating DNA. *Proceedings of the National Academy of Sciences of the United States of America, 113*, E1826–E1834.

Lesko, L. J., & Atkinson, A. J., Jr. (2001). Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: Criteria, validation, strategies. *Annual Review of Pharmacology and Toxicology, 41*, 347.

Leygo, C., Williams, M., Jin, H. C., Chan, M. W. Y., Chu, W. K., Grusch, M., & Cheng, Y. Y. (2017). DNA methylation as a noninvasive epigenetic biomarker for the detection of cancer. *Disease Markers, 2017*, 3726595.

Li, W., Zhang, X., Lu, X., You, L., Song, Y., Luo, Z., Zhang, J., Nie, J., Zheng, W., Xu, D., Wang, Y., Dong, Y., Yu, S., Hong, J., Shi, J., Hao, H., Luo, F., Hua, L., Wang, P., Qian, X., Yuan, F., Wei, L., Cui, M., Zhang, T., Liao, Q., Dai, M., Liu, Z., Chen, G., Meckel, K., Adhikari, S., Jia, G., Bissonnette, M. B., Zhang, X., Zhao, Y., Zhang, W., He, C., & Liu, J. (2017). 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. *Cell Research, 27*, 1243–1257.

Lim, S. M., Hong, M. H., & Kim, H. R. (2020). Immunotherapy for non-small cell lung cancer: Current landscape and future perspectives. *Immune Network, 20*, e10.

Lin, Y., Qian, F., Shen, L., Chen, F., Chen, J., & Shen, B. (2019). Computer-aided biomarker discovery for precision medicine: Data resources, models and applications. *Briefings in Bioinformatics, 20*, 952–975.

Liu, L., Toung, J. M., Jassowicz, A. F., Vijayaraghavan, R., Kang, H., Zhang, R., Kruglyak, K. M., Huang, H. J., Hinoue, T., Shen, H., Salathia, N. S., Hong, D. S., Naing, A., Subbiah, V., Piha-Paul, S. A., Bibikova, M., Granger, G., Barnes, B., Shen, R., Gutekunst, K., Fu, S., Tsimberidou, A. M., Lu, C., Eng, C., Moulder, S. L., Kopetz, E. S., Amaria, R. N., Meric-Bernstam, F., Laird, P. W., Fan, J. B., & Janku, F. (2018). Targeted methylation sequencing of plasma cell-free DNA for cancer detection and classification. *Annals of Oncology, 29*, 1445–1453.

Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik, 29*, 102–127.

Luo, H., Xu, G., Li, C., He, L., Luo, L., Wang, Z., Jing, B., Deng, Y., Jin, Y., & Li, Y. (2019). Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: A multicentre, case-control, diagnostic study. *The Lancet Oncology, 20*, 1645–1654.

Ma, J., Yang, J., Jin, Y., Cheng, S., Huang, S., Zhang, N., & Wang, Y. (2021). Artificial intelligence based on blood biomarkers including CTCs predicts outcomes in epithelial ovarian cancer: A prospective study. *Oncotargets and Therapy, 14*, 3267–3280.

Maisel, A. S., Katz, N., Hillege, H. L., Shaw, A., Zanco, P., Bellomo, R., Anand, I., Anker, S. D., Aspromonte, N., & Bagshaw, S. M. (2011). Biomarkers in kidney and heart disease. *Nephrology Dialysis Transplantation, 26*, 62–74.

Makary, M. A., & Daniel, M. (2016). Medical error-the third leading cause of death in the US. *BMJ, 353*, i2139.

Mayekar, M. K., & Bivona, T. G. (2017). Current landscape of targeted therapy in lung cancer. *Clinical Pharmacology and Therapeutics, 102*, 757–764.

Merker, J. D., Oxnard, G. R., Compton, C., Diehn, M., Hurley, P., Lazar, A. J., Lindeman, N., Lockwood, C. M., Rai, A. J., Schilsky, R. L., Tsimberidou, A. M., Vasalos, P., Billman, B. L., Oliver, T. K., Bruinooge, S. S., Hayes, D. F., & Turner, N. C. (2018). Circulating tumor DNA analysis in patients with cancer: American Society of Clinical Oncology and College of American Pathologists Joint Review. *Journal of Clinical Oncology, 36*, 1631–1641.

Mesko, B. (2017). The role of artificial intelligence in precision medicine. *Expert Review of Precision Medicine and Drug Development, 2*, 239–241.

Miller, R. S., & Wong, J. L. (2018). Using oncology real-world evidence for quality improvement and discovery: The case for ASCO's CancerLinQ. *Future Oncology, 14*, 5–8.

Mondelo-Macía, P., García-González, J., León-Mateos, L., Castillo-García, A., López-López, R., Muinelo-Romay, L., & Díaz-Peña, R. (2021). Current status and future perspectives of liquid biopsy in small cell lung cancer. *Biomedicines, 9*(1), 48. https://doi.org/10.3390/biomedicines9010048

Naito, Y., & Urasaki, T. (2018). Precision medicine in breast cancer. *Chinese Clinical Oncology, 7*, 29.

Narod, S. (2016). Can advanced-stage ovarian cancer be cured? *Nature Reviews. Clinical Oncology, 13*, 255–261.

Niazi, M. K. K., Parwani, A. V., & Gurcan, M. N. (2019). Digital pathology and artificial intelligence. *The Lancet Oncology, 20*, e253–e261.

Ossandon, M. R., Agrawal, L., Bernhard, E. J., Conley, B. A., Dey, S. M., Divi, R. L., Guan, P., Lively, T. G., Mckee, T. C., Sorg, B. S., & Tricoli, J. V. (2018). Circulating tumor DNA assays in clinical cancer research. *Journal of the National Cancer Institute, 110*, 929–934.

Pantel, K. (2016). Blood-based analysis of circulating cell-free DNA and tumor cells for early cancer detection. *PLoS Medicine, 13*, e1002205.

Parwani, A. V., & Amin, M. B. (2020). Convergence of digital pathology and artificial intelligence tools in anatomic pathology practice: Current landscape and future directions. *Advances in Anatomic Pathology, 27*, 221–226.

Peltomäki, P. (2012). Mutations and epimutations in the origin of cancer. *Experimental Cell Research, 318*, 299–310.

Pickhardt, P. J., Graffy, P. M., Zea, R., Lee, S. J., Liu, J., Sandfort, V., & Summers, R. M. (2020). Automated CT biomarkers for opportunistic prediction of future cardiovascular events and mortality in an asymptomatic screening population: A retrospective cohort study. *The Lancet Digital Health, 2*, e192–e200.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*, 1347–1358.

Rezvantalab, A., Safigholi, H. & Karimijeshni, S. (2018). Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms. *arXiv preprint arXiv:1810.10348*.

Sabaawy, H. E. (2013). Genetic heterogeneity and clonal evolution of tumor cells and their impact on precision cancer medicine. *Journal of Leukemia (Los Angel), 1*, 1000124.

Sahutoglu, T., Sakaci, T., Hasbal, N. B., Ahbap, E., Kara, E., Sumerkan, M. C., Sevinc, M., Akgol, C., Koc, Y., & Basturk, T. (2017). Serum VEGF-C levels as a candidate biomarker of hypervolemia in chronic kidney disease. *Medicine, 96*, e6543.

Saltz, J., Gupta, R., Hou, L., Kurc, T., Singh, P., Nguyen, V., Samaras, D., Shroyer, K. R., Zhao, T., Batiste, R., Van Arnam, J., Shmulevich, I., Rao, A. U. K., Lazar, A. J., Sharma, A., & Thorsson, V. (2018). Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Reports, 23*, 181–193.e7.

Skrede, O. J., De Raedt, S., Kleppe, A., Hveem, T. S., Liestøl, K., Maddison, J., Askautrud, H. A., Pradhan, M., Nesheim, J. A., Albregtsen, F., Farstad, I. N., Domingo, E., Church, D. N., Nesbakken, A., Shepherd, N. A., Tomlinson, I., Kerr, R., Novelli, M., Kerr, D. J., & Danielsen, H. E. (2020). Deep learning for prediction of colorectal cancer outcome: A discovery and validation study. *Lancet, 395*, 350–360.

Snyder, M. W., Kircher, M., Hill, A. J., Daza, R. M., & Shendure, J. (2016). Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. *Cell, 164*, 57–68.

Song, C. X., Yin, S., Ma, L., Wheeler, A., Chen, Y., Zhang, Y., Liu, B., Xiong, J., Zhang, W., Hu, J., Zhou, Z., Dong, B., Tian, Z., Jeffrey, S. S., Chua, M. S., So, S., Li, W., Wei, Y., Diao, J., Xie, D., & Quake, S. R. (2017). 5-Hydroxymethylcytosine signatures in cell-free DNA provide information about tumor types and stages. *Cell Research, 27*, 1231–1242.

Strimbu, K., & Tavel, J. A. (2010). What are biomarkers? *Current Opinion in HIV and AIDS, 5*, 463–466.

Tian, X., Sun, B., Chen, C., Gao, C., Zhang, J., Lu, X., Wang, L., Li, X., Xing, Y., Liu, R., Han, X., Qi, Z., Zhang, X., He, C., Han, D., Yang, Y. G., & Kan, Q. (2018). Circulating tumor DNA 5-hydroxymethylcytosine as a novel diagnostic biomarker for esophageal cancer. *Cell Research, 28*, 597–600.

Verma, M. (2012). Personalized medicine and cancer. *Journal of Personalized Medicine, 2*, 1–14.

Wan, J. C. M., Massie, C., Garcia-Corbacho, J., Mouliere, F., Brenton, J. D., Caldas, C., Pacey, S., Baird, R., & Rosenfeld, N. (2017). Liquid biopsies come of age: Towards implementation of circulating tumour DNA. *Nature Reviews. Cancer, 17*, 223–238.

Wang, F., Casalino, L. P., & Khullar, D. (2019). Deep learning in medicine—Promise, progress, and challenges. *JAMA Internal Medicine, 179*, 293–294.

Wang, D., Khosla, A., Gargeya, R., Irshad, H. & Beck, A. H. (2016). Deep learning for identifying metastatic breast cancer. *ArXiv,* abs/1606.05718.

Waseem, M., Ahmad, M. K., Srivatava, V. K., Rastogi, N., Serajuddin, M., Kumar, S., Mishra, D. P., Sankhwar, S. N., & Mahdi, A. A. (2017). Evaluation of miR-711 as novel biomarker in prostate cancer progression. *Asian Pacific Journal of Cancer Prevention: APJCP, 18*, 2185.

Wedge, E., Hansen, J. W., Garde, C., Asmar, F., Tholstrup, D., Kristensen, S. S., Munch-Petersen, H. D., Ralfkiaer, E., Brown, P., Grønbaek, K., & Kristensen, L. S. (2017). Global hypomethylation is an independent prognostic factor in diffuse large B cell lymphoma. *American Journal of Hematology, 92*, 689–694.

Williams, B. J., Lee, J., Oien, K. A., & Treanor, D. (2018). Digital pathology access and usage in the UK: Results from a national survey on behalf of the National Cancer Research Institute's CM-path initiative. *Journal of Clinical Pathology, 71*, 463–466.

Wrzeszczynski, K. O., Frank, M. O., Koyama, T., Rhrissorrakrai, K., Robine, N., Utro, F., Emde, A.-K., Chen, B.-J., Arora, K., & Shah, M. (2017). Comparing sequencing assays and human-machine analyses in actionable genomics for glioblastoma. *Neurology Genetics, 3*, e164.

Wulczyn, E., Steiner, D. F., Xu, Z., Sadhwani, A., Wang, H., Flament-Auvigne, I., Mermel, C. H., Chen, P.-H. C., Liu, Y., & Stumpe, M. C. (2020). Deep learning-based survival prediction for multiple cancer types using histopathology images. *PLoS One, 15*, e0233678.

Wyatt, A. W., Annala, M., Aggarwal, R., Beja, K., Feng, F., Youngren, J., Foye, A., Lloyd, P., Nykter, M., Beer, T. M., Alumkal, J. J., Thomas, G. V., Reiter, R. E., Rettig, M. B., Evans, C. P., Gao, A. C., Chi, K. N., Small, E. J., & Gleave, M. E. (2017). Concordance of circulating tumor DNA and matched metastatic tissue biopsy in prostate cancer. *Journal of the National Cancer Institute, 109*, 78.

Xu, R. H., Wei, W., Krawczyk, M., Wang, W., Luo, H., Flagg, K., Yi, S., Shi, W., Quan, Q., Li, K., Zheng, L., Zhang, H., Caughey, B. A., Zhao, Q., Hou, J., Zhang, R., Xu, Y., Cai, H., Li, G., Hou, R., Zhong, Z., Lin, D., Fu, X., Zhu, J., Duan, Y., Yu, M., Ying, B., Zhang, W., Wang, J., Zhang, E., Zhang, C., Li, O., Guo, R., Carter, H., Zhu, J. K., Hao, X., & Zhang, K. (2017). Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma. *Nature Materials, 16*, 1155–1161.

Yamada, M., Saito, Y., Imaoka, H., Saiko, M., Yamada, S., Kondo, H., Takamaru, H., Sakamoto, T., Sese, J., & Kuchiba, A. (2019). Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Scientific Reports, 9*, 1–9.

Yap, J., Yolland, W., & Tschandl, P. (2018). Multimodal skin lesion classification using deep learning. *Experimental Dermatology, 27*, 1261–1267.

Zeng, H., He, B., Yi, C., & Peng, J. (2018). Liquid biopsies: DNA methylation analyses in circulating cell-free DNA. *Journal of Genetics and Genomics, 45*, 185–192.

Zeng, C., Stroup, E. K., Zhang, Z., Chiu, B. C., & Zhang, W. (2019). Towards precision medicine: Advances in 5-hydroxymethylcytosine cancer biomarker discovery in liquid biopsy. *Cancer Communications (London), 39*, 12.

Zhang, J., Fujimoto, J., Zhang, J., Wedge, D. C., Song, X., Zhang, J., Seth, S., Chow, C. W., Cao, Y., Gumbs, C., Gold, K. A., Kalhor, N., Little, L., Mahadeshwar, H., Moran, C., Protopopov, A., Sun, H., Tang, J., Wu, X., Ye, Y., William, W. N., Lee, J. J., Heymach, J. V., Hong, W. K., Swisher, S., Wistuba, I. I., & Futreal, P. A. (2014). Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science, 346*, 256–259.

Chapter 3
Use of Artificial Intelligence in Implementing Mainstream Precision Medicine to Improve Traditional Symptom-driven Practice of Medicine: Allowing Early Interventions and Tailoring better-personalised Cancer Treatments

Thifhelimbilu Luvhengo, Thulo Molefi, Demetra Demetriou, Rodney Hull, and Zodwa Dlamini

Abstract Cancer was until recently considered a homogenous disease with clearly defined boundaries based on the organ involved and the TNM stage of a tumour. Current challenges in cancer treatment include the shortage of expertise, delay in diagnosis, inaccurate quantitative staging and variable treatment response. An improvement in the understanding of the pathogenesis of cancer, tumour microenvironment and metastatic pathways has led to an increase in the application of precision medicine in the management of malignant tumours. Recent advances in imaging, histopathological analysis, genomics, transcriptomics, epigenomics, proteomics and metabolomics have increased the volume of information available to guide personalised management of patients diagnosed with cancer. The ability to combine demographic, clinical, radiologic and genomic findings has made the

T. Luvhengo (✉)
Department of Surgery, Charlotte Maxeke Johannesburg Academic Hospital, Johannesburg, South Africa
e-mail: Thifhelimbilu.luvhengo@wits.ac.za

T. Molefi
Department of Medical Oncology, Steve Biko Academic Hospital and University of Pretoria, Pretoria, South Africa

SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: thulo.molefi@up.ac.za

D. Demetriou · R. Hull · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: dd.demetriou@up.ac.za; rodney.hull@up.ac.za; zodwa.dlamini@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_3

application of precision medicine in the management of cancer more feasible. This chapter will report on the current uses of AI in the management of common malignancies and how it would facilitate the implementation of mainstream precision medicine for screening, early diagnosis and personalised treatment of common solid and haematological malignancies.

Keywords Artificial Intelligence · Intervention · Oncology · Precision Medicine · Tailored treatment

3.1 Introduction

Cancer is the leading cause of death globally and its incidence is increasing exponentially across the world, especially in younger adults (Kitahara & Sosa, 2016; Marur & Forastiere, 2008; Miller et al., 2020). Most cancers would be curable if they are diagnosed early and treated timeously. However, majority of patients however present when the cancer is either locally advanced or metastatic which reduces the likelihood of cure (Mamelle et al., 1994; Wu et al., 2021). In certain cases, some cancers acquire resistance to previously effective chemotherapy, targeted therapy or immunotherapy (Alqahtani et al., 2019; Mangaj et al., 2021). The options for the treatment of cancer include surgery, radiotherapy, chemotherapy, hormonal therapy, immunotherapy and targeted therapy.

Historically clinicians relied on the "one size fits all" approach for screening, history taking, physical examination, staging, treatment, surveillance of side effects and prognostication of patients who were diagnosed with cancer. Among the limitations of clinical and image-dependent qualitative TNM staging is the subjectivity during the interpretation of findings and its inability to accurately quantify the actual burden of the disease or predict its clinical behaviour. Among the other reasons for variable treatment response and outcome of cancers is the inter- and intra-tumour variability (Borczuk et al., 2009). The interpretation of results of imaging and histopathological investigations is influenced by the experience of a radiologist and pathologist, respectively. Furthermore, once cancer has developed it continues to evolve and becomes heterogeneous within itself, and at its metastatic sites (AlSendi et al., 2021). Analysis of biopsy specimen is usually limited to a small area of the tumour even if an excision biopsy was done (Vo et al., 2020). The diagnostic workup of cancer rarely includes a biopsy of metastatic sites whose microenvironment is likely to be different to that of the primary. The tumour markers that are currently used for screening, diagnosis or follow-up of most cancers are not reliable and often there is no correlation between their serum level and the volume of cancer (Zhang et al., 2020b).

Precision medicine is the ability to offer the most specific, appropriate, effective, efficient and safe treatment to a patient. Originally, precision medicine was limited to the use of results from genomics studies to guide treatment but precision medicine currently incorporates analysis of lifestyle, environment and bio-information to guide treatment (Beckmann & Lew, 2016; Canzoneri et al., 2019; Rogers et al.,

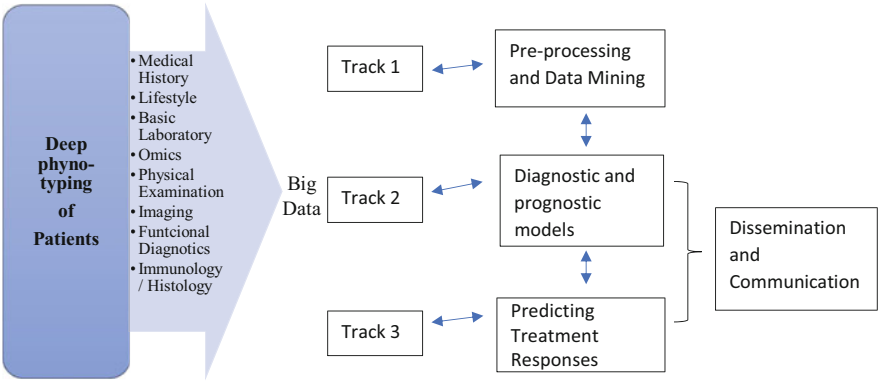


Fig. 3.1 Illustration of the potential application of precision medicine based on a combination of results obtained from imaging, histopathology and genomic investigations to guide treatment. Deep phenotyping of patients and the workup includes the medical history, basic lifestyle, laboratory results, omic results, physical examination, functional diagnosis and immunology results. This leads to enormous amounts of data. The data needs to be pre-processed and selected via data mining. Diagnostic and predictive models can be designed based on the results obtained. The models will allow the prediction of treatment that will be more effective. These models can be shared with the relevant parties

2020; Wu et al., 2021). Precise and personalised treatment of cancer relies on the ability to timeously make an accurate diagnosis, correctly stage the disease, select the most effective treatment, minimise side effects and detect resistance or recurrent disease early when it is still treatable (Bhinder et al., 2021. Precision medicine uses results from imaging, histopathology and/or genomic investigations to guide treatment (Fig. 3.1) (Konig et al., 2017).

Collation and analysis of bio-information, big data generated from electronic health information system and next-generation sequencing is beyond the capabilities of a human brain. Artificial intelligence (AI) has been adopted by several industries to generate algorithms that assist decision-making (Rogers et al., 2020). The use of AI to assist decision-making began in the 1950s but an exponential increase in the use of AI in the healthcare industry started recently (Ahmed et al., 2021; Bourcier et al., 2021; Connor, 2019; Keshinbora & Guven, 2020; Kroner et al., 2021; Luchini et al., 2022; Nensa et al., 2019; Weidlich & Weidlich, 2018). Artificial intelligence in the diagnostic workup, treatment planning and follow-up of common cancers such as carcinoma of the breast, colon, prostate and lung (Luchini et al., 2022) (Fig. 3.2).

3.2 Artificial Intelligence

Artificial intelligence was introduced in the 1950s but its use in the healthcare industry only started within the last 15 years (Bhinder et al., 2021; He et al., 2019). Artificial intelligence (AI) is the use of computer applications to imitate

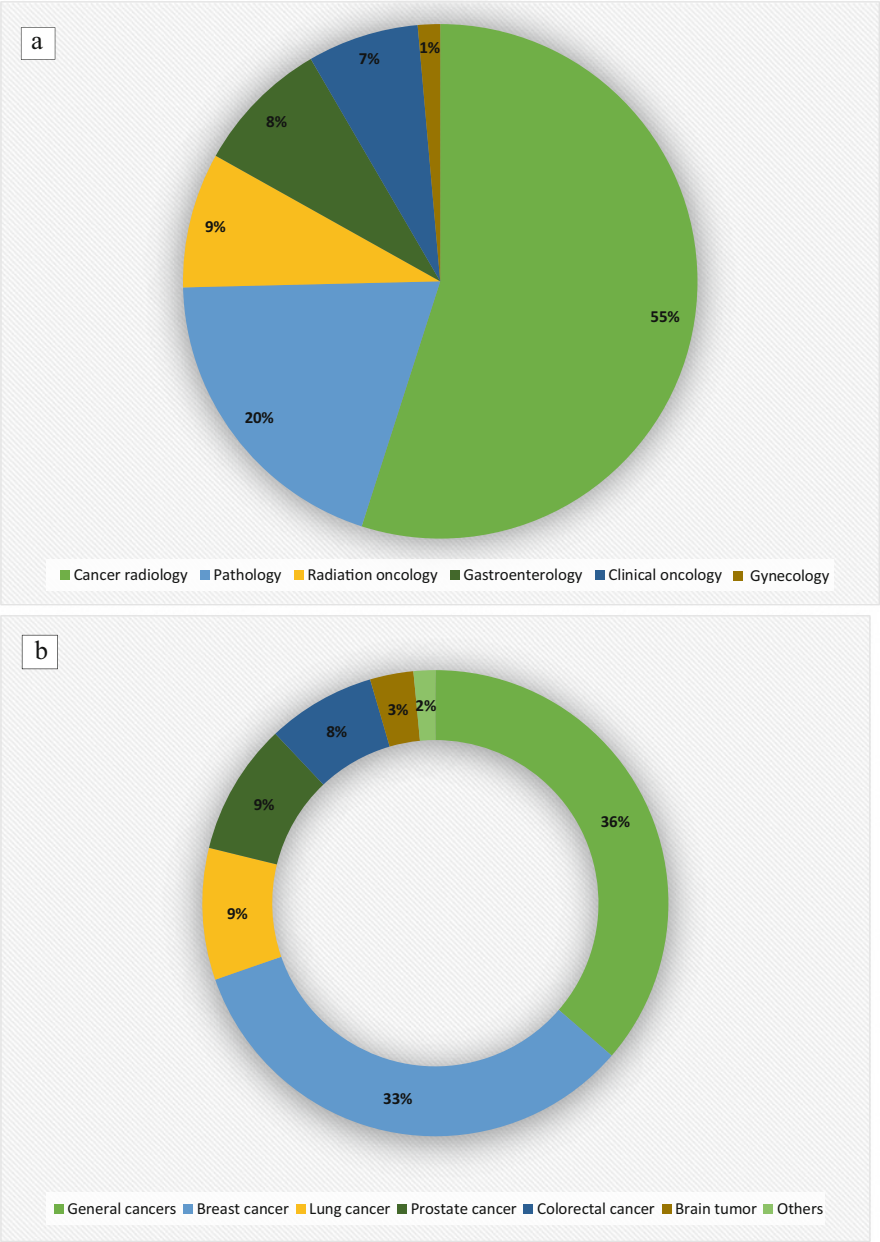


Fig. 3.2 Status of Artificial Intelligence in oncology. 3.2a and 3.2b provide the representation of AI-based devices that is expressed by oncology-related specialities. 3.2a Cancer radiology at—55%, Pathology at—20%, Radiation oncology at—9%, Gastroenterology at—8%, Clinical oncology at 7% and Gynaecology at—1%. 1b) General cancer at—36%, Breast cancer at—33%, Lung cancer at—9%, Prostate cancer at—9%, Colorectal cancer at ~8%, Brain tumours at—3% and Others 2%

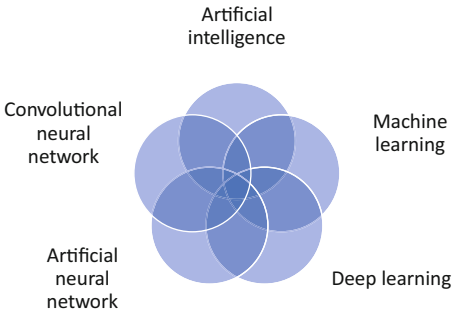


Fig. 3.3 The scope of artificial intelligence ranges from simple machine learning to the more advanced convolutional neural network

human intelligence (Liu et al., 2019; Shimizu & Nakayama, 2020). A more basic form of AI is machine learning which uses computers to develop algorithms (Fig. 3.3).

The simplest form of deep learning is artificial neural network (ANN). Convolutional neural network (CNN) is more advanced as it includes more than one hidden layer. A convolutional neural network relies on the input of previously collected clinical, radiological, histopathologicalor genomic data to develop a recognition model for future encounters (Bi et al., 2019; Honsy et al., 2018). Among the reported benefits of AI in oncology is the facilitation of early detection of cancer, virtual biopsy, prediction of metastasis, assessment of the adequacy of tumour excision margin and determination of the prognosis. A combination of findings from personal, lifestyle, clinical, radiomics, epigenomics, proteomics and metabolomics (medomics) may be used for risk-stratification of cancer and facilitate the application of precision medicine (Dlamini et al., 2020; Wu et al., 2021) (Fig. 3.4).

3.3 Use of Artificial Intelligence for Early Interventions and Tailoring Better-personalised Treatment of Common Cancers

Breast, colon, lung, prostate, skin, stomach, lung, cervical, oesophageal and hepatocellular carcinoma are among the 10 most commonly diagnosed cancers in adults. The other common cancers include head and neck, ovarian, anal, pancreas, brain and haematological malignancies. Common to all cancers is a presentation at advanced in close to 70% of the cases, inter- and intra-tumoural heterogeneity, the plurality of molecular subtypes, variation in their response to treatment and frequent acquisition of resistance to therapies during follow-up.

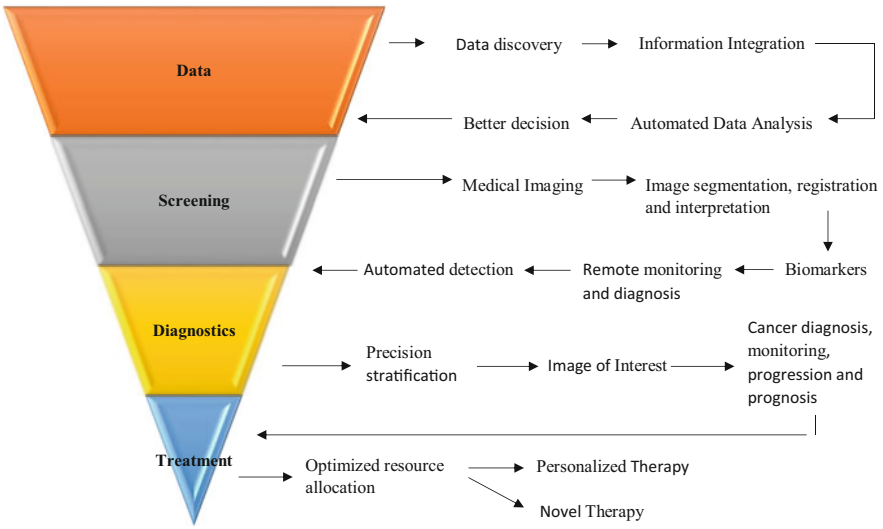


Fig. 3.4 An illustration of areas of collection of information that can be enhanced by AI and used to facilitate the delivery of precision medicine. Artificial intelligence enhances data discovery and allows the integration of information. This process can be automated and lead to improved decision-making. AI assists in image segmentation, registration and interpretation. AI allows the identification of biomarkers that allows remote monitoring and diagnosis. AI also allows automated detection. Regarding the diagnosis, AI allows precision stratification and the selection of the image of interest. This assists in cancer diagnosis, monitoring, progression and prognosis. AI allows for optimised resource allocation and leads to personalised therapy and novel therapies

3.3.1 Breast Cancer

Breast cancer is the most common cancer in women worldwide and its incidence is increasing (Miller et al., 2020). While screening using imaging modalities such as mammography, tomosynthesis, ultrasound and/or MRI are reliable methods for the detection of breast cancer, the expertise for interpretation of the findings is not universally available (Geras et al., 2019). The diagnosis of breast cancer requires a core needle biopsy. Core needle biopsy is invasive and only sample a small area of the tumour. Radical or modified radical mastectomy was the standard of care for curable breast cancer, but breast-conserving surgery is currently appropriate for most cases of early breast cancer. A select group of patients who have oligometastatic breast cancer may benefit from curative treatment (AlSendi et al., 2021).

The stage and molecular subtypes of breast cancer, the microenvironment in the tumour and its genetic landscape influence the necessity of neo-adjuvant and/or adjuvant (Garrido-Castro et al., 2019). The molecular subtype, presenting symptoms and sites of metastases from breast cancer influence the selection of palliative treatment (Garrido-Castro et al., 2019). Chemotherapy and trastuzumab are for all HER-2/neu enriched tumours irrespective of the TNM stage (Lerebours et al., 2021). Categorisation of breast cancer into molecular subtypes uses the differentiation of a

tumour, oestrogen and progesterone receptor status, HER-2 neu expression, mitotic count and Ki-67 index (Garrido-Castro et al., 2019). The genetic landscape and tumour microenvironments such as overexpression of growth factors and the existence of tumour infiltrating lymphocytes also influence the behaviour of breast cancer inclusive of its response to treatment (Low et al., 2018). The behaviour of some cases of breast cancer is not predictable through manual collation and interpretation of assembled information, despite the involvement of a multidisciplinary team.

Early diagnosis and targeted treatment based on the molecular subtype and mutational analysis of breast cancer offers the greatest hope for a cure. The volume of data obtained from clinical, radiological, histopathological and molecular testing during the evaluation of a patient who has breast cancer is beyond the integrative ability of humans. The introduction of radiomics has improved the accuracy of the interpretation of mammography where there is limited availability of experienced radiologists (Geras et al., 2019). Challenges associated with the histopathological assessment of a tissue specimen from the breast include the heterogeneity of the tumour, limited availability of pathologists, inter-observer variability and delay in communicating the results due to workload (Ibrahim et al., 2020). The application of AI through the digitalization of histology slides and the use of deep learning-based interpretation provides more accurate results and categorisation of subtypes of breast cancer and facilitates timeous reporting of histology results (Ibrahim et al., 2020). Furthermore, it is easier to select areas of interest while using digital slides. Analysis of digital slides is, therefore, more likely to be comprehensive and lead to a better characterisation of cancer including its microenvironment and the molecular subtypes (Ibrahim et al., 2020; Low et al., 2018). Artificial intelligence can predict the histological type, tumour grade and molecular subtype of breast cancer without a tissue biopsy. The other reported benefits of AI in breast cancer include the prediction of the likelihood of additionally involved lymph nodes if a sentinel lymph node is positive. Artificial intelligence programs can accurately predict the presence of distant metastases.

3.3.2 Colorectal Cancer

Colorectal cancer is the third most common cancer globally and is the second leading cause of cancer-related mortalities (Mitsala et al., 2021). The incidence of colorectal cancer is increasing globally, especially among young adults including in low- and middle-income countries. Close to 90% of colorectal cancer are sporadic and develop from an adenomatous polyp through a stepwise process, which takes on average 10–15 years (Kather et al., 2018). The genomic landscapes changes and new mutations develop as colorectal cancer grows and metastasises (Mitsala et al., 2021). The prognosis of colorectal is better if diagnosed before it has metastasised to lymph nodes or the liver. Approximately 20% of colorectal cancer present when the tumour has already metastasized (Aigner et al., 2017; Lin et al., 2020; Maclean et al., 2021;

Rocca et al., 2022; Rompianesi et al., 2022). Some cases of metastatic colorectal cancer are curable (Aigner et al., 2017).

Colonoscopy is the most reliable method for screening colorectal cancer whereas the diagnostic modalities include colonoscopy, endoscopic ultrasound, CT scan and MRI. Management of colorectal cancer depends on the stage of the tumour at presentation and site of the tumour, histological differentiation, genetic make-up, molecular subtype and tumour microenvironment (Kather et al., 2018). Options for the treatment of early colorectal cancer include endoscopic mucosal resection, endoscopic submucosal dissection, trans-anal excision, and trans-anal endoscopic microsurgery (Mitsala et al., 2021). The choice of endoscopic treatment is dependent on early diagnosis, accurate pre-operative histological diagnosis and tumour staging. Chemotherapy, radiotherapy and targeted therapy when the cancer is either locally advanced or metastatic.

The introduction of AI has led to a reduction of missed polyps during colonoscopy, accurate assessment of cancer risk and depth of tumour invasion in adenomatous polyps, accurate assessment of lymph node status, early detection of liver metastases and prediction of response to treatment (Bedrikovetski et al., 2021; Dayde et al., 2017; Rompianesi et al., 2022). The liver is the most common site for metastases from colorectal cancer and 25–50% of patients have liver metastases at first presentation (Lin et al., 2020; Liu et al., 2020; Rocca et al., 2022). Colorectal cancer with limited metastases in a selected group of patients is curable. Artificial intelligence combining CT, MRI, mass-spectrometry of exhaled volatile compounds, demographics, CEA and tumour stage are useful for timeous detection of liver metastases when they are still resectable or suitable for liver transplant (Rompianesi et al., 2022).

Additional challenges in the management of rectal cancer are the risk of local recurrence and the need to preserve the anal sphincter. The interpretation of findings from endo-luminal ultrasound and MRI is subjective and reliant on the experience of the radiologist. Although minimal surgery options are preferred, they are not suitable for all patients. Artificial intelligence programs combining demographic information, radiomics, genomics, tumour markers and proteomics may guide treatment selection and predict treatment response (Liu et al., 2020; Wang et al., 2021).

3.3.3 Lung Cancer

Lung cancer is the most common malignancy in men and the third most common globally, although its incidence is decreasing (Jones & Baldwin, 2018). Lung cancer is heterogenous as it has various subtypes which differ in terms of risk factors, molecular and genetic profile, response to treatment and prognosis (Borczuk et al., 2009). Around 62–75% of patients present when the tumour is advanced or metastatic (Jones & Baldwin, 2018; Wu et al., 2021). The one-year survival of metastatic lung cancer is below 40%. Screening and early diagnosis of lung cancer using plain x-ray, low-dose CT scan, MRI and PET/CT significantly improves the chance of

cure (Tunali et al., 2021). Sampling of mediastinal lymph nodes is required in some cases (Tunali et al., 2021). Surgery is the mainstay of treatment for lung cancer, but curative radiotherapy is effective for early disease in non-operable patients (Jones & Baldwin, 2018). Radiofrequency and microwave ablation are beneficial for irresectable diseases (Jones & Baldwin, 2018). Selected patients with limited oligometastatic lung cancer may benefit from a trial of curative treatment (Mentink et al., 2021). Artificial intelligence can predict the response to chemotherapy, targeted therapy or immunotherapy Tunali et al., 2021).

3.3.4 Cancer of the Cervix

Cancer of the cervix remains the most common cancer in some of the LMICs and its rate is increasing despite the expansion of the screening program using Pap smears and vaccination programs against the human papillomavirus (HPV) (Xue et al., 2020). Over 85% of mortalities associated with cancer of the cervix occur in LMICs (Bedell et al., 2020; Hu et al., 2019). Prevention and treatment of cancer are labour intensive and are reliant on the availability of personnel, expertise and specific equipment (Baleydier et al., 2021; Holmstrom et al., 2021; Hu et al., 2019). Interpretation of Pap smear is subjective and based on experience that may lead to excessive referral of patients for colposcopy (Xue et al., 2020). The interpretation and reporting of findings following colposcopy are subjective and rely on the level of experience of the practitioner (Xue et al., 2020). Screening for high-risk HPV serotypes is more sensitive than a pap smear but its accuracy unless its evaluation and interpretation are automated.

Deep learning was recently introduced into the screening program for cancer of the cervix including digitalization of glass slides, whole-slide scanning and automated evaluation of pap smears such as the CYTOREADER (Hu et al., 2019). The other addition to the screening program is using dual stained slides for p16/Ki-67 for the identification of high-risk intra-epithelial neoplasm or early advanced cancer followed by scanning of the entire slide for regions of interest and automated evaluation of Pap smear slides (Bedell et al., 2020). Artificial intelligence-guided digital colposcopy improves diagnostic accuracy and expedites treatment (Hu et al., 2019). Smart phone-based computer-aided screening and diagnosis of cancer of the cervix is effective and feasible (Baleydier et al., 2021). Another potential role of AI in the selection of treatment and/or prediction of tumour response in patients who have oligometastatic, oligo-progressive or oligo-recurrent cancer of the cervix (Mangaj et al., 2021).

3.3.5 Gastric Cancer

Adenocarcinoma of the stomach (gastric cancer) is the third to the fifth most common cause of cancer deaths in the world. Gastric cancer is highly heterogeneous

and is classified based on the age of onset, underlying cause, stage at presentation, histological finding, molecular subtype and extent of distant metastases (Ahadi et al., 2020; Bergquist et al., 2019; Birkman et al., 2017; Chevallay et al., 2022; Machlowska et al., 2020; Skierucha 2016; Song et al., 2017). Pathologists' workload, discrepancy, and heterogeneity of the tumour may affect the quality of histopathology reports (Niu et al., 2020). Identification of signet ring cells and their proportion depends on the thoroughness of a pathologist (Pernot et al., 2021). Digital slides and new-generation sequencing have led to an improvement in the accuracy of pathology reports (Niu et al., 2020).

Gastric cancer may be early, locally advanced or metastatic. The prognosis of gastric cancer depends on the stage at presentation and, the histological, genomic and molecular subtype of the tumour (Machlowska et al., 2020). Early gastric cancer is a tumour that has not invaded beyond the submucosa regardless of the lymph node status. Options for the management of low and intermediate-risk early gastric cancer include endoscopic submucosal excision and endoscopic submucosal resection. High-risk patients and histological features mandate a gastrectomy with appropriate lymph node dissection. The 5-year survival of early gastric cancer following curative surgery is over 90%. Staging laparoscopy with lavage cytology should precede curative resection for locally advanced gastric cancer as peritoneal metastases are sometimes missed. Surgery and chemotherapy for early gastric cancer are associated with 90% 5-year survival, inclusive of the signet ring cell subtype. Prognosis is good even for SRCC. Endoscopic mucosal resection or endoscopic submucosal dissection. Another area where AI may help in the selection of patients for curative treatment in cases of oligometastatic gastric cancer (Chevallay et al., 2022).

3.3.6 Prostate Cancer

Prostate cancer is among the six most commonly occurring cancers and the third most common cause of cancer-related mortality in men globally (Liberini et al., 2022). The diagnosis of prostate cancer requires a needle biopsy that can be finger-, ultrasound or MRI-guided (Kasivisvanathan et al., 2018). Prostate cancer is heterogeneous, and its subtypes include indolent, early, locally advanced and metastatic (Haffner et al., 2021; Mateo et al., 2020). The age of the patient, Gleason score, tumour microenvironment, genetic landscape and extent of distance metastases can influence the behaviour of prostate cancer (Foster et al., 2019; Haffner et al., 2021; Mateo et al., 2020). The microenvironment and behaviour of prostate cancer are highly variable as it changes as the tumour grows, and extensive heterogeneity occurs even within an individual tumour (Liberini et al., 2022; Wang et al., 2018).

Options for the management of prostate cancer include observation, brachytherapy or radical prostatectomy with or without adjuvant chemotherapy (Huang et al., 2018). The use of anti-androgen (castration) therapy, chemotherapy, targeted therapy or immunotherapy depends on the molecular sub-types of the tumour (Mateo et al., 2020). Application of AI program combining clinical, radiomics, genomics

and epigenomics is useful for screening, predicting tumour grade and Gleason score, detection of pelvic lymph node and distant metastases using PET/CT, and prognostication of prostate cancer (Huang et al., 2018; Liberini et al., 2022). Additionally, a multi-parameter CT scan can accurately predict the Gleason score of prostate cancer (Liberini et al., 2022).

3.3.7 Malignant Melanoma

Although malignant melanoma is relatively rare compared to other skin cancers, it is both the most aggressive and most fatal of the primary skin tumours. It however among the top 10 most commonly occurring cancers in some countries, including the USA (Bobos, 2021). The types of cutaneous malignant melanoma include superficial spreading, nodular, lentigo malignant melanoma and acral lentiginous melanoma (Bobos, 2021). The prognosis of melanoma is good with expected 5-year survival above 90% if is detected early. Deep learning-guided dermoscopy using smartphones can accurately confirm the malignant change in a skin lesion (Phillips et al., 2019). Malignant melanoma is a highly heterogeneous disease (Bobos, 2021).

The prognosis of a patient who has malignant melanoma is influenced by gender, age, site of the tumour, Breslow's thickness, presence and number of mitotic counts, evidence of tumour ulceration, lymphovascular infiltration, neurotropism, presence of tumour infiltration lymphocytes, the existence of microscopic or macroscopic satellites, in-transit metastases and lymph node or distant metastases (Beasley, 2020; Bobos, 2021). Management of patients diagnosed with metastatic malignant melanoma is dependent on the stage at presentation and the molecular profile of the tumour (Eroglu et al., 2020). Mortality of patients presenting with stage 3 or 4 melanoma remains high with variable responses to targeted and immunomodulation therapy. A combination of PET/CT is useful for monitoring the response of metastatic melanoma to targeted therapy and immunomodulation treatment (Filippi et al., 2022). Artificial intelligence program using a combination of demographic, clinical, radiomics, genomics, epigenomic and proteomics data is used for early diagnosis, screening for metastases and prediction of response to treatment and overall prognosis (Eroglu et al., 2020; Filippi et al., 2022).

3.3.8 Ovarian Cancer

Ovarian cancer includes stromal, germ cell and epithelial tumours. Epithelial ovarian cancer accounts for 90% of malignancies of the ovary and is subdivided into Type 1 and Type 2 (Kroeger & Drapkin, 2017). Type 2 epithelial tumours are fast-growing and the most aggressive with a reduced chance of cure. The symptoms of ovarian malignancies are non-specific and around 70% of patients have an irresectable disease at presentation (Matulonis et al., 2020). Management of advanced ovarian

cancer includes de-bulking followed by paclitaxel or platinum-based chemotherapy (Matulonis et al., 2020). Less than 50% of ovarian cancer respond to paclitaxel or platinum-based chemotherapy (Ma et al., 2021). Deep learning program using CNN can predict the treatment response of ovarian cancer by combining demographics, clinic-pathological, radiomics, biomarkers and genomics to predict histology of the tumour, tumour stage, response to chemotherapy, the likelihood of recurrence and overall prognosis (Gong et al., 2020). Radiomics based on CT and MRI can accurately diagnose the type of ovarian cancer and its histological grade, presence of peritoneal metastases and predict response to treatment (Wang et al., 2022; Zhou et al., 2022). The model for predicting tumour response and the overall outcome combines age, BMI, neutrophils and lymphocytes count and the level of CRP, albumin, fibrinogen, CA-125, CEA, α- fetoprotein and circulating tumour cells (CTC) (Ma et al., 2021).

3.3.9 Hepatocellular Carcinoma

Hepatocellular carcinoma is increasing in incidence globally (Khemlina et al., 2017). Risk factors of HCC include Hepatitis B and C infection, liver cirrhosis, non-alcoholic steatohepatitis and obesity (Khandekar et al., 2011; Khemlina et al., 2017). Majority of patients who have HCC present late when the tumour is neither resectable nor suitable for liver transplant (Khemlina et al., 2017). The diagnosis of HCC relies on the level of α-fetoprotein, imaging or needle biopsy (Zhou et al., 2019). Needle biopsy is rarely necessary for the evaluation of a liver mass and all HCCs are associated with an elevation of α-fetoprotein. Although diagnosis based on CT and/or MRI findings is usually relied on, some of the HCCs do not show classical findings. Hepatocellular carcinoma may be confused with intrahepatic cholangiocarcinoma, liver metastasis or focal nodular hyperplasia (Perez & Grande, 2020). Furthermore, the interpretation of CT or MRI of the liver depends on the expertise that may not be available in under-resourced settings (Honsy et al., 2018; Liu et al., 2019).

The use of radiomics can provide an accurate classification of liver lesions according to the Liver Imaging Reporting and Data System (Feng et al., 2021). Radiomic can preoperatively predict the histological grade, Ki-67 index, the existence of microvascular invasion and molecular profile of HCC (Feng et al., 2021). The other use of radiomics is in the immune profiling of HCC in preparation for immunotherapeutic agents (Moldogazieva et al., 2021). A combination of radiomics and α-fetoprotein levels can prognosticate HCC. Integration of radiomics, genomics, epigenomics, transcriptomics and proteomics is more accurate for the prediction of tumour response or recurrence following resection and selection of options for palliative treatment (Perez & Grande, 2020; Zhang et al., 2020b). Among the available palliative, options are TACE, RFA and sorafenib (Zhang et al., 2020b).

3.3.9.1 *Carcinoma of Oesophagus*

Carcinoma of the oesophagus is among the top eight most commonly occurring malignancies (Huang et al., 2020). The main subtypes of carcinoma of the oesophagus are squamous cell carcinoma and adenocarcinoma. The incidence of adenocarcinoma of the oesophagus is increasing. Majority of patients present when the tumour is advanced and the expected five-year survival is less than 20% malignancies (Huang et al., 2020). Screening of individuals who have premalignant lesions leads to early diagnosis and an increased chance of cure. The performance and interpretation of screen results of screening test is however laborious and subjective. Potential benefits of AI have been demonstrated in the entire pathway from screening to provisioning of personalised care of individuals who have oesophageal cancer (Syed et al., 2020). Image-guided AI can accurately diagnose early cancer including the ability to perform endocytoscopy and virtual biopsy (Syed et al., 2020). Furthermore, AI can predict the presence of local and distant metastasis, treatment planning, response to treatment and prognostication (Syed et al., 2020; Zhang et al., 2020a). Application of AI allows for a comprehensive assessment of the microenvironment and genomics of the tumour personalised care of patients who have oesophageal cancer (Visaggi et al., 2021; Zhao et al., 2021.

3.3.10 Pancreatic Adenocarcinoma

Although pancreatic adenocarcinoma (PDAC) is rare and its rate is increasing. Risk factors of PDAC include smoking, chronic pancreatitis and familial. Treatment options include surgical resection, chemotherapy and targeted therapy. More than 80% of patients with PDAC present when the tumour is not resectable (Pereira et al., 2020). The tumour microenvironment and genetic landscape of PDAC are highly variable (Hayashi et al., 2021). The prognosis of PDAC remains decimal despite an improvement in surgical expertise and current advances in the understanding of the genomics, epigenetics and the microenvironment of pancreatic adenocarcinoma (Waddell et al., 2015). The application of AI for the screening of high-risk individuals has generated hope for early diagnosis and cure for patients who have PDAC (Kenner et al., 2021). Artificial intelligence is also helpful for staging, prediction of response to treatment and prognostication of pancreatic cancer (Hayashi et al., 2021).

3.3.11 Other Cancers

The beneficial use of AI extends to carcinoma of the uterus (Ravegnini et al., 2022), urinary bladder (Malinaric et al., 2022), renal cell carcinoma (Peng et al., 2021) and

gastrointestinal stromal tumour (GIST) (Yang et al., 2020). Artificial intelligence is useful for screening and early diagnosis of basal cell carcinoma of the skin (Santilli et al., 2020), oesophageal carcinoma, thyroid cancer, neuroendocrine tumours (Clift et al., 2020), phyllodes tumour of the breast (Niu et al., 2021; Rayzah et al., 2020), cholangiocarcinoma, dermatofibrosarcoma protuberans (Li et al., 2021b) and haematological malignancies (El Alaoui et al., 2022; Salama et al., 2020). Another area of the potential benefit of AI in oncology is radiomics and virtual biopsy for cholangiocarcinoma (Yang & Shu, 2021), gall bladder cancer (Jeong et al., 2020), primary and metastatic brain tumour and neuroblastoma (Giglia et al., 2020; Rudie et al., 2019), soft tissue sarcoma (Gitto et al., 2021; Vibhakar et al., 2021) and adrenal tumours (Liu et al., 2022a). Artificial intelligence provides more accurate information for treatment planning in head and neck cancer (Kearney et al., 2018; Mahmood et al., 2021; Resteghini et al., 2018; van Dijk & Fuller, 2021) and malignant tumours in paediatric patients (Vo et al., 2020) and determination of the adequacy of excision margin for basal cell carcinoma (Santilli et al., 2020).

Additional benefits of AI and its potential use for precision medicine is in the prediction of fibrosarcomatous variant and the possible presence of metastases dermatofibrosarcoma protuberans (Li et al., 2021b) and lung metastases in patients who have papillary carcinoma of the thyroid (Zhao et al., 2019; Liu et al., 2022b). Patients who have head and neck (Shen et al., 2019), neuroendocrine tumours (Partouche et al., 2021, sarcomas (Tian et al., 2018; Zhu et al., 2019; Gitto et al., 2021; Li et al., 2021a), pheochromocytoma (Buffet et al., 2019; Guo et al., 2020) and thyroid cancer Yoon et al., 2020) can benefit from AI as it can predict the response to treatment. Similarly, AI is useful for the prediction of side effects and prognostication of patients who have head and neck cancers (Kearney et al., 2018). Table 3.1 is a summary of areas of potential benefits of AI in 27 commonly diagnosed solid and hemopoietic malignancies.

Collection, collation, analysis and matching with the existing AI developed prediction models improve the ability to deliver precision medicine (Fig. 3.5).

3.4 Limitations

The application of AI is resource intensive as it requires advanced computing which may not be available in LMICs. The resources or expertise to offer the treatment recommended following the application of AI may also not be available in LMICs. The big data that is available for use to generate algorithms for decisions are likely to have been sourced from HICs. The robustness of the algorithms is likely to be influenced by the quality of the data which has been fed into the computer. The size of a sample that is used for training and validation of the treatment algorithms might not have been adequate or appropriate. A possibility of either false positives or false negatives exists. Deep learning especially CNN is beyond the capability of a human brain. A computer is not able to explain how a diagnosis was made or why a certain

Table 3.1 The current scope of AI with potential benefits for precision medicine in 27 solid and haematopoietic malignancies

Cancer type	Screening	Early Diagnosis	Treatment response assessment	Prediction of tumour recurrence	Prognostication
Breast cancer	√	√	√	√	√
Cancer of cervix	√	√	√	√	√
Uterine malignancies	√	√	√	√	√
Colorectal carcinoma	√	√	√	√	√
Gastric cancer	√	√	√	√	√
Lung cancer	√	√	√	√	√
Prostate cancer	√	√	√	√	√
Head & neck cancer	√	√	√	√	√
Oesophageal carcinoma	√	√	√	√	√
Hepatocellular carcinoma	√	√	√	√	√
Melanoma	√	√	√	√	√
Ovarian cancer	√	√	√	√	√
Neuroendocrine tumours	√	√	√	√	√
Basal cell carcinoma	√	√	√	√	–
Pancreatic adenocarcinoma	√	√	√	√	√
Oesophagus	√	√	√	√	√
Thyroid	√	√	√	√	√
Soft tissue sarcoma	–	√	√	√	√
Dermatofibrosarcoma protuberans	–	√	√	√	√
Gastrointestinal stromal tumour	√	√	√	√	√
Lymphoma	√	√	√	√	√
Leukaemia	√	√	√	√	√
Multiple myeloma	√	√	√	√	√
Nephroblastoma	–	√	√	√	√
Neuroblastoma	–	√	√	√	√
Osteogenic sarcoma	–	√	√	√	√
Adrenal tumours	√	√	√	√	√

treatment would be the most ideal, the so-called "black-boxes" of AI. Decisions which are guided by AI may not be defensible in case of litigation. The application of AI to guide personalised may threaten confidentiality and the ability to receive informed consent and therefore autonomy of patients.

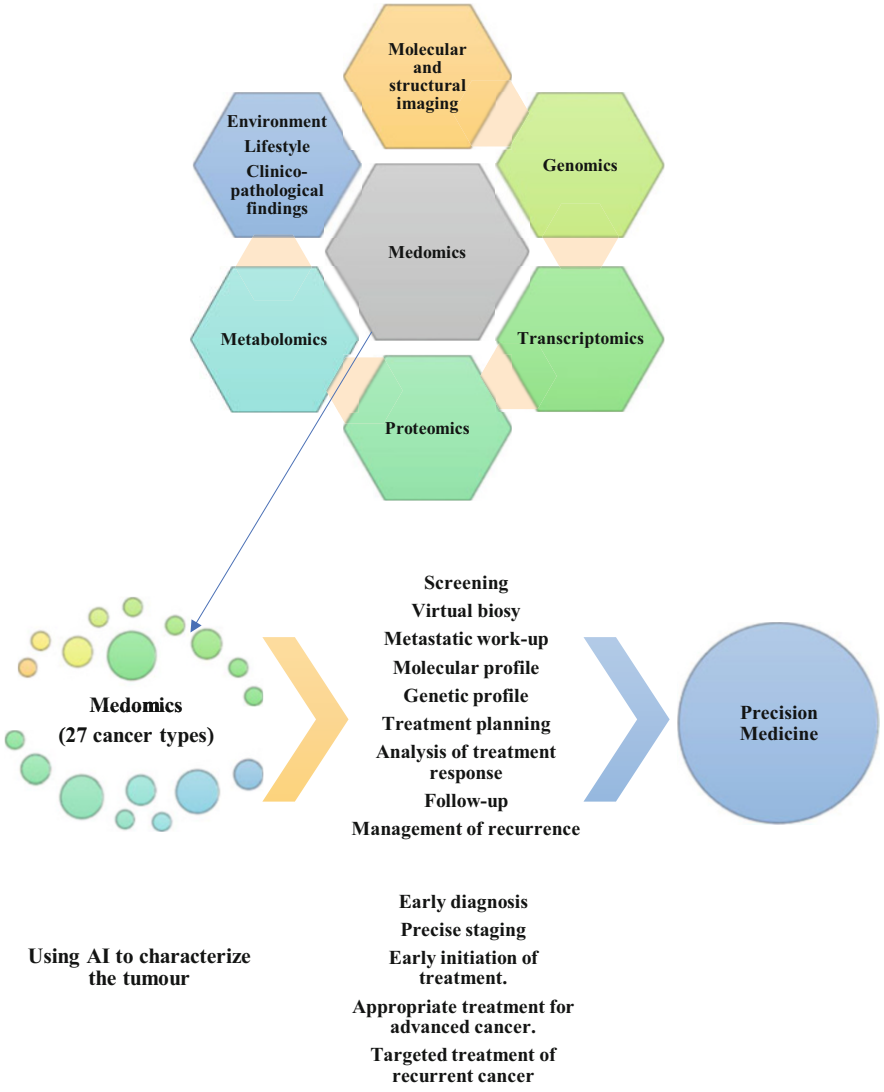


Fig. 3.5 Illustration of the linkage of medomics, AI benefits and precision medicine to enable early intervention and tailoring better-personalised cancer treatments

3.5 Conclusion

The traditional symptoms and TNM decisions are archaic and lead to either under or over-treatment of cancer. The genomic and epigenomics in the microenvironment of cancer makes it a heterogenous disease within itself and in different persons. The need for personalised medicine is more relevant in oncology than any other field in

medicine and related professions in the healthcare industry. The last 10 years have seen a dramatic in the uptake and use of AI in oncology the workup and treatment of almost all malignancies in humans. The use of AI extends from the screening of common malignancies or high-risk individuals to allow for early diagnosis with confidence and exclusion of differentials. The ability to combine demographic, clinic-pathological, radiological and molecular profiles in a patient allows for a more enhanced prediction of the likelihood of metastasis, tumour grade, response to treatment, the possibility of tumour recurrence and long-term survival. Therefore, AI in oncology is likely to facilitate the delivery of precision medicine in the management of cancer.

References

Ahadi, M., Moradi, A., Musavinejad, L., Movafagh, A., & Moradi, A. (2020). The expression of p53, CD44, Ki-67, and HER-2/neu markers in gastric cancer and its association with histopathological indicators: A retrospective study. *Asian Pacific Journal of Cancer Prevention, 21*(6), 1607–1614. https://doi.org/10.31557/APJCP.2020.21.6.1607

Ahmed, Z., Zeeshan, S., Foran, D. J., Kleinman, L. C., Wondisford, F. E., & Dong, X. (2021). Integrative clinical, genomics and metabolomics data analysis for mainstream precision medicine to investigate COVID-19. *BMJ Innovations, 7*, 6–10. https://doi.org/10.1136/bmjinnov-2020-000444

Aigner, F., Pratschke, J., & Schmelzle, M. (2017). Oligometastatic disease in colorectal cancer – How to proceed? *Visceral Medicine, 33*, 23–28. https://doi.org/10.1159/000454688

Alqahtani, A., Khan, Z., Alloghbi, A., Ahmed, T. S. S., Ashraf, M., & Hammouda, D. M. (2019). Hepatocellular carcinoma: Molecular mechanisms and targeted therapies. *Medicina, 55*, 526. https://doi.org/10.3390/medicina55090526

AlSendi, M., O'Reilly, D., Zeidan, Y. H., & Kelly, C. M. (2021). Oligometastatic breast cancer: Are we there yet? *International Journal of Cancer, 149*(8), 1520–1528. https://doi.org/10.1002/ijc.33693

Baleydier, I., Vassilakos, P., Vinals, R., Wisniak, A., Kenfack, B., Fouogue, J. T., et al. (2021). Study protocol for a two-site clinical trial to validate a smartphone-based artificial intelligence classifier identifying cervical precancer and cancer in HPV-positive women in Cameroon. *PLoS One, 16*, e0260776. https://doi.org/10.1371/journal.pone.0260778

Beasley, G. M. (2020). Acral melanomas of the sole may have worse prognosis compared to other sites of acral melanoma. *Annals of Surgical Oncology, 27*(9), 3121–3122. https://doi.org/10.1245/s10434-020-08460-3

Beckmann, J. S., & Lew, D. (2016). Reconciling evidence-based medicine and precision medicine in the era of big data: Challenges and opportunities. *Genome Medicine, 8*, 134. https://doi.org/10.1186/s13073-016-0388-7

Bedell, S. L., Goldstein, L. S., Goldstein, A. R., & Goldstein, A. T. (2020). Cervical cancer screening: Past, present, and future. *Sexual Medicine Reviews, 8*, 28–37. https://doi.org/10.1016/j.sxmr.2019.09.005

Bedrikovetski, S., Dudi-Venkata, N. N., Kroon, H. M., Seow, W., Vather, R., Carneiro, G., et al. (2021). Artificial intelligence for pre-operative lymph node staging in colorectal cancer: A systematic review and meta-analysis. *BMC Cancer, 21*, 1058. https://doi.org/10.1186/s12885-021-08773-w

Bergquist, J. R., Leiting, J. L., Habermann, E. B., Clearly, S. P., Kendrick, M. L., Smoot, R. L., et al. (2019). Early-onset gastric cancer is a distinct disease with worrisome trends and oncogenic features. *Surgery, 166*, 547–555. https://doi.org/10.1016/j.surg.2019.04.036

Bhinder, B., Gilvary, C., Madhukar, N. S., & Elemento, O. (2021). Artificial intelligence in cancer research and precision medicine. *Cancer Discovery, 11*(4), 900–915. https://doi.org/10.1158/2159-8290.CD-21-0090

Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., et al. (2019). Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA: A Cancer Journal for Clinicians, 69*, 127–157. https://doi.org/10.3322/caac.21552

Birkman, E., Mansuri, N., Kurki, S., Ålgars, A., Lintunen, M., et al. (2017). Gastric cancer: Immunohistochemical classification of molecular subtypes and their association with clinicopathological characteristics. *Virchows Archiv, 472*(3), 369–382.

Bobos, M. (2021). Histopathologic classification and prognostic factors of melanoma: A 2021 update. *Italian Journal of Dermatology and Venereology, 156*(3), 300–321. https://doi.org/10.23736/S2784-8671.21.06958-3

Borczuk, A. C., Toonkel, R. L., & Powell, C. A. (2009). Genomics of lung cancer. *Proceedings of the American Thoracic Society, 6*, 152–158. https://doi.org/10.1513/pats.200807-076LC

Bourcier, S., Klug, J., & Nguyen, L. S. (2021). Non-occlusive mesenteric ischemia: Diagnostic challenges and perspectives in the era of artificial intelligence. *World Journal of Gastroenterology, 27*(26), 4088–4103. https://doi.org/10.3748/wjg.v27.i26.4088

Buffet, A., Ben Aim, L., Leboulleux, S., Drui, D., Vezzosi, D., Libe, R., et al. (2019). Positive impact of genetic test on the management and outcome of patients with paraganglioma and/or pheochromocytoma. *The Journal of Clinical Endocrinology and Metabolism, 104*(4), 1109–1118. https://doi.org/10.1210/jc.2018-02411

Canzoneri, R., Lacunza, E., & Abba, M. C. (2019). Genomics and bioinformatics as pillars of precision medicine in oncology. *Medicina (Buenos Aires), 79*(6/1), 587–592.

Chevallay, M., Wassmer, C. H., Iranmanesh, P., Jung, M. K., & Monig, S. P. (2022). Multimodal treatment in oligometastatic gastric cancer. *World Journal of Gastrointestinal Oncology, 14*(2), 434–449. https://doi.org/10.4251/wjgo.v14.i2.434

Clift, A. K., Kidd, M., Bodei, L., Toumpanakis, C., Baum, R. P., Oberg, K., et al. (2020). Neuroendocrine neoplasms of the small bowel and pancreas. *Neuroendecrinology, 110*(6), 444–476. https://doi.org/10.1159/000503721

Connor, C. W. (2019). Artificial intelligence and machine learning in anesthesiology. *Anesthesiology, 131*(6), 1346–1359. https://doi.org/10.1097/ALN.0000000000002694

Dayde, D., Tanaka, I., Jain, R., Tai, M. C., & Taguchi, A. (2017). Predictive and prognostic molecular biomarkers for response to neoadjuvant chemoradiation in rectal cancer. *International Journal of Molecular Sciences, 18*, 573. https://doi.org/10.3390/ijms18030573

Dlamini, Z., Francies, F. Z., Hull, R., & Marima, R. (2020). Artificial intelligence (AI) and big data in cancer and precision oncology. *Computational and Structural Biotechnology Journal, 18*, 2300–2311. https://doi.org/10.1016/j.csbj.2020.08.019

El Alaoui, Y., Elomri, A., Qaraqe, M., Padmanabhan, R., Taha, R. Y., El Omri, H., et al. (2022). A review of artificial intelligence applications in hematology management: Current practices and future prospects. *Journal of Medical Internet Research, 24*(7), e36490. https://doi.org/10.2196/36490

Eroglu, Z., Eatrides, J., Naqvi, S. M. H., Kim, Y., Rich, J., Babacan, N. A., et al. (2020). Neoadjuvant BRAF-targeted therapy in regionally advanced and oligometastatic melanoma. *Pigment Cell & Melanoma Research, 33*(1), 86–95. https://doi.org/10.1111/pcmr.12813

Feng, B., Ma, X. H., Wang, C. W., Liu, X. B., & Zhao, X. M. (2021). Application of artificial intelligence in preoperative imaging of hepatocellular carcinoma: Current status and future perspectives. *World Journal of Gastroenterology, 27*(32), 5341–5350. https://doi.org/10.3748/wjg.v27.i32.5341

Filippi, L., Bianconi, F., Schillaci, O., Spanu, A., & Palumbo, B. (2022). The role and potential of ^{18}F-FDG PET/CT in malignant melanoma: Prognostication, monitoring response to targeted and immunotherapy, and radiomics. *Diagnostics, 12*, 929. https://doi.org/10.3390/diagnostics120409929

Foster, C. C., Weichselbaum, R. R., & Pitroda, S. P. (2019). Oligometastatic prostate cancer: Reality or figment of imagination? *Cancer, 125*, 340–352. https://doi.org/10.1002/cncr.31860

Garrido-Castro, A. C., Lin, N. U., & Polyak, K. (2019). Insights into molecular classifications of triple-negative breast cancer: Improving patient selection for treatment. *Cancer Discovery, 9*(2), 176–198. https://doi.org/10.1158/2159-8290.CD-18-1177. (Heterogeity of cancer. Subtypes of breast cancer. Even subtypes of subtypes). luminal A, luminal B, HER2-enriched, basal-like, normal breast-like group.

Geras, K. J., Mann, R. M., & Moy, L. (2019). Artificial intelligence for mammography and digital tomosynthesis: Current concepts and future perspectives. *Radiology, 293*, 246–259. https://doi.org/10.1148/radiol.2019182627

Giglia, G., Gambino, G., & Sardo, P. (2020). Through predictive personalized medicine. *Brain Sciences, 10*, 594. https://doi.org/10.3390/brainsci10090594

Gitto, S., Cuocolo, R., Albano, D., Morelli, F., Pescatori, L. C., Messina, C., et al. (2021). CT and MRI radiomics of boneand soft-tissue sarcomas: A systematic review and reproducibility and validation strategies. *Insights Into Imaging, 12*, 68. https://doi.org/10.1186/s13244-021-01008-3

Gong, G., Lin, T., & Yuan, Y. (2020). Integated analysis of gene expression and DNA methylation profiles in ovarian cancer. *Journal of Ovarian Research, 13*, 30. https://doi.org/10.1186/s13048-020-00632-9

Guo, J., Gu, Y., Ma, X., Zhang, L., Li, H., Yan, Z., et al. (2020). Identification of hub genes and pathways in adrenocortical carcinoma by integrated bioinformatics analysis. *Journal of Cellular and Molecular Medicine, 24*, 4428–4438. https://doi.org/10.1111/jcmm.15102

Haffner, M. C., Zwart, W., Roudier, M. P., True, L. D., Nelson, W. G., Epstein, J. I., De Marzo, A. M., et al. (2021). Genomic and phenotypic heterogeneity in prostate cancer. *Nature Reviews. Urology, 18*(2), 79–92. https://doi.org/10.1038/s41585-020-00400-w

Hayashi, H., Uemura, N., Matsumura, K., Zhao, L., Sato, H., Shiraishi, Y., et al. (2021). Recent advances in artificial intelligence for pancreatic ductal adenocarcinoma. *World Journal of Gastroenterology, 2743*, 7480–7496. https://doi.org/10.3748/wjg.v27.i43.7480

He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of artificial intelligence in medicine. *Nature Medicine, 25*(1), 30–36. https://doi.org/10.1038/s41591-018-0307-0

Holmstrom, O., Linder, N., Kaingu, H., Mbuuko, N., Mbete, J., Kinyua, F., et al. (2021). Point-of care digital cytology with artificial intelligence for cervical cancer screening in a resource-limited setting. *JAMA Network Open, 4*(3), e211740. https://doi.org/10.1001/jamanetworkopen.2021.1740

Honsy, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. (2018). Artificial intelligence in radiology. *Nature Reviews. Cancer, 18*(8), 500–510. https://doi.org/10.1038/s41568-018-0016-5

Hu, L., Bell, D., Antani, S., Xue, Z., Yu, K., Horning, M. P., et al. (2019). An observational study of deep learning and automated evaluation of cervical images for cancer screening. *Journal of the National Cancer Institute, 111*(9), djy225. https://doi.org/10.1093/jnci/djy225

Huang, L. M., Yang, W. J., Huang, Z. Y., Tang, C. W., & Li, J. (2020). Artificial intelligence technique in detection of early esophageal cancer. *World Journal of Gastroenterology, 26*(39), 5959. https://doi.org/10.3748/wjg.v26.i39.5959

Huang, S., Yin, L., Yue, J. I., Li, Y., Yang, Y., & Lin, Z. (2018). Direct comparison of choline PET/CT and MRI in the diagnosis of lymph node metastases in patients with prostate cancer. *Medicine, 97*, 50(e13344). https://doi.org/10.1097/MD.0000000000013344

Ibrahim, A., Gamble, P., Jaroensri, R., Abdelsamea, M. M., Mermel, C. H., Chen, P. H. C., et al. (2020). Artificial intelligence in digital breast pathology: Techniques and applications. *The Breast, 49*, 267–273. https://doi.org/10.1016/j.breast.2019.12.007

Jeong, Y., Kim, J. H., Chae, H. G., Park, S. J., Bae, J. S., & Joo, I. (2020). Deep learning-based decision support system for the diagnosis of neoplastic gallbladder polyps on ultrasonograpy: Preliminary results. *Scientific Reports, 10*, 7700. https://doi.org/10.1038/s41598-020-64205-y

Jones, G. S., & Baldwin, D. R. (2018). Recent advances in the management of lung cancer. *Clinical Medicine, 18*(2), s41–s46.

Kasivisvanathan, V., Ranniko, A. S., Borghi, M., Panebianco, V., Mynderse, L. A., Vaarala, M. G. H., et al. (2018). MRI-targeted or standard biopsy for prostate-cancer diagnosis. *The New England Journal of Medicine, 378*(19), 1767–1177. https://doi.org/10.1056/NEJMoa1801993

Kather, J. N., Halama, N., & Jaeger, D. (2018). Genomics and emerging biomarkers for immunotherapy of colorectal cancer. *Seminars in Cancer Biology, 52*, 189–197. https://doi.org/10.1016/j.semcancer.2018.02.010

Kearney, V., Chan, J. W., Valdes, G., Solberg, T. D., & Yom, S. S. (2018). The application of artificial intelligence in the IMRT planning process for head and neck cancer. *Oral Oncology, 87*, 111–116.

Kenner, B., Chari, S. T., Kelsen, D., Klimstra, D. S., Pandol, S. J., Rosenthal, M., et al. (2021). Artificial intelligence and early detection of pancreatic cancer. *Pancreas, 50*, 251–279. https://doi.org/10.1097/MPA.0000000000001762

Keshinbora, K., & Guven, F. (2020). Artificial intelligence and ophthalmology. *Turkish Journal of Ophthalmology, 50*, 37–43. https://doi.org/10.4274/tjo.galenos.2020.78989

Khandekar, M. J., Cohen, P., & Spiegelman, B. M. (2011). Molecular mechanisms of cancer development in obesity. *Nature Reviews. Cancer, 11*(12), 886–895. https://doi.org/10.1038/nrc3174

Khemlina, G., Ikeda, S., & Kurzrock, R. (2017). The biology of GHepatocellular carcinoma: Implications for genomic and immune therapies. *Molecular Cancer, 16*, 149. https://doi.org/10.1186/s12943-017-0712-x

Kitahara, C. M., & Sosa, J. A. (2016). The changing incidence of thyroid cancer. *Nature Reviews Endocrinology, 12*, 646. https://www.nature.com/articles/nrendo.2016.110

Konig, I. R., Fuchs, O., Hansen, G., von Mutius, E., & Kopp, M. V. (2017). What is precision medicine? *The European Respiratory Journal, 50*, 1700391. https://doi.org/10.1183/13993003.00391-2017

Kroeger, P. T., & Drapkin, R. (2017). Pathogenesis and heterogeneity of ovarian cancer. *Current Opinion in Obstetrics & Gynecology, 29*, 26–34. https://doi.org/10.1097/GCO.0000000000000340

Kroner, P. T., Engels, M. M. L., Glicksberg, K. W., Johnson, K. W., Mzaik, O., van Hooft, J. E., et al. (2021). Artificial intelligence in gastroenterology: A state-of-the-art review. *World Journal of Gastroenterology, 27*(40), 6794–6824. https://doi.org/10.3748/wjg.v27.i40.6794

Lerebours, F., Cabel, L., & Pierga, J. Y. (2021). Neoadjuvant endocrine therapy in breast cancer management: State of the art. *Cancers, 13*, 902. https://doi.org/10.3390/cancers13040902

Li, B., Fang, L., Wang, B., Yang, Z., & Zhao, T. (2021a). Identification of prognostic RBPs in osteosarcoma. *Technology in Cancer Research & Treatment, 20*, 1–11. https://doi.org/10.1177/15330338211004918

Li, Y., Liang, J., Xu, X., Jiang, X., Wang, C., Chen, S., et al. (2021b). Clinicopathological features of fibrosarcomatous dermatofibrosarcoma protuberans and contruction of a back-propagation neural network recognition model. *Orphanet Journal of Rare Diseases, 16*, 48. https://doi.org/10.1186/s13023-021-01698-4

Liberini, V., Laudicella, R., Balma, M., Nicolotti, D. G., Buschiazzo, A., Grimaldi, S., et al. (2022). Radiomics and artificial intelligence in prostate cancer: New tools for molecular hybrid imaging and theragnostics. *European Radiology Experimental, 6*, 27. https://doi.org/10.1186/s41747-022-00282-0

Lin, C., Sun, Y. D., Yu, G. Y., Cui, J. R., Lou, Z., Zhang, H., et al. (2020). Integrated omics of metastatic colorectal cancer. *Cancer Cell, 38*(5), 734–747. https://doi.org/10.1016/j.ccell.2020.08

Liu, H., Guan, X., Xu, B., Zeng, F., Chen, C., Yin, H. I., et al. (2022a). Computed tomography-based machine learning differentiates adrenal pheochromocytoma from lipid-poor adenoma. *Frontiers in Endocrinology, 13*, 833413. https://doi.org/10.3389/fendo.2022.8333413

Liu, Z., Meng, X., Zhang, H., Li, Z., Liu, J., Sun, K., et al. (2020). Predicting distant metastasis and chemotherapy benefit in locally advanced rectal cancer. *Nature Communications, 11*, 4308. https://doi.org/10.1038/s41467-020-18162-9

Liu, Z., Wang, S., Dong, D., Wei, J., Fang, C., Zhou, X., et al. (2019). The applications of radiomics in precision diagnosis and treatment of oncology: Opportunities and challenges. *Theranostics, 9*(5), 1303–1322. https://doi.org/10.7150/thno.30309

Liu, W., Wang, S., Ye, Z., Xu, P., Xia, X., & Guo, M. (2022b). Prediction of lung metastases in thyroid cancer using machine learning based on SEER database. *Cancer Medicine, 11*, 2503–2515. https://doi.org/10.1002/cam4.4617

Low, S. K., Zembutsu, H., & Nakamura, Y. (2018). Breast cancer: The translation of big genomic data to cancer precision medicine. *Cancer Science, 109*, 497–506. https://doi.org/10.1111/cas.13463

Luchini, C., Pea, A., & Scarpa, A. (2022). Artificial intelligence in oncology: Current applications and future perspectives. *British Journal of Cancer, 126*, 4–9. https://doi.org/10.1038/s41416-021-01633-1

Ma, J., Yang, J., Jin, Y., Cheng, S., Huang, S., Zhang, N., et al. (2021). Artificial intelligence based on blood biomarkers including CTCs predicts outcomes in epithelial ovarian cancer: A prospective study. *Oncotargets and Therapy, 14*, 3267–3280.

Machlowska, J., Baj, J., Sitarz, M., Maciejewski, R., & Sitarz, R. (2020). Gastric cancer: Epidemiology, risk factors, classification, genomic characteristics and treatment strategies. *International Journal of Molecular Sciences, 21*, 4012. https://doi.org/10.3390/ijms21114012

Maclean, D., Tsakok, M., Gleeson, F., Breen, D. J., Goldin, R., Primrose, J., et al. (2021). Comprehensive imaging characterization of colorectal liver metastases. *Frontiers in Oncology, 11*, 730854. https://doi.org/10.3389/fonc.2021.730854

Mahmood, H., Shaban, M., Rajpoot, N., & Khurram, S. A. (2021). Artificial intelligence-based methods in head and neck cancer diagnosis: An overview. *British Journal of Cancer, 124*, 1934–1940. https://doi.org/10.1038/s41416-021-01586-x

Malinaric, R., Mantica, G., Lo Monaco, L., Mariano, F., Leonardi, R., Simonato, A., et al. (2022). The role of novel bladder cancer diagnostic and surveillance biomarkers—What should a urologist really know? *International Journal of Environmental Research and Public Health, 19*, 9648. https://doi.org/10.3390/ijerph19159648

Mamelle, G., Pampurik, J., Luboinski, B., Lancar, R., Lusinchi, A., & Bosq, J. (1994). Lymph node prognostic factors in head and neck squamous cell carcinomas. *American Journal of Surgery, 168*(5), 494–498. https://doi.org/10.1016/s0002-9610(05)80109-6

Mangaj, A., Chopra, S., & Nout, R. A. (2021). Defining the role of high-dose radiation in oligometastatic & oligorecurrent cervical cancer. *The Indian Journal of Medical Research, 154*, 303–318. https://doi.org/10.4103/ijmr.IJMR_298_21

Marur, S., & Forastiere, A. A. (2008). Head and neck cancer: Changing epidemiology, diagnosis, and treatment. *Mayo Clinic Proceedings, 83*(4), 489–501.

Mateo, J., Seed, G., Bertan, C., Rescigno, P., Dolling, D., Figueiredo, I., et al. (2020). Genomics of lethal prostate cancer at diagnosis and castration resistance. *Journal of Clinical Investigation, 130*(4), 1743–1751. https://doi.org/10.1172/JCI132031

Matulonis, U. A., Sood, A. K., Fallowfield, L., Howitt, B. E., Sehouli, J., & Karlan, B. Y. (2020). Nature reviews. *Disease Primers, 2*, 16061. https://doi.org/10.1038/nrdp.2016.61

Mentink, J. F., Paats, M. S., Dumoulin, W., Cornelissen, R., Elbers, J. B. W., Maat, A. P. W. M., et al. (2021). Defining oligometastatic non-small cell lung cancer: Concept versus biology, a literature review. *Translational Lung Cancer Research, 10*(7), 3329–3338. https://doi.org/10.21037/tlcr-21265

Miller, K. D., Keegan, T. H., Hipp, H. S., Jemal, A., & Siegel, R. L. (2020). Cancer statistics for adolescents and young adults, 2020. *CA: A Cancer Journal for Clinicians, 70*(6), 443–459. https://doi.org/10.3322/caac.21637

Mitsala, A., Tsalikidis, C., Pitiakoudis, M., Simopoulos, C., & Tsaroucha, A. K. (2021). Artificial intelligence in colorectal cancer screening, diagnosis and treatment. A New Era. *Current Oncology, 28*, 1581–1607. https://doi.org/10.3390/curroncol28030149

Moldogazieva, N. T., Mokhosoev, I. M., Zavadskiy, S. P., & Terentiev, A. A. (2021). Proteomics profiling and artificial intelligence for hepatocellular carcinoma translational medicine. *Biomedicine, 9*, 159. https://doi.org/10.3390/biomediines9020159

Nensa, F., Demircioglu, A., & Rischpler, C. (2019). Artificial intelligence in nuclear medicine. *The Journal of Nuclear Medicine, 60*(9), 29S–37S. https://doi.org/10.2967/jnumed.118.220590

Niu, S., Huang, J., Li, J., Liu, X., Wang, D., Wang, Y., et al. (2021). Differential diagnosis between small breast phyllodes tumors and fibroadenomas using artificial intelligence and ultrasound data. *Quantitative Imaging in Medicine and Surgery, 11*(5), 2052–2061. https://doi.org/10.21037/qims-20-919

Niu, P. H., Zhao, L. L., Wu, H. L., Zhao, D. B., & Chen, Y. T. (2020). Artificial intelligence in gastric cancer: Application and future perspectives. *World Journal of Gastroenterology, 26*(36), 5408–5419. https://doi.org/10.3748/wjg.v26.i36.5408

Partouche, E., Yeh, R., Eche, T., Rozenblum, L., Carrere, N., Guimbaud, R., et al. (2021). Updated trends in imaging practices for pancreatic neuroendocrine tumours (PNETs): A systematic review and meta-analysis to pave the way for standardization in the new era of big data and artificial intelligence. *Frontiers in Oncology, 11*, 628408.

Peng, Q., Shen, Y., Fu, K., Dai, Z., Jin, L., Yang, D., et al. (2021). Artificial intelligence prediction model for overall survival of clear cell renal cell carcinoma based on a 21-gene molecular prognostic score system. *Aging, 13*(5), 7361–7381.

Pereira, S. P., Oldfield, L., Ney, A., Hart, P. A., Keane, M. G., Pandol, S. J., et al. (2020). Early detection of pancreatic cancer. *The Lancet Gastroenterology & Hepatology, 5*(7), 698–710. https://doi.org/10.1016/S2468-1253(19)30416-9

Perez, M. J., & Grande, R. G. (2020). Application of artificial intelligence in the diagnosis and treatment of hepatocellular carcinoma: A review. *World Journal of Gastroenterology, 26*(37), 5617–5628. https://doi.org/10.3748/wjg.v26.i37.5617

Pernot, S., Voran, T., Perkins, G., Lagorce-Pages, C., Berger, A., & Taleb, J. (2021). Signet-ring cell carcinoma of the stomach: Impact on prognosis and specific therapeutic challenge. *World Journal of Gastroenterology, 21*(40), 11428–11438. https://doi.org/10.3748/wjg.v21.i40.11428

Phillips, M., Marsden, H., Jaffe, W., Matin, R. N., Wali, G. N., Greenhalgh, J., et al. (2019). Assessment of accuracy of an artificial intelligence algorithm to detect melanoma in images of skin lesions. *JAMA Network Open, 2*(11), e1916430. https://doi.org/10.1001/jamanetworkopen.2019.13436

Ravegnini, G., Ferioli, M., Pantaleo, F. M., Morganti, A. G., De Leo, A., De Iaco, P., Rizzo, S., et al. (2022). Radiomics and artificial intelligence in malignant uterine body cancers: Protocol for a systematic review. *PLoS One, 17*(6), e0267727. https://doi.org/10.1371/journal.pone.0267727

Rayzah, M. (2020). Phyllodes tumours of the breast: A literature review. *Cureus, 12*(9), e10288. https://doi.org/10.7759/cureus.10288

Resteghini, C., Trama, A., Borgonovi, E., Hosni, H., Corrao, G., Orlandi, E., et al. (2018). Big data in head and neck cancer. *Current Treatment Options in Oncology, 19*, 62. https://doi.org/10.1007/s11864-018-0585-2

Rocca, A., Brunese, M. C., Santone, A., Avella, P., Biancho, P., Scacchi, A., et al. (2022). Early diagnosis of liver metastases from colorectal cancer through CT radiomics and formal methods: A pilot study. *Journal of Clinical Medicine, 11*, 31. https://doi.org/10.3390/jcm11010031

Rogers, W., Seetha, S. T., Refaee, T. A. G., Lieverse, R. I. Y., Granzier, R. W. Y., Ibrahim, A., et al. (2020). Radiomics: From qualitative to quantitative imaging. *The British Journal of Radiology, 93*, 20190948. https://doi.org/10.1259/bjr.20190948

Rompianesi, G., Pegoraro, F., Ceresa, C. D. L., Montalti, R., & Troisi, R. I. (2022). Artificial intelligence in the diagnosis and management of colorectal cancer liver metastases. *World Journal of Gastroenterology, 28*(1), 108–122. https://doi.org/10.3748/wjg.v28.i1.108

Rudie, J. D., Rauschecker, A. M., Bryan, R. N., Davatzikos, C., & Mohan, S. (2019). Emerging applications of artificial intelligence in neuro-oncology. *Radiology, 290*(3), 607–618. https://doi.org/10.1148/radiol.2018181928

Salama, M. E., Macon, W. R., & Pantanowitz, L. (2020). Is the time right to start using digital pathology and artificial intelligence for the diagnosis of lymphoma? *Journal of Pathology Informatics, 1*, 16. https://doi.org/10.4103/jpi.jpi_16_20

Santilli, A. M. L., Jamzad, A., Janssen, N. N. Y., Kaufmann, M., Connolly, L., Vanderbeck, K., et al. (2020). Perioperative margin detection in basal cell carcinoma using deep learning framework: A feasibility study. *International Journal of Computer Assisted Radiology and Surgery, 15*, 887–896. https://doi.org/10.1007/s11548-020-02152-9

Shen, Y., Liu, J., Zhang, L., Dong, S., Zhang, J., Liu, Y., et al. (2019). Identification of potential biomarkers and survival analysis for head and neck squamous cell carcinoma using bioinformatics strategy: A study on TCGA and GEO datasets. *BioMed Research International, 2019*. https://doi.org/10.1155/2019/7376034

Shimizu, H., & Nakayama, K. I. (2020). Artificial intelligence in oncology. *Cancer Science, 111*, 1452–1460. https://doi.org/10.1111/cas.14377

Skierucha, M. (2016). Molecular alterations in gastric cancer with special reference to the early-onset subtype. *World Journal of Gastroenterology, 22*(8), 2460.

Song, Z., Wu, Y., Yang, J., Yang, D., & Fang, X. (2017). Progress in the treatment of advanced gastric cancer. *Tumor Biology, 39*, 1–7. https://doi.org/10.1177/1010428317714626

Syed, T., Doshi, A., Guleria, S., Syed, S., & Shah, T. (2020). Artificial intelligence and its role in identifying Esophageal neoplasia. *Digestive Diseases and Sciences, 65*(12), 3448–3455. https://doi.org/10.1007/s10620-020-06643-2

Tian, H., Guan, D., & Li, J. (2018). Identifying osteosarcoma metastasis associated genes by weighted gene co-expression network analysis (WGCNA). *Medicine, 97*, 24. (e10781). https://doi.org/10.1097/MD.0000000000010781

Tunali I, Gillies RJ, Schabath MB. Application of radiomics and AI for lung cancer precision medicine. Cold Spring Harb Perspect Med 2021; 11 (8), a039537 https://doi.org/10.1101/cshperspect.a039537.

Van Dijk, L. V., & Fuller, C. D. (2021). Artificial intelligence and radiomics in head and neck cancer care: Opportunities, mechanics, and challenges. *American Society of Clinical Oncology Educational Book, 41*, 1–11. https://doi.org/10.1200/EDBK_320951

Vibhakar, A. M., Cassels, J. A., Botchu, R., Rennie, W. J., & Shah, A. (2021). Imaging update on soft tissue sarcoma. *Journal of Clinical Orthopaedics and Trauma, 22*, 101568. https://doi.org/10.1016/j.jcot.2021.101568

Visaggi, P., Barberio, B., Gnisa, M., Rebolsi, M., Savarino, V., Fassa, M., et al. (2021). Modern diagnosis of early Esophageal cancer: From blood biomarkers to advanced endoscopy and artificial intelligence. *Cancers, 13*, 3162. https://doi.org/10.3390/cancers13133152

Vo, T. K., Parsons, D. W., & Seibel, N. L. (2020). Precision medicine in pediatric oncology. *Surgical Oncology Clinics of North America, 29*(1), 63–72. https://doi.org/10.1016/j.soc.2019.08.005

Waddell, N., Pajic, M., Patch, A. M., Chang, D. K., Kassahn, K. S., Bailey, P., et al. (2015). Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature, 518*(7540), 495–501. https://doi.org/10.1038/nature14169

Wang, G., Zhao, D., Spring, D. J., & Depinho, R. A. (2018). Genetics and biology of prostate cancer. *Genes & Development, 32*, 1105–1140. http://www.genesdev.org/cgi/doi/10.1101/gad.315739

Wang, P. P., Deng, C. L., & Wu, B. (2021). Magnetic resonance imaging-based artificial intelligence model in rectal cancer. *World Journal of Gastroenterology, 27*(18), 2122–2130. https://doi.org/10.3748/wjg.v27.118.2122

Wang, T., Wang, H., Wang, Y., Liu, X., Ling, L., Zhang, G., et al. (2022). MR-based radiomics-clinical nomogram in epithelial ovarian tumor prognosis prediction: Tumor body texture analysis across various acquisition protocols. *Journal of Ovarian Research, 15*, 6. https://doi.org/10.1186/s13048-021-00941-7

Weidlich, V., & Weidlich, G. A. (2018). Artificial intelligence in medicine and radiation oncology. *Cureus, 10*(4), e2475. https://doi.org/10.7759/cureus.2475

Wu, G., Jochems, A., Refaee, T., Ibrahim, A., Yan, C., Sanduleanu, S., et al. (2021). Structural and functional radiomics for lung cancer. *European Journal of Nuclear Medicine and Molecular Imaging, 48*, 3961–3974. https://doi.org/10.1007/s00259-021-05242-1

Xue, P., Alexander, M. T., & Qiao, Y. (2020). The challenges of colposcopy for cervical cancer screening in LMICs and solutions by artificial intelligence. *BMC Medicine, 18*, 169. https://doi.org/10.1186/s12916-020-01613.x

Yang, C. W., Liu, X. J., Liu, S. Y., Wan, S., Ye, Z., & Song, B. (2020). Current and potential applications of artificial intelligence in gastrointestinal stromal tumor imaging. *Contrast Media & Molecular Imaging, 133*, 6058159. https://doi.org/10.1155/2020/6058159

Yang, C. M., & Shu, J. (2021). Cholangiocarcinoma Evaluation via Imaging and Artificial Intelligence. *Oncology, 99*, 72–83. https://doi.org/10.1159/000507449

Yoon, J., Lee, E., Koo, J. S., Yoon, J. H., Nam, K. H., Lee, J., et al. (2020). Artificial intelligence to predict $BRAF^{V600E}$ mutation in patients with thyroid cancer. *PLoS One, 15*, e0242806. https://doi.org/10.1371/journal.pone.0242806

Zhang, Y. H., Guo, L. J., Yuan, X. L., & Hu, B. (2020a). Artificial intelligence-assisted esophageal cancer management: Now and future. *World Journal of Gastroenterology, 26*(35), 5256–5271. https://doi.org/10.3748/wjg.v26.i35.5256

Zhang, Q., Rong, Y., Yi, K., Huang, L., Chen, M., & Wang, F. (2020b). Circulating tumor cells in hepatocellular carcinoma: Single-cell based analysis, preclinical models, and clinical applications. *Theranostics, 10*(26), 12060–12071. https://doi.org/10.7150/thno.48918

Zhao, W. J., Fu, L. R., Huang, Z. M., Zhu, J. Q., & Ma, B. Y. (2019). Effectiveness evaluation of computer-aided diagnosis system for the diagnosis of thyroid nodules on ultrasound. *Medicine, 98*, 32. (e16379). https://doi.org/10.1097/MD.0000000000016379

Zhao, Y., Xu, L., Wang, X., Niu, S., Chen, H., & Li, C. G. (2021). A novel prognostic mRNA/miRNA signature for esophageal cancer and its immune landscape in cancer progression. *Molecular Oncology, 15*, 1088–1109. https://doi.org/10.1002/1878-0261.12902

Zhou, K. J., Cao, W., Wang, L., Pan, Z., & Fu, Y. (2022). Application of artificial intelligence in the diagnosis and prognostic prediction of ovarian cancer. *Computers in Biology and Medicine, 146*, 105608. https://doi.org/10.1016/jcompbiomed.2022.105608

Zhou, L. Q., Wang, J. Y., Yu, S. Y., Wu, G. G., Wei, Q., Deng, Y. B., et al. (2019). Artificial intelligence in medical imaging of the liver. *World Journal of Gastroenterology, 25*(6), 672–682. https://doi.org/10.3748/wjg.v25.i6.672

Zhu, K. P., Zhang, C. L., Ma, X. L., Hu, J. P., Cai, T., & Zhang, L. (2019). Analyzing the interactions of mRNAs and ncRNAs to predict competing endogenous RNA networks in osteosarcoma chemo-resistance. *Molecular Therapy, 27*(3), 518. https://doi.org/10.1016/j.ymthe.2019.01.001

Chapter 4
AI as a Novel Approach for Exploring ccfNAs in Personalized Clinical Diagnosis and Prognosis: Providing Insight into the Decision-Making in Precision Oncology

Lesetja Motadi, Mzubanzi Mabongo, Demetra Demetriou, Precious Mathebela, and Zodwa Dlamini

Abstract For many years the idea of personalized medicine has toyed around. It is characterized as an innovative method to alter disease treatment and prevention that considers the differences in individuals' genes and lifestyles. The purpose of precision medicine is to provide patient-specific treatments at the right time for quick and cost-effective recovery. While world technological advances provide the opportunity for many in medical science to access big data that no human brain can collate, the use of artificial intelligence (AI) may help improve cancer screening and diagnosis and planned treatment. In this chapter, although we will not cover all technologies already available for patient care, we highlight the progress made in using technology combined with liquid biopsy in precision oncology. We will also be discussing AI as a novel approach for exploring circulating cell-free nucleic acids (ccfNAs) in personalized clinical diagnosis and prognosis.

L. Motadi (✉)
Department of Biochemistry, University of Johannesburg, Johannesburg, South Africa
e-mail: lmotadi@uj.ac.za

M. Mabongo
Department of Maxillofacial and Oral Surgery, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
e-mail: mzubanzi.mabongo@up.ac.za

D. Demetriou · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: dd.demetriou@up.ac.za; zodwa.dlamini@up.ac.za

P. Mathebela
Department of Surgery, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: mathebela.precious@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_4

Keywords Artificial intelligence · CcfNAs · Precision oncology · Novel · Diagnosis · Prognosis · Clinical decision-making

4.1 Introduction

Cancer is one of the common global causes of mortality and morbidity. Its burden is putting a lot of pressure on the health systems. Management of cancer is still mainly based on the TNM staging with the current being the sixth edition of AUJC. This classification informs decision-making and planning that includes treatment and prognosis. In most cases, in an attempt to pull together the skills and resources from different fields, multidisciplinary clinics are formed to improve the quality of decision-making and prognostication of cancer. The unfortunate aspect of this data is that it is not always digitalized in some units. The need for digitalization of the data has always been a goal for clinicians. Unfortunately, scant resources make this a pie in the sky, especially in low-income countries. The digitalization of oncology data would bring the service closer to the use of artificial intelligence (AI).

Artificial intelligence refers to the use of computer techniques to simulate clinician intelligence and cognitive function in health sciences (Elkhader & Elemento, 2022). AI plays a vital role in personalized medicine which developed as a result of the collaboration of experts in medicine, mathematics, and physics (La Porta & Zapperi, 2018). In medicine, AI has several applications including diagnosis, early detection of cancer using imaging and genomics disease monitoring, and new drug development. AI comprises two subfields; machine learning (ML) and deep learning (DL). ML enables computers to develop the necessary skills for problem-solving and learn without being precisely programmed (Ilhan et al., 2021). ML is classified into unsupervised -, supervised learning, and reinforcement learning (Dutton & Conroy, 1997; Jain et al., 1999). Unsupervised learning algorithms can analyze unlabeled data and can be used to identify structures that are hidden. The algorithm itself is able to detect patterns in the data that it can use for learning as the system does not have expectable results or prior labeled data. Supervised learning is the learning process of the machine which is based on known or labeled data. In reinforcement learning, the software is exposed to negative and/or positive learning within a powerful environment (García-Pola, 2021). ML uses different types of classifiers which include artificial neural networks (ANN), support vector machines (SVM), and decision trees. These require input data and structured hierarchical learning networks that are accurately imported to assist the computer systems to build a mathematical-based model for decision-making.

In deep learning, complex architectural analogs are constructed to the interconnected neurons of the human brain (Hunter et al., 2022). Deep learning algorithms that are commonly applied in the genomics field include feedforward neural networks (FNN), natural language processing (NLP), convolutional neural networks (CNNs), recurrent neural networks (RNNs), bidirectional long short-term memory networks (BLSTMs), long short-term memory networks (LSTMs), and gated recurrent unit (Alharbi & Rashid, 2022). AI may provide a one-stop on

which the captured data like signs and symptoms, histology images and reports, biochemistry results, and radiology images and reports can be collated and integrated. This expediates the diagnosis, planning, improve accuracy and consistency in decision-making (Greatbatch et al., 2019). AI in oncology has many applications which include liquid biopsy extraction, identification, detection, classification, and prognosis. Liquid biopsy assays continuously rely on ML and AI to isolate cellular and molecular signatures that can predict the outcomes or presence of tumors (Elkhader & Elemento, 2022).

Precision medicine has been used currently to achieve greater personalized care by the development of the latest updated diagnostic methods to study individual variability (Kaur et al., 2017). The initial significant step to this personalized medicine is the collection of comprehensive real-time data. In oncology, this includes liquid biopsy, appropriate diagnostic imaging, biochemistry results, and clinical records. The analytes from the liquid biopsy include circulating free nucleic acids (cfNA), which is the DNA and RNA material released into circulation cells that have undergone apoptosis and necrosis (Mandel, 1948; Pös et al., 2018). Useful markers within the body fluids have been identified as helpful in personalized treatment resulting in increased survival (Kosaka et al., 2010). Liquid biopsies are being used for their non-invasive, prognostic, and predictive nature. Under normal homeostasis, ccfNAs are produced from the hematopoietic system. In clinical conditions, ccfNAs are released by circulating tumor cells into body fluids like serum, plasma, and cerebrospinal fluids by apoptosis, necrosis, phagocytosis, and exocytosis (De Rubis et al., 2018; Zhou et al., 2017).

Cell-free Nucleic Acids (cfNA) include circulating tumor DNA (ctDNA) and RNA which are released by apoptotic tumor cells. The cfNA analytes from the liquid biopsy consist of cell-free DNA (cfDNA) and RNA. The RNA components include long non-coding RNAs (lncRNAs), microRNA (miRNA), and exomes (Sorber et al., 2017). These analytes act as diagnostic and prognostic markers as well as therapeutic targets providing significant clinical benefits to patients. The captured data from these analytes, for example, the quantities, are used for both diagnosis and disease monitoring. This provides valuable information for AI. Figure 4.1 shows the use of cfNA in cancer diagnosis and treatment.

4.2 Cancer Liquid Biopsies and Their Use in Precision Oncology

Liquid biopsies have recently played a crucial role in diagnosing several cancers and screening in which a patient could benefit from a specific therapy. In addition, they are considered to be a new gold standard of care for cancer patients (Torres et al., 2020). These biopsies contain isolated tumor-derived products such as ctDNA which is present in the fluids drawn from cancer patients. Liquid biopsies tests have been proven to be highly accurate as they give an idea of what is happening in the tumor

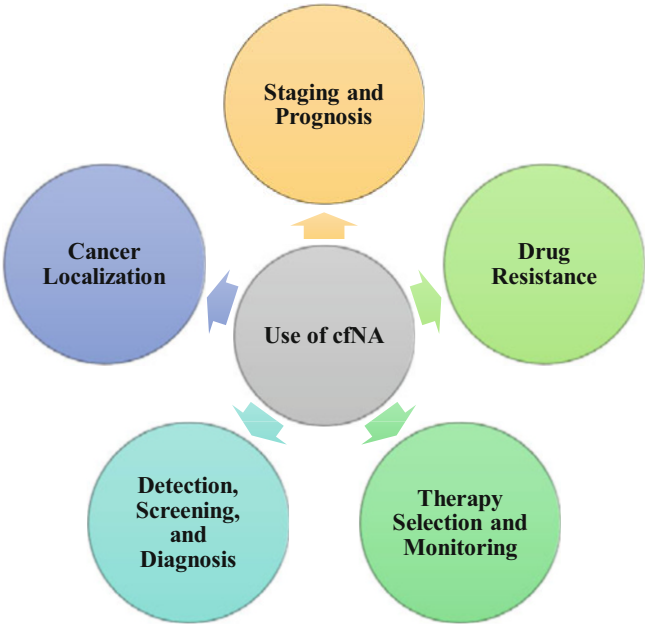


Fig. 4.1 Uses of cell-free nucleic acids (cfNA) in cancer diagnosis and treatment. CfNAs can be used to screen, detect and diagnose cancer. It can also be used for the localization of cancer through cfDNA pattern fragmentation, nucleosome spacing, methylation analysis, and non-coding RNA detection. Additionally, cfNAs are used in staging, prognosis, drug resistance, and treatment response monitoring

microenvironment (TME). The well-known liquid biopsy markers in cancer scenarios are circulating cell-free nucleic acids (ccfNAs), extracellular vesicles (EV), exfoliated tumor cells, cell-free proteins, exosomes, and peptides (Di Meo et al., 2017). In this chapter, we are only going to focus on the role of ccfNAs in cancer and how they can be integrated with AI to guide decision-making in precision oncology.

Radiomics is a specialized branch of AI that enables the extraction of image features obtained by clinicians (Gillies et al., 2016). These features can give an idea of pathophysiological processes and represent the phenotypic characteristics of the tumor. Because of their simple workflows, replicability and minimal invasiveness, the combined use of radiomics and liquid biopsies make them very attractive in oncology. ML approaches play an important role in ensuring the success of radiomic applications in clinical settings. In a radiomic study by Parma et al., they reported that two methods; the Wilcoxon test-based feature selection method (WLCX) and random forest (RF) had the best prognostic performance when compared to the other 12 methods (Parmar et al., 2015). Additionally, the methods were very stable against data variations. They went on to conclude that one must accurately choose ML methods that are optimal for radiomic applications as this can assist in obtaining stable and specific biomarkers.

4.2.1 Cell-Free DNAs

By definition cell-free DNAs (cfDNA) are fragments of DNA released into the plasma/serum following circulating tumor cell (CTC) lysis by apoptosis or necrosis which carries genome-wide DNA information. AI using cfDNA can achieve a mean sensitivity and specificity of 85% in early colorectal cancer in stages I and II (Wan et al., 2019). CancerSEEK detected eight common cancer types through the analysis of cfDNA and the use of a random forest model by evaluating eight proteins, 1933 gene positions and predicted malignancy with an area under the curve (AUC) of 91%. It also identified a very high proportion of ovarian and liver cancers (Cohen et al., 2018). Mitochondrial DNA (mtDNA) are detectable fragments released from the cell as a result of apoptosis. Some studies reported increased levels of circulating DNA in the blood of patients suffering from cancer (Meddeb et al., 2019).

4.2.2 Circulating Tumor DNA

Circulating tumor DNA (ctDNA) is a nucleic material isolated from circulation and can be accessed and analyzed from the liquid biopsy sample. CtDNA is a form of cfDNA that is expelled by the tumor cells into circulation (Wang et al., 2020). It is reported to be released into the peripheral blood after tumor cells outgrow their blood supply, become hypoxic and undergo apoptosis or necrosis. CtDNA in body fluids exist in two forms, a free single or double-stranded DNA and DNA–protein complex (Cheng et al., 2016; Zhang, Liang, et al., 2019). The half-life of ctDNA is reported to range from 15 minutes to 2 hours, and is rapidly cleared by kidneys, liver, and spleen (Perakis & Speicher, 2017). Like cfDNA, ctDNA molecules circulating in the blood are emerging as critical non-invasive biomarkers for pathological processes including cancer. CtDNA is differentiated from cfDNA by the presence of gene mutations. These mutations can be diagnosed by use of next-generation sequencing (NGS). NGS generates a large amount of data from ctDNA by whole genome sequencing and methylation sequencing (Hunter et al., 2022). This large genomic data generated from NGS can be integrated with the disease phenotype and other clinical records through AI algorithms. Combining AI and epigenomic analysis of ctDNA has been successful in the identification of genes that are differentially methylated in lung cancer (Bahado-Singh et al., 2022). This was archived through DL and with 100% sensitivity and specificity. A recent test in *EGFR* mutation through plasma droplet digital PCR (ddPCR) was proven to be hundred percent specific in detecting these mutations in lung cancer (Guo et al., 2019). CtDNA-based liquid biopsies present hope that it may guide precision medicine treatment through the identification of unique molecular characteristics of an individual's cancer. With a liquid biopsy, NGS is the best tool to sequence ctDNA to provide a molecular profile of cancer leading to clear diagnostic outcomes.

4.2.3 Cell-free Mitochondrial DNA in Cancer

Mitochondrial DNA (mtDNA) are detectable fragments of DNA of mitochondrial origin and are also released from the cell as a result of apoptosis. Several studies report elevated levels of circulating mtDNA in the blood of patients suffering from various diseases especially cancer (Meddeb et al., 2019). This elevation can be attributed to the fact that during various diseases the affected cells activate cell machinery like apoptosis which will attack the infected or affected cell. As described previously in many studies apoptosis results in the fragmentation of DNA and that of mitochondria (Mahmoud et al., 2016; Motadi et al., 2007). Like any other gene, mtDNAs are prone to mutations and such mutations have been implicated in the pathogenesis of multiple cancers thereby making circulating cell-free mtDNA as potential non-invasive tumor biomarker (Liu et al., 2016). These epigenetic marks on cfDNA have been referenced as indication of existence and location of tumors. They would also be used to monitor tumor burden noninvasively through estimating the percentage of ctDNA in the total cfDNAs (Snyder et al., 2016; Sun et al., 2015). Several studies also suggested that mitochondria-originated cfDNA fragments in cancer patients are more fragmented than those in healthy individuals (Cristiano et al., 2019). This suggestion provided knowledge in the diagnosis and also prognosis of cancer which can advise on the stages of cancer (An et al., 2019; Ma et al., 2017; Underhill et al., 2016).

In a study by Liu et al., they have shown that using NGS substantial fraction of tumor-specific mtDNA mutations in plasma cf-mtDNA specifically from hepatocellular carcinoma but none from colorectal cancer (Liu & Geng, 2022). Similarly, the examination in thyroid cancer even though not conclusive, displayed that the quantity of cfDNA was higher in patients affected by nodular thyroid diseases than healthy individuals (Salvianti et al., 2017). This evidence presents potential cancer-specific difference of tumor-derived mtDNA. This information suggests that monitoring the size of mt-cfDNAs in cancer patients would be a useful tool to estimate tumor burden and cancer progression.

4.3 AI and Cell-free RNAs in Cancer

Circulating cell-free ribonucleic acids (ccfRNAs) including mRNA and microRNA (miRNA) are present in significant levels in the blood of cancer patients which makes them potential targets for medical diagnosis. Well-known circulating RNAs are mainly represented by miRNAs, lncRNAs and messenger RNAs (mRNAs). Cell-free RNAs (cfRNA) are highly stable in body fluids because they are protected from endogenous RNase activity by encapsulation in lipoprotein complexes (Rapado-González, 2019). Unlike ctDNA, ccfRNA-based AI programs are not reliable or accurate. However, one of the most applied programs is RNA-seq has revealed a high accuracy to classify cancer subtypes and predict cancer progression (Elbashir

et al., 2019). There are several other ccfRNA-based AI programs that would not all be covered and have shown progress in improving precision oncology. CcfRNAs are endogenous non-coding small RNA molecules that are secreted into the blood circulation and are involved in several biological process regulations. There is increasing evidence that plasma and serum RNA could serve as tumor-specific markers for cancer detection diagnosis (Kolenda et al., 2020). Several studies have shown that cancer cells transmit their intercellular tumor-suppressive miRNAs to the extracellular environment, thereby modifying the microenvironment of the tumor and supporting cell proliferation of cancer cells (Bahrami et al., 2018; Jamali et al., 2018; Shekari et al., 2018; Wang et al., 2018). Therefore, extracellular miRNAs can be classified as both oncogenic and suppressors by different stimuli. Circulating miR-373 and miR-214 were associated with lymph node metastasis while similar miRNA molecules were reported to be valuable in the treatment of head and neck squamous carcinoma (Chen et al., 2013; Summerer et al., 2013).

4.3.1 Non-coding RNA in Cancer

A non-coding RNA is an RNA that is not translated into a functional protein. This represents a large segment of the human transcriptome that plays important roles in disease pathogenesis. There are several known non-coding RNAs such as small interfering RNA (siRNA) that bind to some mRNA and lead to their degradation in that way promoting or blocking the progression of diseases including cancer. The other type is miRNA which is the type of RNA that controls the expression of certain proteins. MiRNA regulates protein expression by targeting the formation of mRNA of a particular gene and this takes place in the nucleus. The identified role of ncRNAs promising applications in cancer diagnosis, prognosis, and therapy. In most cancers, the expression of ncRNAs correlates with cancer survival, metastasis and tumor grade, therefore providing great potential as a prognostic marker (Bray et al., 2018; Chandra Gupta & Nandan Tripathi, 2017). TP53 plays a major role in cancer prevention and positive response to chemotherapy (Lane, 1992). In a study by Jiang et al., they observed that the hsa-miR-125a-5p miRNA overexpression led to elevated p53 expression. Additionally, the miRNA was able to suppress cell proliferation and induce apoptosis via the p53 pathway in lung cancer (Jiang et al., 2011). Similar results were observed in metastatic colorectal cancer whereby treatment with bevacizumab resulted in increased expression of miR-125a-5p and miR-92b-3p with a positive response, suggesting that the antibody has an effect on regulation of miRNA (Kiss et al., 2021). Moreover, bevacizumab therapy resulted in increased survival of the patients. Figure 4.2 depicts how cfRNA and cfDNA are used in the identification of patient-targeted therapy.

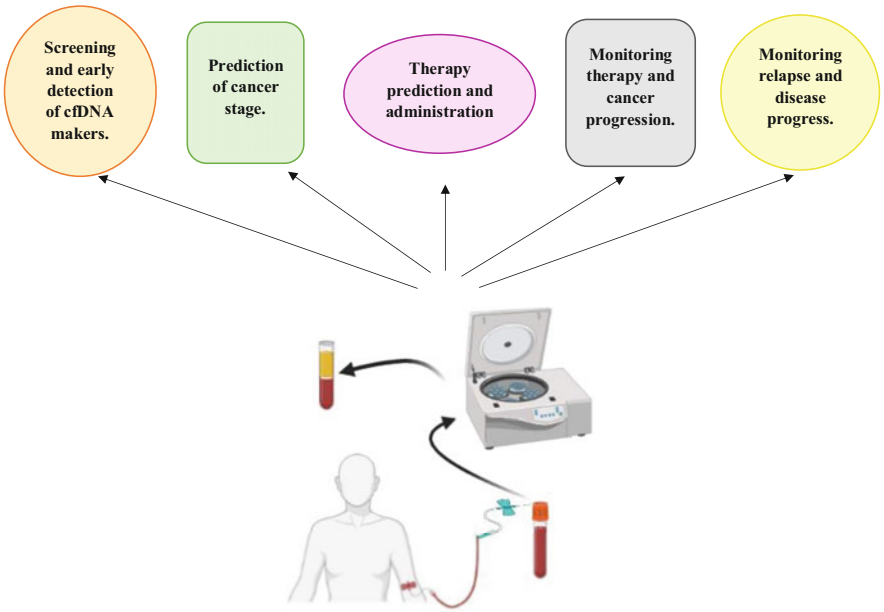


Fig. 4.2 Representation of how cfRNA and cfDNA in plasma/serum can be used in the identification of patient-targeted therapy. Created with BioRender.com

4.3.2 Long Non-coding RNA in Cancer

Long non-coding RNA (lncRNA) are non-coding RNAs that are said to be longer than 200 nt (Li & Chen, 2013). These RNAs interact with protein, DNA, mRNA, and miRNA to control gene expression through epigenetic modifications. This has been identified as a potential target for monitoring cancer treatment and progression. There is an estimate of about 16,000 lncRNA genes contained in the human genome (Fang et al., 2018; Uszczynska-Ratajczak et al., 2018). In cancer, lncRNAs are said to be involved in multiple mechanisms which include chromatin interactions, chromatin remodeling, and natural antisense transcripts. Studies have shown that lncRNAs may modulate transcription by seizing regulatory factors such as transcription factors and catalytic proteins or miRNAs (Kallen et al., 2013). In prostate cancer, there was a relationship between enhancer RNAs (eRNAs) levels produced by upstream enhancers of the prostate-specific antigen (*PSA*) gene and the expression of the actual *PSA* gene which suggests a relationship between chromatin remodeling and eRNA (Hsieh et al., 2014). Additionally, there were identified enhancer regions that bind transcriptional p53 factors thereby affecting the cell cycle arrest (Melo et al., 2013). Similarly, lncRNA *HULC* was reported to be upregulated in hepatocellular carcinoma with multiple binding sites of miR-372 (Wang et al., 2010). Overexpression of the *HULC* gene reduced miR-372 expression leading to the downregulation of translation of its target transcript *PRKACB* by

activating the phosphorylation of the cAMP-responsive element (CRE)-binding protein (CREB).

In some instances, lncRNAs act as competitive endogenous RNA (ceRNAs) to promote cancer, particularly pancreatic cancer. This was evident in studies where HLA complex group 11 (HCG11) was identified as the contributing factor in the suppression of apoptosis in various malignant tumors by facilitating cancer progression (Xu et al., 2017; Zhang, Huang, et al., 2019). Similarly, some studies have shown and supported that DLEU2L wiped miR-210-3p through competing with BRCA2 via ceRNA mechanism (Xu et al., 2021). Additionally, miR-210-3p served as an oncogene, which was presented by its direct correlation with malignant biochemical activities in pancreatic cancer cells, including proliferation, invasion, and migration (Ni et al., 2019). Yan et al. showed that in gastric cancer HOX transcript antisense RNA (*HOTAIR)* binds to miR-126 directly and inhibits its expression, resulting in enhanced expression of VEGFA and PIK3R2 and activating the PI3K/AKT/MRP1 pathway (Yan et al., 2016). *HOTAIR* acts as a ceRNA to promote cisplatin resistance. *HOTAIR* was reported to have targeted miR-17-5p while modifying PTEN which affects the proliferation and apoptosis of gastric cancer cells (Jia et al., 2019). MALAT1 is a nuclear-reserved lncRNA with over 8000 nucleotide bases located on chromosome 11q13. Studies suggested that MALAT1 is overexpressed in numerous human cancers, and the ceRNA network based on MALAT1 plays a crucial role in several cancer processes (Brown et al., 2012; Li et al., 2016). In lung cancer, MALAT1 is considered to be an early marker of metastasis and tumorigenesis (Li et al., 2016).

4.3.3 The Role of MicroRNAs in Human Cancer

MicroRNAs (miRNAs) are a family of small non-coding RNAs that function to control a wide range of biochemical mechanisms such as carcinogenesis. In cancer cells, miRNAs are said to be unregulated. MiRNAs function either as oncogenes or tumor suppressors during cancer development (Fig. 4.2). Several studies have shown that modification of specific miRNA through the use of miRNA mimics was able to normalize the gene signaling pathways and reverse the phenotype in cancerous cells (Borchert et al., 2006; Lee et al., 1993; Reinhart et al., 2000). The synthesis of miRNA begins with the transcription of the gene by either RNA polymerase II or III into a large primary transcript (Lee et al., 2004). The pri-miRNAs are then cleaved by a microprocessor complex and transported by Ran/GTP/Exportin 5 complex from the nucleus to the cytoplasm as a duplex structure. The mature miRNA is combined with protein complex termed RNA-induced silencing complex (RISC) and guides RISC to target mRNA (Macfarlane & Murphy, 2010).

Over the past decade, miRNA expression has been discovered to be deregulated in human cancers and this is owing to mechanisms such as epigenetic changes and defects in the miRNA synthesis. Overexpression of miRNA in cancer cells when compared to normal cells is often associated with alterations in genomic miRNA

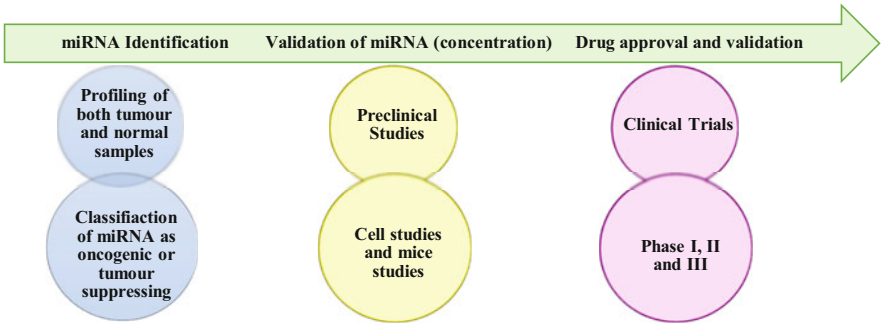


Fig. 4.3 Sequential use of miRNA in patient care and drug development

copy numbers and gene locations. Equally, amplification of *miR-17–92* cluster gene first observed in B-cell lymphomas and lung cancers and their translocation was observed in T-cell acute lymphoblastic leukemia (Hayashita et al., 2005; Mavrakis et al., 2010). Also, miRNA expression is tightly regulated by several transcription factors, so abnormal expression of miRNA in cancer was reported to be associated with dysregulation of some key transcription factors, such as c-Myc and p53 (O'Donnell et al., 2005). In the same study, oncogenic *miR-17–92* was activated in several cancers by c-myc to regulate apoptosis and cell proliferation. The following several cancer-suppressing miRNAs were all reported to be inactivated by c-myc, *mir-15a*, *miR-26*, *miR-29*, and *mir-30* (Chang et al., 2008). Figure 4.3 shows the sequential use of miRNA in patient care and drug development.

The *p53* gene is a tumor suppressor and is the most mutated gene in human cancers. *P53* regulates the expression of many carcinogenic genes through its transcriptional activity, including miR-34a/b/c that promotes cell-cycle arrest, cell senescence, and apoptosis in cancer (Chang et al., 2007; Hermeking, 2010; Raver-Shapira et al., 2007). MiR-223 is expressed in the hematopoietic system, and its expression is suppressed in many cancers including hepatocellular cancer and acute myeloid leukemia (Eyholzer et al., 2010; Stamatopoulos et al., 2009). The hallmarks of human cancer comprise six biochemical capabilities acquired during tumor development, this including sustaining proliferation, evading growth suppressors, avoiding cell death, and activating invasion and metastasis (Hanahan & Weinberg, 2011). Given that miRNA is abnormally expressed in cancers, it is suggested that the suppressed miRNAs could affect some of the cancer hallmarks. Depending on their validation and targets, miRNA may act as oncogene or tumor suppressor providing a clear target for cancer diagnosis and therapy.

4.3.4 Gene Silencing in Diagnosis and Prognosis of Cancer

RNA interference (RNAi) is a biochemical process that inhibits gene expression in cancer. In recent years it has been used/targeted to improve the efficiency, accuracy,

Table 4.1 Selected examples of siRNA, miRNA, and lncRNA and their targets in cancer therapy

Type of cancer	siRNA	microRNA	Non-coding RNA
Lung cancer	siRNA-EphA2-DOPC	Increase Apoptosis p53, STAT3	Phase I clinical trial NCT02369198; NCT02221999 for phase II/III.
Colorectal cancer	Doublecortin-like kinase 1 (DCLK1), STAT6	Increase bevacizumab; miR-216b-5p / ZNF14, miR-497/Bcl2, NF90/ VEGFA	KCNQ1OT1, NEAT1
Pancreatic cancer	Atu027 and siG12D LODER both at phase II.	HCG11/cisplatin resistance. Oncogenic/miR-132 and miR-212 act on RB1.	GAS5
Breast cancer	linc0015226	Trichostatin A targets over 22 upregulated miRNAs.	GAS5
Breast cancer	siRNA-mediated HOTAIR	MiR-27b and miR-892a act on CYP1B1 and CYP1A1 expression.	AK023948 as positive regulator of AKT.
Colorectal carcinoma, Ovarian cancer, Hepatoblastoma, Breast cancer.	siRNA-based inhibition of: MALAT1, TKM-PLK1 (TKM-080301)	miR-25, miR-504 and miR-30d act on suppressing p53.	miRNA182-5p, MEG3

and stability of treatments, particularly genetic therapies. SiRNAs are double-stranded RNAs (dsRNAs) with 21–23 nucleotides in length, which act by silencing their target genes through enzymatic cleavage of target mRNA. This targeted therapeutic system was designed to prevent tumor growth and it has been working in clinical studies in patients bearing solid tumors. SiRNA and miRNA-based therapies are now in clinical trials and recently novel siRNA-based therapeutics approved by the Food and Drug Administration (FDA), indicating the beginning of a new era of targeted therapeutics (Table 4.1).

RNAi-mediated silencing is activated in the cytoplasm through the conversion of double-stranded RNAs (dsRNAs) into small interfering RNAs (siRNAs) or directly via cytosolic delivery of chemically-synthesized siRNA drugs (Van de Vyver et al., 2021). The RNAi mechanism is triggered initially by the enzyme Dicer that slices dsRNAs into short double-stranded siRNAs of 21–25 nt. In recent years, AI algorithms have been introduced to synthesize numerous chemical siRNAs to increase efficacy and potency in RNAi for in vivo use (Behlke, 2008). This was evident in current Covid-19 vaccines. This has increased in vivo delivery of siRNA including the current FDA-approved Onpattro and Givlaari for the treatment of ATTR amyloidosis and acute hepatic porphyria, respectively (Gangopadhyay & Gore, 2022).

Table 4.1 *shows selected examples of siRNA, miRNA, and lncRNA and their targets in cancer therapy. SiRNA has been in use for several years now, however, the delivery of synthetic siRNA has been a challenge to reach target sites. In recent years several studies have identified* different types of systems for administrating

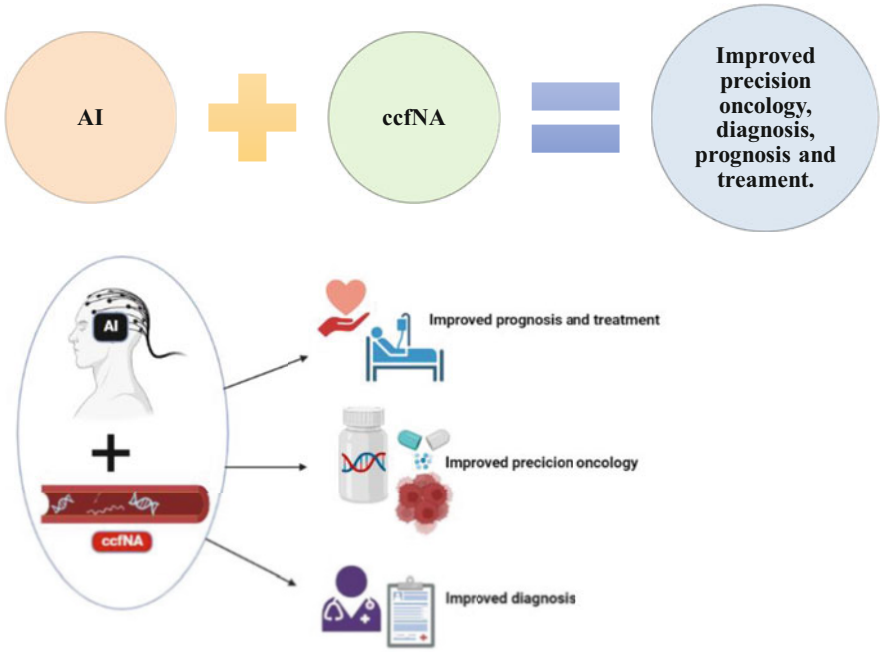


Fig. 4.4 The use of AI as a novel approach for exploring ccfNAs in personalized clinical diagnosis and prognosis provides insight into the decision-making in precision oncology. Created with BioRender.com

siRNA such as micelles, antibody conjugates, microparticles, synthetic cationic polymers, and peptides (Semple et al., 2010; Tam et al., 2013). Neutral lipid-based formulations have successfully enabled the delivery of siRNA in vivo in mouse models displaying tumor growth inhibition and consequently the downregulation of targeted genes (Oh & Park, 2009). Similarly, Ozcan et al. showed a successful siRNA-based treatment against ovarian cancer (Ozcan et al., 2015). In addition to lipids, several nanoparticles are approved as carriers for human drug use (Wagner, 2012). One of the anticancer systems approved is CALAA-01, which is a targeted nanocomplex that includes an anti-R2 siRNA that inhibits cancer growth by targeting Bcl2 family of anti-apoptotic genes (Rahman et al., 2013). Another form of delivery includes a combination of Poly-L-lactic-co-glycolic acid (PLGA) and hyaluronic acid (HA) which plays a crucial role as a binding ligand against CD44 receptor which is highly overexpressed in tumor cells (Byeon et al., 2018). The combination was used to deliver paclitaxel together with a siRNA against focal adhesion kinase (FAK) that is overexpressed in breast, colon and ovarian cancers which resulted in highly targeted delivery to CD44+ cells (Byeon et al., 2018). However, the limitation to siRNA therapeutic agents is their vulnerability to degradation in serum. Figure 4.4 shows a summary of ccfNAs and AI to improve precision oncology, diagnosis, prognosis, and treatment.

4.4 Limitations and Future Perspectives

Artificial intelligence is changing the healthcare industry, with new predictions carried out daily and unlimited. With new technology comes new challenges. One of the challenges of AI in precision oncology is the biasness, lack of clarity for some AI algorithms. Although this technology assists in developing early detections in mutations and predictions of when one might develop cancer, in poor and developing countries due to lack of resources and the cost of the technology, people would hardly benefit from it. There are challenges in the development, implementation, and maintenance of AI models are substantial. One of the limitations of ccfNAs is that they have a relatively short half-life and requires precautious handling to retain their integrity and stability.

Other limitations include unregulated set algorithms, unsupervised learning implementations and patient data confidentiality. A major problem in the development of AI models is the absence of structured, cancer-related data, as well as the nonexistence of standardization. AI can be used as a diagnostic tool routinely as part of medical check-ups for adult patients. This can be achieved through ML algorithms by digital diagnosis. AI algorithms allow data to be analyzed from DNA, RNA, metabolites and other analytes from liquid biopsy at the same time, cutting costs and financial implications. AI can also be used as a non-invasive tool that can easily diagnose genomic aberration by using a probe over the large surface blood vessels of the body and can be used as a reliable tool to easily acquire liquid biopsies such as saliva and urine.

4.5 Conclusion

Despite the results achieved so far, the application of AI to cancer for valuable precision oncology is still limited. Oncologic radiographic imaging AI is currently being used for detection and diagnosis. However, there has not been any highly clinical value to it. Genomic profiling has revolutionized precision oncology for eligibility to targeted therapy, identification of chemotherapy-induced genomic alterations as compared to the initial tissue biopsy. On the other hand, the development of AI-based technologies for biomedical applications opened a new era in the field of personalized clinical diagnosis and prognosis. AI can be used as novel approach for exploring ccfNAs in personalized clinical diagnosis and prognosis that will provide insight into the decision-making in precision oncology.

References

Alharbi, W. S., & Rashid, M. (2022). A review of deep learning applications in human genomics using next-generation sequencing data. *Human Genomics, 16*, 26.

An, Q., Hu, Y., Li, Q., Chen, X., Huang, J., Pellegrini, M., Zhou, X. J., Rettig, M., & Fan, G. (2019). The size of cell-free mitochondrial DNA in blood is inversely correlated with tumor burden in cancer patients. *Precision Clinical Medicine, 2*, 131–139.

Bahado-Singh, R., Vlachos, K. T., Aydas, B., Gordevicius, J., Radhakrishna, U., & Vishweswaraiah, S. (2022). Precision oncology: Artificial intelligence and DNA methylation analysis of circulating cell-free DNA for lung cancer detection. *Frontiers in Oncology, 12*, 790645.

Bahrami, A., Aledavood, A., Anvari, K., Hassanian, S. M., Maftouh, M., Yaghobzade, A., Salarzaee, O., Shahidsales, S., & Avan, A. (2018). The prognostic and therapeutic application of microRNAs in breast cancer: Tissue and circulating microRNAs. *Journal of Cellular Physiology, 233*, 774–786.

Behlke, M. A. (2008). Chemical modification of siRNAs for in vivo use. *Oligonucleotides, 18*, 305–320.

Borchert, G. M., Lanier, W., & Davidson, B. L. (2006). RNA polymerase III transcribes human microRNAs. *Nature Structural & Molecular Biology, 13*, 1097–1101.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 68*, 394–424.

Brown, J. A., Valenstein, M. L., Yario, T. A., Tycowski, K. T., & Steitz, J. A. (2012). Formation of triple-helical structures by the 3′-end sequences of MALAT1 and MENβ noncoding RNAs. *Proceedings of the National Academy of Sciences, 109*, 19202–19207.

Byeon, Y., Lee, J.-W., Choi, W. S., Won, J. E., Kim, G. H., Kim, M. G., Wi, T. I., Lee, J. M., Kang, T. H., & Jung, I. D. (2018). CD44-targeting PLGA nanoparticles incorporating paclitaxel and FAK siRNA overcome chemoresistance in epithelial ovarian cancer overcoming chemoresistance by HA-PLGA-NP in ovarian cancer. *Cancer Research, 78*, 6247–6256.

Chandra Gupta, S., & Nandan Tripathi, Y. (2017). Potential of long non-coding RNAs in cancer patients: From biomarkers to therapeutic targets. *International Journal of Cancer, 140*, 1955–1967.

Chang, T.-C., Wentzel, E. A., Kent, O. A., Ramachandran, K., Mullendore, M., Lee, K. H., Feldmann, G., Yamakuchi, M., Ferlito, M., & Lowenstein, C. J. (2007). Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Molecular Cell, 26*, 745–752.

Chang, T.-C., Yu, D., Lee, Y.-S., Wentzel, E. A., Arking, D. E., West, K. M., Dang, C. V., Thomas-Tikhonenko, A., & Mendell, J. T. (2008). Widespread microRNA repression by Myc contributes to tumorigenesis. *Nature Genetics, 40*, 43–50.

Chen, W., Cai, F., Zhang, B., Barekati, Z., & Zhong, X. Y. (2013). The level of circulating miRNA-10b and miRNA-373 in detecting lymph node metastasis of breast cancer: Potential biomarkers. *Tumor Biology, 34*, 455–462.

Cheng, F., Su, L., & Qian, C. (2016). Circulating tumor DNA: A promising biomarker in the liquid biopsy of cancer. *Oncotarget, 7*, 48832–48841.

Cohen, J. D., Li, L., Wang, Y., Thoburn, C., Afsari, B., Danilova, L., Douville, C., Javed, A. A., Wong, F., & Mattox, A. (2018). Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science, 359*, 926–930.

Cristiano, S., Leal, A., Phallen, J., Fiksel, J., Adleff, V., Bruhm, D. C., Jensen, S., Medina, J. E., Hruban, C., White, J. R., Palsgrove, D. N., Niknafs, N., Anagnostou, V., Forde, P., Naidoo, J., Marrone, K., Brahmer, J., Woodward, B. D., Husain, H., Van Rooijen, K. L., Ørntoft, M. W., Madsen, A. H., Van De Velde, C. J. H., Verheij, M., Cats, A., Punt, C. J. A., Vink, G. R., Van Grieken, N. C. T., Koopman, M., Fijneman, R. J. A., Johansen, J. S., Nielsen, H. J., Meijer,

G. A., Andersen, C. L., Scharpf, R. B., & Velculescu, V. E. (2019). Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature, 570*, 385–389.

De Rubis, G., Krishnan, S. R., & Bebawy, M. (2018). Circulating tumor DNA–current state of play and future perspectives. *Pharmacological Research, 136*, 35–44.

Di Meo, A., Bartlett, J., Cheng, Y., Pasic, M. D., & Yousef, G. M. (2017). Liquid biopsy: A step forward towards precision medicine in urologic malignancies. *Molecular Cancer, 16*, 1–14.

Dutton, D. M., & Conroy, G. V. (1997). A review of machine learning. *The Knowledge Engineering Review, 12*, 341–367.

Elbashir, M. K., Ezz, M., Mohammed, M., & Saloum, S. S. (2019). Lightweight convolutional neural network for breast cancer classification using RNA-seq gene expression data. *IEEE Access, 7*, 185338–185348.

Elkhader, J., & Elemento, O. (2022). Artificial intelligence in oncology: From bench to clinic. *Seminars in Cancer Biology, 84*, 113–128.

Eyholzer, M., Schmid, S., Schardt, J. A., Haefliger, S., Mueller, B. U., & Pabst, T. (2010). Complexity of miR-223 regulation by CEBPA in human AML. *Leukemia Research, 34*, 672–676.

Fang, S., Zhang, L., Guo, J., Niu, Y., Wu, Y., Li, H., Zhao, L., Li, X., Teng, X., & Sun, X. (2018). NONCODEV5: A comprehensive annotation database for long non-coding RNAs. *Nucleic Acids Research, 46*, D308–D314.

Gangopadhyay, S., & Gore, K. R. (2022). Advances in siRNA therapeutics and synergistic effect on siRNA activity using emerging dual ribose modifications. *RNA Biology, 19*, 452–467.

García-Pola, M. P. (2021). *Cancers* [Online], 13.

Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images are more than pictures, they are data. *Radiology, 278*, 563.

Greatbatch, O., Garrett, A., & Snape, K. (2019). The impact of artificial intelligence on the current and future practice of clinical cancer genomics. *Genetics Research, 101*, e9.

Guo, Q.-M., Wang, L., Yu, W.-J., Qiao, L.-H., Zhao, M.-N., Hu, X.-M., Sun, Y.-M., Ni, S., Xu, Y.-H., & Lou, J.-T. (2019). Detection of plasma EGFR mutations in NSCLC patients with a validated ddPCR lung cfDNA assay. *Journal of Cancer, 10*, 4341.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell, 144*, 646–674.

Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., Yatabe, Y., Kawahara, K., Sekido, Y., & Takahashi, T. (2005). A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer Research, 65*, 9628–9632.

Hermeking, H. (2010). The miR-34 family in cancer and apoptosis. *Cell Death & Differentiation, 17*, 193–199.

Hsieh, C.-L., Fei, T., Chen, Y., Li, T., Gao, Y., Wang, X., Sun, T., Sweeney, C. J., Lee, G.-S. M., & Chen, S. (2014). Enhancer RNAs participate in androgen receptor-driven looping that selectively enhances gene activation. *Proceedings of the National Academy of Sciences, 111*, 7319–7324.

Hunter, B., Hindocha, S., & Lee, R. W. (2022). The role of artificial intelligence in early cancer diagnosis. *Cancers [Online], 14*.

Ilhan, B., Guneri, P., & Wilder-Smith, P. (2021). The contribution of artificial intelligence to reducing the diagnostic delay in oral cancer. *Oral Oncology, 116*, 105254.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys (CSUR), 31*, 264–323.

Jamali, L., Tofigh, R., Tutunchi, S., Panahi, G., Borhani, F., Akhavan, S., Nourmohammadi, P., Ghaderian, S. M., Rasouli, M., & Mirzaei, H. (2018). Circulating microRNAs as diagnostic and therapeutic biomarkers in gastric and esophageal cancers. *Journal of Cellular Physiology, 233*, 8538–8550.

Jia, J., Zhan, D., Li, J., Li, Z., Li, H., & Qian, J. (2019). The contrary functions of lnc RNA HOTAIR/miR-17-5p/PTEN axis and Shenqifuzheng injection on chemosensitivity of gastric cancer cells. *Journal of Cellular and Molecular Medicine, 23*, 656–669.

Jiang, L., Huang, Q., Chang, J., Wang, E., & Qiu, X. (2011). MicroRNA HSA-miR-125a-5p induces apoptosis by activating p53 in lung cancer cells. *Experimental Lung Research, 37*, 387–398.

Kallen, A. N., Zhou, X.-B., Xu, J., Qiao, C., Ma, J., Yan, L., Lu, L., Liu, C., Yi, J.-S., & Zhang, H. (2013). The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Molecular Cell, 52*, 101–112.

Kaur, J., Rahat, B., Thakur, S., & Kaur, J. (2017). Chapter 15 - trends in precision medicine. In M. Verma & D. Barh (Eds.), *Progress and challenges in precision medicine*. Academic Press.

Kiss, D., Machackova, T., Souckova, K., Fabian, P., Krepelkova, I., Svoboda, M., & Kiss, I. (2021). An independent validation study of candidate microRNAs as predictive biomarkers for bevacizumab-based therapy in patients with metastatic colorectal cancer. *In Vivo, 35*, 2809–2814.

Kolenda, T., Guglas, K., Baranowski, D., Sobocińska, J., Kopczyńska, M., Teresiak, A., Bliźniak, R., & Lamperska, K. (2020). cfRNAs as biomarkers in oncology–still experimental or applied tool for personalized medicine already? *Reports of Practical Oncology and Radiotherapy, 25*, 783–792.

Kosaka, N., Iguchi, H., & Ochiya, T. (2010). Circulating microRNA in body fluid: A new potential biomarker for cancer diagnosis and prognosis. *Cancer Science, 101*, 2087–2092.

La Porta, C. A. M., & Zapperi, S. (2018). Explaining the dynamics of tumor aggressiveness: At the crossroads between biology, artificial intelligence and complex systems. *Seminars in Cancer Biology, 53*, 42–47.

Lane, D. P. (1992). p53, guardian of the genome. *Nature, 358*, 15–16.

Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). Elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell, 75*, 843–854.

Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., & Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *The EMBO Journal, 23*, 4051–4060.

Li, C. H., & Chen, Y. (2013). Targeting long non-coding RNAs in cancers: Progress and prospects. *The International Journal of Biochemistry & Cell Biology, 45*, 1895–1910.

Li, J., Wang, J., Chen, Y., Li, S., Jin, M., Wang, H., Chen, Z., & Yu, W. (2016). LncRNA MALAT1 exerts oncogenic functions in lung adenocarcinoma by targeting miR-204. *American Journal of Cancer Research, 6*, 1099.

Liu, Y., & Geng, X. (2022). Long non-coding RNA (lncRNA) CYTOR promotes hepatocellular carcinoma proliferation by targeting the microRNA-125a-5p/LASP1 axis. *Bioengineered, 13*, 3666–3679.

Liu, T., Zhang, X., Gao, S., Jing, F., Yang, Y., Du, L., Zheng, G., Li, P., Li, C., & Wang, C. (2016). Exosomal long noncoding RNA CRNDE-h as a novel serum-based biomarker for diagnosis and prognosis of colorectal cancer. *Oncotarget, 7*, 85551–85563.

Ma, X., Zhu, L., Wu, X., Bao, H., Wang, X., Chang, Z., Shao, Y. W., & Wang, Z. (2017). Cell-free DNA provides a good representation of the tumor genome despite its biased fragmentation patterns. *PLoS One, 12*, e0169231.

Macfarlane, L.-A., & Murphy, P. (2010). MicroRNA: Biogenesis, function and role in cancer. *Current Genomics, 11*, 537–561.

Mahmoud, E. H., Fawzy, A., Ahmad, O. K., & Ali, A. M. (2016). Plasma circulating cell-free nuclear and mitochondrial DNA as potential biomarkers in the peripheral blood of breast cancer patients. *Asian Pacific Journal of Cancer Prevention, 16*, 8299–8305.

Mandel, P. (1948). Les acides nucleiques du plasma sanguin chez 1 homme. *Comptes Rendus des Seances de la Societe de Biologie et de Ses Filiales, 142*, 241–243.

Mavrakis, K. J., Wolfe, A. L., Oricchio, E., Palomero, T., De Keersmaecker, K., Mcjunkin, K., Zuber, J., James, T., Khan, A. A., & Leslie, C. S. (2010). Genome-wide RNA-mediated

interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia. *Nature Cell Biology, 12*, 372–379.

Meddeb, R., Dache, Z. A. A., Thezenas, S., Otandault, A., Tanos, R., Pastor, B., Sanchez, C., Azzi, J., Tousch, G., & Azan, S. (2019). Quantifying circulating cell-free DNA in humans. *Scientific Reports, 9*, 1–16.

Melo, C. A., Drost, J., Wijchers, P. J., Van De Werken, H., De Wit, E., Vrielink, J. A. O., Elkon, R., Melo, S. A., Léveillé, N., & Kalluri, R. (2013). eRNAs are required for p53-dependent enhancer activity and gene transcription. *Molecular Cell, 49*, 524–535.

Motadi, L. R., Misso, N. L., Dlamini, Z., & Bhoola, K. D. (2007). Molecular genetics and mechanisms of apoptosis in carcinomas of the lung and pleura: Therapeutic targets. *International Immunopharmacology, 7*, 1934–1947.

Ni, J., Zhou, S., Yuan, W., Cen, F., & Yan, Q. (2019). Mechanism of miR-210 involved in epithelial–mesenchymal transition of pancreatic cancer cells under hypoxia. *Journal of Receptors and Signal Transduction, 39*, 399–406.

O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V., & Mendell, J. T. (2005). C-Myc-regulated microRNAs modulate E2F1 expression. *Nature, 435*, 839–843.

Oh, Y.-K., & Park, T. G. (2009). siRNA delivery systems for cancer treatment. *Advanced Drug Delivery Reviews, 61*, 850–862.

Ozcan, G., Ozpolat, B., Coleman, R. L., Sood, A. K., & Lopez-Berestein, G. (2015). Preclinical and clinical development of siRNA-based therapeutics. *Advanced Drug Delivery Reviews, 87*, 108–119.

Parmar, C., Grossmann, P., Bussink, J., Lambin, P., & Aerts, H. J. (2015). Machine learning methods for quantitative radiomic biomarkers. *Scientific Reports, 5*, 1–11.

Perakis, S., & Speicher, M. R. (2017). Emerging concepts in liquid biopsies. *BMC Medicine, 15*, 75.

Pös, O., Biró, O., Szemes, T., & Nagy, B. (2018). Circulating cell-free nucleic acids: Characteristics and applications. *European Journal of Human Genetics, 26*, 937–945.

Rahman, M. A., Amin, A. R. M., Wang, D., Koenig, L., Nannapaneni, S., Chen, Z., Wang, Z., Sica, G., Deng, X., & Chen, Z. G. (2013). RRM2 regulates Bcl-2 in head and neck and lung cancers: A potential target for cancer TherapyRRM2 regulates Bcl-2 in human cancer. *Clinical Cancer Research, 19*, 3416–3428.

Rapado-González, Ó. L. (2019). *Cells* [Online], 8.

Raver-Shapira, N., Marciano, E., Meiri, E., Spector, Y., Rosenfeld, N., Moskovits, N., Bentwich, Z., & Oren, M. (2007). Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. *Molecular Cell, 26*, 731–743.

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., & Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature, 403*, 901–906.

Salvianti, F., Giuliani, C., Petrone, L., Mancini, I., Vezzosi, V., Pupilli, C., & Pinzani, P. (2017). Integrity and quantity of total cell-free DNA in the diagnosis of thyroid cancer: Correlation with cytological classification. *International Journal of Molecular Sciences, 18*, 1350.

Semple, S. C., Akinc, A., Chen, J., Sandhu, A. P., Mui, B. L., Cho, C. K., Sah, D. W. Y., Stebbing, D., Crosley, E. J., & Yaworski, E. (2010). Rational design of cationic lipids for siRNA delivery. *Nature Biotechnology, 28*, 172–176.

Shekari, N., Baradaran, B., Shanehbandi, D., & Kazemi, T. (2018). Circulating MicroRNAs: Valuable biomarkers for the diagnosis and prognosis of gastric cancer. *Current Medicinal Chemistry, 25*, 698–714.

Snyder, M. W., Kircher, M., Hill, A. J., Daza, R. M., & Shendure, J. (2016). Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. *Cell, 164*, 57–68.

Sorber, L., Zwaenepoel, K., Deschoolmeester, V., Van Schil, P., Van Meerbeeck, J., Lardon, F., Rolfo, C., & Pauwels, P. (2017). Circulating cell-free nucleic acids and platelets as a liquid biopsy in the provision of personalized therapy for lung cancer patients. *Lung Cancer, 107*, 100–107.

Stamatopoulos, B., Meuleman, N., Haibe-Kains, B., Saussoy, P., Van Den Neste, E., Michaux, L., Heimann, P., Martiat, P., Bron, D., & Lagneaux, L. (2009). microRNA-29c and microRNA-223 down-regulation has in vivo significance in chronic lymphocytic leukemia and improves disease risk stratification. *Blood, The Journal of the American Society of Hematology, 113*, 5237–5245.

Summerer, I., Niyazi, M., Unger, K., Pitea, A., Zangen, V., Hess, J., Atkinson, M. J., Belka, C., Moertl, S., & Zitzelsberger, H. (2013). Changes in circulating microRNAs after radiochemotherapy in head and neck cancer patients. *Radiation Oncology, 8*, 1–9.

Sun, K., Jiang, P., Chan, K. C. A., Wong, J., Cheng, Y. K. Y., Liang, R. H. S., Chan, W.-K., Ma, E. S. K., Chan, S. L., & Cheng, S. H. (2015). Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. *Proceedings of the National Academy of Sciences, 112*, E5503–E5512.

Tam, Y. Y. C., Chen, S., & Cullis, P. R. (2013). Advances in lipid nanoparticles for siRNA delivery. *Pharmaceutics, 5*, 498–507.

Torres, S., González, Á., Cunquero Tomas, A. J., Calabuig Farinas, S., Ferrero, M., Mirda, D., Sirera, R., Jantus-Lewintre, E., & Camps, C. (2020). A profile on cobas® EGFR mutation test v2 as companion diagnostic for first-line treatment of patients with non-small cell lung cancer. *Expert Review of Molecular Diagnostics, 20*, 575–582.

Underhill, H. R., Kitzman, J. O., Hellwig, S., Welker, N. C., Daza, R., Baker, D. N., Gligorich, K. M., Rostomily, R. C., Bronner, M. P., & Shendure, J. (2016). Fragment length of circulating tumor DNA. *PLoS Genetics, 12*, e1006162.

Uszczynska-Ratajczak, B., Lagarde, J., Frankish, A., Guigó, R., & Johnson, R. (2018). Towards a complete map of the human long non-coding RNA transcriptome. *Nature Reviews Genetics, 19*, 535–548.

Van De Vyver, T., De Smedt, S. C., & Raemdonck, K. (2021). Modulating intracellular pathways to improve non-viral delivery of RNA therapeutics. *Advanced Drug Delivery Reviews, 181*, 114041.

Wagner, E. (2012). Polymers for siRNA delivery: Inspired by viruses to be targeted, dynamic, and precise. *Accounts of Chemical Research, 45*, 1005–1013.

Wan, N., Weinberg, D., Liu, T.-Y., Niehaus, K., Ariazi, E. A., Delubac, D., Kannan, A., White, B., Bailey, M., & Bertin, M. (2019). Machine learning enables detection of early-stage colorectal cancer by whole-genome sequencing of plasma cell-free DNA. *BMC Cancer, 19*, 1–10.

Wang, J., Liu, X., Wu, H., Ni, P., Gu, Z., Qiao, Y., Chen, N., Sun, F., & Fan, Q. (2010). CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Research, 38*, 5366–5383.

Wang, H., Peng, R., Wang, J., Qin, Z., & Xue, L. (2018). Circulating microRNAs as potential cancer biomarkers: The advantage and disadvantage. *Clinical Epigenetics, 10*, 1–10.

Wang, H.-C., Yeh, T.-J., Chan, L.-P., Hsu, C.-M., & Cho, S.-F. (2020). Exploration of feasible immune biomarkers for immune checkpoint inhibitors in head and neck squamous cell carcinoma treatment in real world clinical practice. *International Journal of Molecular Sciences, 21*, 7621.

Xu, J., Wang, J., He, Z., Chen, P., Jiang, X., Chen, Y., Liu, X., & Jiang, J. (2021). LncRNA CERS6-AS1 promotes proliferation and metastasis through the upregulation of YWHAG and activation of ERK signaling in pancreatic cancer. *Cell Death & Disease, 12*, 1–13.

Xu, Y., Zheng, Y., Liu, H., & Li, T. (2017). Modulation of IGF2BP1 by long non-coding RNA HCG11 suppresses apoptosis of hepatocellular carcinoma cells via MAPK signaling transduction. *International Journal of Oncology, 51*, 791–800.

Yan, J., Dang, Y., Liu, S., Zhang, Y., & Zhang, G. (2016). LncRNA HOTAIR promotes cisplatin resistance in gastric cancer by targeting miR-126 to activate the PI3K/AKT/MRP1 genes. *Tumor Biology, 37*, 16345–16355.

Zhang, H., Huang, H., Xu, X., Wang, H., Wang, J., Yao, Z., Xu, X., Wu, Q., & Xu, F. (2019). LncRNA HCG11 promotes proliferation and migration in gastric cancer via targeting miR-1276/CTNNB1 and activating Wnt signaling pathway. *Cancer Cell International, 19*, 1–12.

Zhang, L., Liang, Y., Li, S., Zeng, F., Meng, Y., Chen, Z., Liu, S., Tao, Y., & Yu, F. (2019). The interplay of circulating tumor DNA and chromatin modification, therapeutic resistance, and metastasis. *Molecular Cancer, 18*, 36.

Zhou, S., Zhang, Z., Zheng, P., Zhao, W., & Han, N. (2017). MicroRNA-1285-5p influences the proliferation and metastasis of non-small-cell lung carcinoma cells via downregulating CDH1 and Smad4. *Tumour Biology, 39*, 1010428317705513.

Chapter 5
AI-Enhanced Digital Pathology and Radiogenomics in Precision Oncology

Demetra Demetriou, Rodney Hull, Mmamoletla Kgoebane-Maseko, Zarina Lockhat, and Zodwa Dlamini

Abstract Precision medicine is the personalization of medicine based on the molecular or genetic profiling of an individual and can lead to improved diagnosis, prognosis, and treatment. Personalized medicine requires key features to be extracted from large data sets, normally using artificial intelligence (AI). Artificial intelligence can also be used to perform digital pathology and to identify key features in cytology, hematology, and histology based on medical images. Digital pathology consists of the acquisition of information from slides prepared from patient samples and data analysis using specific software tools. AI can automate the process of digital pathology by learning to discriminate the regions of interest from background tissue and finding features of interest to make accurate diagnoses. All this can be done rapidly, with high repeatability and no bias. Medical imaging is used as a diagnostic tool in cancer screening and disease monitoring. Imaging techniques like CT, MRI, and PET/CT scans can visualize the interior of a patient's body non-invasively and with a skilled radiologist can be used as an accurate diagnostic tool in cancer. The application of AI to this field, combined with information obtained from large omics data sets, has led to an improved tool for the screening and monitoring of cancer in the form of radiogenomics. This technique uses AI to associate features detected in medical radiological imaging with disease phenotypes in patients based on their genomic, transcriptomic, or proteomic profiles. AI, however, still relies heavily on

D. Demetriou (✉) · R. Hull · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: dd.demetriou@up.ac.za; rodney.hull@up.ac.za; zodwa.dlamini@up.ac.za

M. Kgoebane-Maseko
Department of Anatomical Pathology, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
e-mail: mw.kgoebane-maseko@up.ac.za

Z. Lockhat
Department of Radiology, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: zarina.lockhat@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_5

good basic clinical practice. This starts with the investigation of patients' symptoms, taking an adequate clinical history, thorough clinical examination, baseline laboratory, and radiology/ultra-sonographic work-up. This chapter will discuss the role played by AI in the advancement of personalized patient care, through digital pathology and radiogenomics which is the aim of precision medicine.

Keywords Cancer · Genetics · CT · PET/CT · MRI · Precision medicine · Artificial intelligence

5.1 Introduction

Cancer is a major public health problem globally. Numerous studies have recorded that treatment is given at later stages of cancer progression leading to poor survival. It is crucial to detect cancer early and to provide the correct treatment. For this reason, diagnosis using digital pathology, radiogenomics and artificial intelligence (AI) is an important sub-field of precision medicine. Cancer is a complex disease. Various parameters must be considered for accurate decision-making (Abernethy et al., 2010). It is difficult for physicians to consider all these parameters. For this reason, AI is likely to become an important part of cancer management by providing accurate and faster interpretations of patient data (Fogel & Kvedar, 2018). Molecular biomarkers are a promising technique for cancer detection. These biomarkers can be used for diagnosis, prognosis, survival rate prediction, assistance in making a treatment decision, and in the management and surveillance of the disease (Dlamini et al., 2022). AI is used for advanced statistical and data analysis in conjunction with other methods such as digital pathology and radiogenomics. By increasing the accuracy of data analysis and knowledge of cancer through these methods, the implementation of precision oncology will be improved as the results will be tailored to a patient. This improves diagnosis, prognosis, and treatment. Figure 5.1 shows areas in basic patient work-up, where AI (through Next-Generation Sequencing (NGS), digital pathology, and radiogenomics) can be integrated to improve the speed and accuracy of diagnosis and treatment.

Precision medicine is a rising field for disease prevention and treatment that considers gene variation, environment and lifestyles of individuals. Precision medicine, also referred to as personalized medicine, aims to treat patients through tailor-made therapies based on unique traits. These traits include the patient's genome, transcriptome, and proteome from a genetics aspect and also include the patient's environment, lifestyle, and socio-economic status (Dlamini et al., 2022). Numerous studies are trying to find a way to improve the diagnosis, prognosis, treatment, and survival rate of cancer patients. Clinicians and scientists aim to find non-invasive imaging biomarkers that can be associated with genomic features and clinical outcomes (Liu & Hu, 2022). For this reason, digital pathology and radiogenomics, have proven of interest in the fields of cancer diagnosis and prognosis. AI, with the use of these methods, improves diagnosis and prognosis.

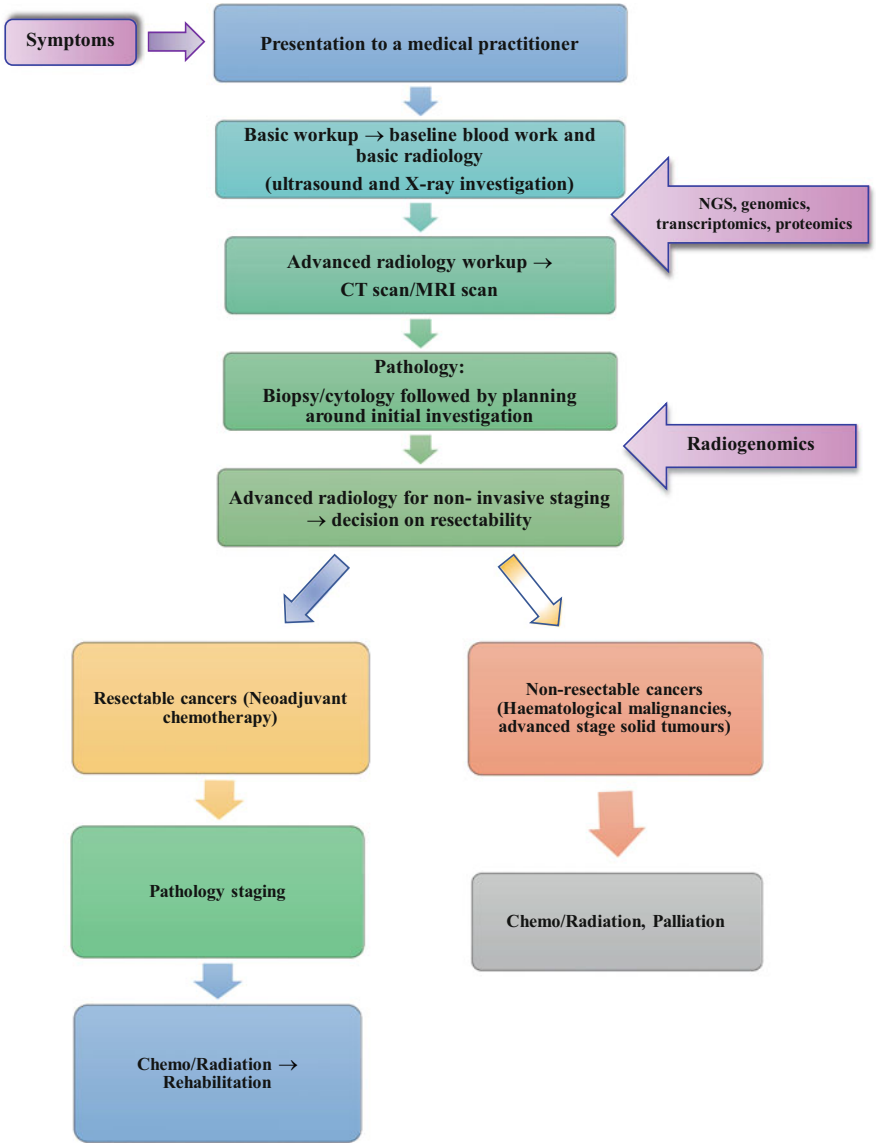


Fig. 5.1 Sequential steps of basic patient work-up and how AI will be factored in. The patient presents their symptoms to the medical practitioner. The medical practitioner will perform a physical examination. Basic blood work and basic radiology will be performed. This includes ultrasounds and X-ray investigations. NGS, Genomics and proteomics will be done leading to advanced radiology work-up in combination with a CT scan or MRI scan. Radiogenomics will enter at this point allowing informed decision-making. A biopsy will occur followed by planning around the initial investigation. A decision on resectability will be made. If the cancer is resectable, staging will be performed followed by chemotherapy or radiation and rehabilitation. If the cancer is non-resectable, chemotherapy or radiation will be given to the patient for the relief of symptoms and suffering

5.2 Medical Imaging in Precision Medicine

5.2.1 Magnetic Resonance Imaging (MRI) in Precision Medicine

Magnetic resonance imaging (MRI) uses radio waves and strong magnetic fields to produce detailed images of the body. MRI is a non-invasive method and is used to scan the brain, bones, spinal cord, heart, joints, and blood vessels. The results can be used for diagnosis, treatment plans and to analyze treatment efficacy. It uses non-ionizing electromagnetic radiation and employs radio frequency (RF) radiation in the presence of controlled magnetic fields (Katti et al., 2011). The patient will lie inside a tunnel that contains strong magnets. The human body is made of water molecules which consist of hydrogen and oxygen atoms. MRIs interact with the protons at the center of hydrogen molecules within the body. During the scan, the protons in the body line up in the same direction. Short lengths of radio waves are sent to a focus area of the body, distributing the alignment of the protons. When the radio waves stop, the protons re-align. Radio signals are sent out that are received by the receiver. The received information can provide the exact location of the protons in the body. These signals provide information about the exact proton location within the body. The signals from the protons are combined to create a detailed image of the inside of the body.

The advantages of MRI include non-invasiveness, no ionizing radiation, contrast manipulation to distinguish tissue, multiplanar image to obtain direct, sagittal, coronal, and oblique images and no adverse effects, to name a few (Pekar, 2006). Some disadvantages of MRI include claustrophobia, expensive equipment, implants such as cardiac pacemakers that might not be safe and image distortion in patients with surgical clips or stents (Pekar, 2006). MRI scans are used in precision medicine as it can provide detailed data about the individual patient.

5.2.2 Computed Tomography (CT) Scan in Precision Medicine

Computed tomography (CT) is a widely used cross-sectional imaging method. CT was a ground-breaking development in the 1970s. CT derives from computed (computer), tomo (to cut), and graph(y) (Caldemeyer & Buckwalter, 1999). CT uses X-rays or ionizing radiation that is coupled with an electronic detector array to record patterns of densities. This creates an image of a "slice" or "cut" of tissue. The X-ray beam rotates around the object and allows multiple X-ray projections to pass through the object (Rivas et al., 2011). The internal structure is reconstructed from multiple projections (König et al., 2015). The multiple projections produce a detailed 3-D image of the body. Doctors can evaluate hard tissues and soft tissues. CT scans are commonly done in hospitals and centers to evaluate life-threatening

conditions in a limited time. CT scans are used in advanced breast cancer to detect distant metastasis, especially in the chest, abdomen, and pelvis. Only a limited anatomic area can be scanned at a time. This is a disadvantage and at present, CT scans are not the most useful for whole-body screening. CT scans in precision medicine can evaluate hard tissues and soft tissues of the individual patient leading to improved diagnosis and treatment.

5.2.3 Positron Emission Tomography (PET)/Computed Tomography (CT) in Precision Medicine

Positron emission tomography (PET) evaluates tissue and organ function by using small amounts of radioactive materials or radiopharmaceuticals. PET/CT scans can be used to detect the early onset of disease by identifying changes at the cellular level. PET/CTs are commonly used to detect and monitor cancer. Currently, 18F-FDG PET/CT using a glucose analog, fluorine-18 labeled FDG is part of staging and monitoring treatment response in breast cancer. FDG PET imaging exploits the increased glycolysis found in most tumors, due to the over-expression of glucose transporters (GLUT) and increased hexokinase activity. GLUT expression is an independent prognostic marker and is associated with more aggressive disease and poor survival in various cancer types (Boellaard et al., 2010; Meyer et al., 2019). PET/CT is often chosen in precision medicine as it provides detailed data of the individual that leads to improved diagnosis, prognosis, and treatment.

5.2.4 CT vs. PET/CT Comparisons: The Preferred Choice

CT and PET/CT scans are both diagnostic tools that provide accurate and clear views of the body. The main difference is the focus of the tools. A CT scan creates a detailed stationary image of organs, bones, and tissues. A PET/CT scan shows how the tissues work on a cellular level within the body. A CT scan is used to evaluate hard and soft tissues. CT scans pass X-rays through the body allowing the creation of images. The radioactive material used in the PET/CT scan emits energy. The energy is detected by a unique camera to produce detailed images. A CT scan is performed in minutes. For this reason, CT scans are used in emergencies, for example, car accidents. PET/CT scans require more time. With a CT scan, radiation does not stay in the body. With PET/CT scans, a small amount of radiation may stay in the body for a limited time. CT scans display problems after a disease changes the tissue or organ's structure. PET/CT scans can detect cancer earlier than other tests as it looks at the cellular level. For this reason, PET/CT scans are the preferred choice in precision medicine.

5.3 Digital Pathology and AI

Pathology has always played a critical role in drug development including target identification, identification of drug mode of action and pharmacodynamics, etc. (Jubb et al., 2014; Kramer et al., 2007). Today, pathology has improved and has led to new areas of interest such as digital pathology. Digital pathology (DP) plays a significant role in laboratories and modern clinical practice (Tizhoosh & Pantanowitz, 2018; Farahani et al., 2015). DP consists of the acquisition of information via the microscopic examination of patient samples and the analysis, interpretation, and management of data via software tools. Technological advances allowed improved computing power, faster networks and increased storage. This has enabled pathologists to manage digital slide images more efficiently. Whole-slide imaging (WSI) aids as a platform for artificial intelligence applications. These images produce multiple sources of information, which include the presence of color information, anatomical orientation, and multiple scale information (Niazi et al., 2019). Advanced AI algorithms integrated with digital pathology, allow in-depth analysis of samples. Unique imaging markers associated with diseases can be identified earlier which will lead to improved prognosis and treatment selection. DP and AI have great potential for improving cancer management through precision oncology. DP reduces the turnaround time for pathologists. The digital aspect changes the way cancer is diagnosed, by allowing image and data sharing, integrated diagnostics, increased efficiency, modernization of pathology, cost saving, and improved patient care (Niazi et al., 2019). AI tools can assist pathologists by providing immediate, interactive, and standardized analysis of digital slides, that can be shared with multiple users (Farahani et al., 2015; Zarella et al., 2019). AI can also provide automated annotations and can play a vital role in quality assurance. Figure 5.2 shows the basic digital pathology workflow.

Tissue preparations for digital pathology remain the same as with traditional pathology. The pathologist will perform a gross examination of the tissue, assessing the color, size, and consistency of the biopsied tissue and select a region of interest to examine under a microscope. A selected area of interest will undergo a multi-step preparation that can take up to several days. The prepared slide is then digitized using a WSI scanner. This scanner allows the capturing of an image of the slide and creates an electronic replica known as a virtual slide. Virtual slides are easy to duplicate, edit, store, catalog, and share. The scanner automatically pre-processes the virtual slide and saves it onto cloud storage. A compression algorithm is applied to reduce the file size before saving. The pathologist will examine the magnified tissue samples using slide viewing and management software for digital pathology. Many modes of action can be used such as angle change, comment additions, and multiple slide views with the use of the software. DP allows for the ability to examine tissue together with other medical data associated such as radiology scans or clinical history. The slide viewing software must be integrated with the radiology information system (RIS) and electronic health record (EHR).

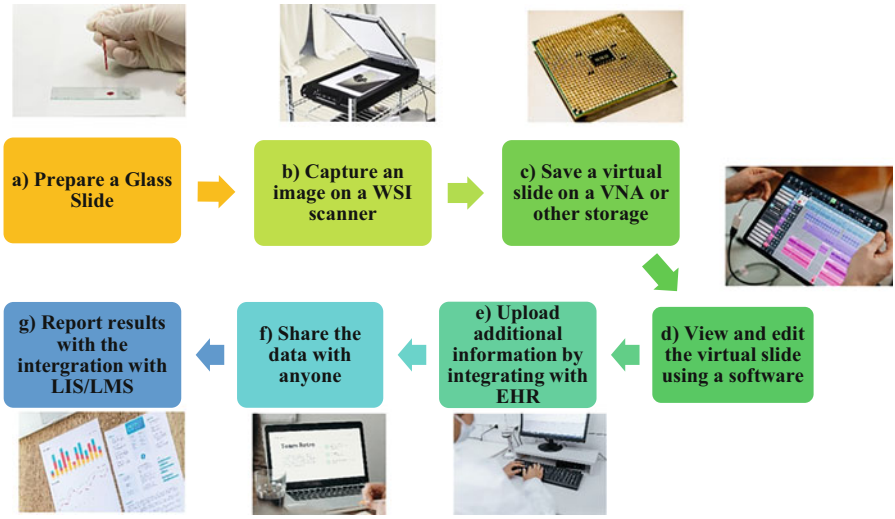


Fig. 5.2 Basic digital pathology workflow. (**a**) Sample preparation on a glass slide. (**b**) Sample capturing using a WSI scanner. (**c**) The virtual slide can be saved on the cloud or another storage system. (**d**) The virtual slide can be viewed or edited using the appropriate software. (**e**) Additional information can be uploaded and integrated with EHR. (**f**) The software allows the virtual slide and data to be shared with anyone, anytime, anywhere. (**g**) The results can be reported with the integration with LIS/LMS

5.3.1 Reporting the Results

Some software has a reporting functionality. Reporting can also be performed via the integration of laboratory information systems (LIS) or laboratory information management systems (LIMS). EHRs also automate sending reports to a physician that can be used for treatment. Advanced digital workflows incorporate machine learning (ML) and AI methods to recognize patterns in tissue samples. This has led to the new discipline of computational pathology.

AI in pathology enables the geographical analysis of data and quantitative accuracy using spatial algorithms (Baxi et al., 2022). DP and AI approaches provide improvements over traditional methods, for example, the enabling of spatial analysis. It provides unbiased, highly precise, and consistent results that are accessed by pathologists (Mroz et al., 2013). DP can assist in the implementation of precision medicine and can be used in conjunction with other techniques such as medical imaging and radiogenomics for improved diagnosis, prognosis, and treatment.

5.4 Radiogenomics and Artificial Intelligence and Its Use in Precision Medicine

Radiogenomics combines medical imaging and genomic profiles for analysis (Liu & Hu, 2022). Radiogenomics has been referred to as a significant advancement in medical imaging, training, analysis, and high-throughput methods that are used to

correlate multiple imaging parameters with genomic data (Pinker et al., 2018) non-invasively. Radiogenomics aims to provide a better understanding of tumor biology as it combines quantitative data from medical images with an individual's genomic data by constructing prediction models (Shui et al., 2020; Liu & Hu, 2022). The goal will be to develop imaging biomarkers that incorporate phenotypic and genotypic components to improve cancer outcomes. Radiogenomics is favored for its non-invasive nature of imaging as well as being cost-effective as it can result in early cancer detection (Pinker et al., 2018). These images, selected features, and associated omics data are used to make in-depth decisions on genomic biomarkers, prognosis, and treatment. Since radiogenomics is a combination of radiomics and genomics, it is important to understand each aspect on its own and how it can be combined.

Radiomics is defined as the "high throughput extraction of quantitative features that results in the conversion of images into mineable data" (Visvikis et al., 2019; Gillies et al., 2015). Genomics refers to the study of the genome of an organism that can be used to find variations. This knowledge can be used to determine health, disease, or drug response (Del Giacco & Cattaneo, 2012; Vailati Riboni et al., 2017). Figure 5.3 shows the basic, overall methodology of radiomics, genomics, and radiogenomics.

The Cancer Imaging Archive (CIA) is the largest medical imaging dataset that is publicly available (Clark et al., 2013). Data from 125 projects have been collected. CIA covers various disease types including Coronavirus Disease 2019 (COVID-19) and cancer. The Cancer Genome Atlas (CGA) platform contains contributing genomic sources (Tomczak et al., 2015). Image information used includes the results from CT, MRI, PET with CT (PET/CT) scans, radiomic features, tumor segmentation of the CT images, and quantitative values measured from the PET/CT images (Liu & Hu, 2022; Shui et al., 2020). Figure 5.4 shows the basic Radiogenomic workflow.

5.4.1 Acquisition of Raw Images

PET/CT and single photon emission CT (SPECT) provide the anatomical and functional details of a tumor. Recently, combinations of quantitative functional assessments such as multiple PET tracers, MRI contrast mechanisms, and PET-MRI, have revealed multi-dimensional tumor phenotypic features (Yankeelov et al., 2014; Matthew et al., 2015; Tixier et al., 2020). For example, diffusion-weighted MRI can provide tumor density and cellularity information. This can be used for cytotoxic treatment monitoring (Anderson et al., 2000). Fluorodeoxyglucose (FDG)-PET is a molecular imaging tool that characterizes metabolic activity changes within the tumor. The uptake, metabolism, and accumulation rate can be used to assess the disease progression and therapeutic effects (Tixier et al., 2020; Eisenhauer et al., 2009; Mu et al., 2020).

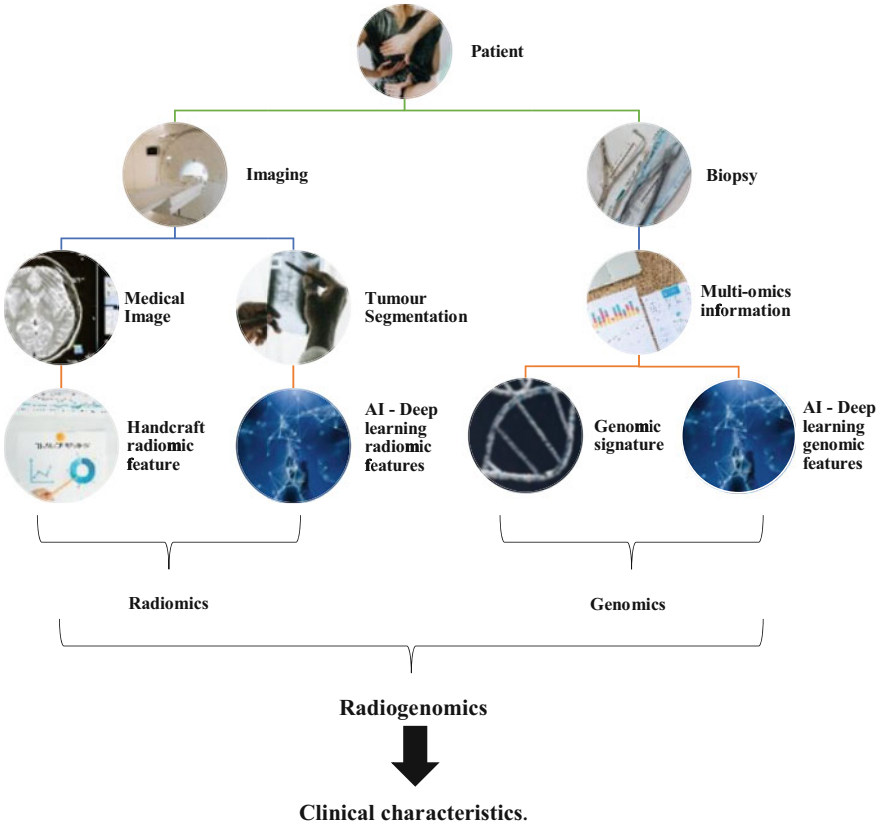


Fig. 5.3 The relationship between radiomics, genomics, and radiogenomics. With radiomics, the patient is taken for medical imaging via CT, MRI, or PET/CT. The medical image is used for data interpretation via graphs and statistical analysis. Based on the image, the tumor can be virtually isolated in a process called tumor segmentation. Tumor segmentation represents the correct identification of the spatial location of a tumor. The sample can be combined with AI to provide improved diagnosis, prognosis, and treatment. Regarding the genomics pathway, the patient undergoes a biopsy. The biopsy material is sent for genetic testing. Various genetic tests provide multiple omics data and genomic signature is evaluated. The information can be combined with AI for improved diagnosis, prognosis, and treatment. Radiomic and genomic data are combined into the new field of radiogenomics. The combination and the knowledge gained by the individual pathways have led to increased survival rates

5.4.2 Pre-processing of Information

It is important that raw image data must be pre-processed to vindicate homogenous and reliable traits. The imaging signals within the region of interest (ROI) can be filtered. However, manual segmentation is often used. Experienced clinicians will extract sufficient and optimal data from the ROI. Sufficient information cannot be

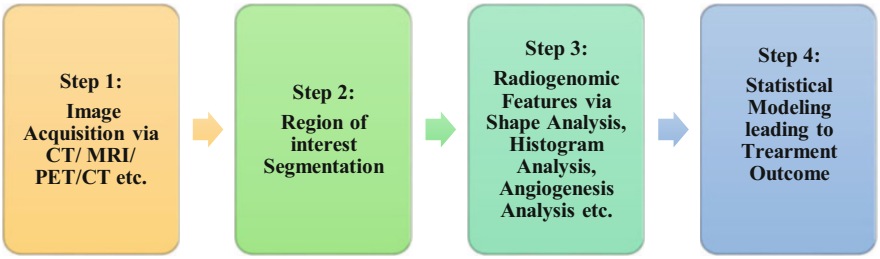


Fig. 5.4 The basic radiogenomics study workflow. Step 1 consists of image acquisition achieved via CT, MRI, PET/CT, etc. Step 2 consists of the identification and segmentation of the region of interest. Step 3 consists of quantitative imaging feature extraction using various types of analysis. Step 4 consists of bioinformatics analysis and data mining

provided if the ROI is minute. A large ROI can lead to bias due to the heterogeneity of the tumor. Some clinicians prefer full manual segmentation, but this is time-consuming and can show inter-observer variability (Parmar et al., 2014; Yip et al., 2017). Automatic segmentation is preferred to manual segregation. Automatic segmentation performance is dependent on the algorithm's accuracy and the ability to differentiate the ROIs from the surrounding tissues. Various machines can manage automatic segregation (Dorador & Rodríguez-Tovar, 2020; Mouawad et al., 2020; Velazquez et al., 2013). For this reason, various studies have proven that the preferred mode of segmentation is semi-automatic (Sensakovic et al., 2011). Tixier et al. compared the robustness of 108 radiomic features using a semi-automatic and an interactive segmentation method (Tixier et al., 2019). The results demonstrated that the interactive method resulted in more robust features than the semi-automatic method (Tixier et al., 2019).

Um et al. used five image pre-processing techniques to extract 420 features from 161 cases (Um et al., 2019). Histogram standardization contributed the most to reducing radiomic feature variability. The results showed that patients can be grouped based on their survival rates (Um et al., 2019). Veeraraghavan et al. developed a novel semi-automatic approach by combining cancer-specific multiparametric Gaussian Mixture Model (GCGMM) and GrowCut (GC) to produce reproducible and accurate segmentations (Veeraraghavan et al., 2018). Segmentation performances were compared in a sample of 75 patients with invasive breast carcinoma. GCGMM's segmentations were shown to be more reproducible when compared to manual delineations and other analyzed segmentation methods (Veeraraghavan et al., 2018).

5.4.3 Extraction of Features

The important factor of radiomics is the extraction of high-dimensional feature sets. These sets can be used to describe the attributes of cancer phenotypes and can be

used to develop prediction models. Radiomic features can be processed by unique software, including PyRadiomics (Liu et al., 2019; Fedorov et al., 2012), CERR (Chen et al., 2020; Apte et al., 2018) or IBEX (Bettinelli et al., 2020; Au-Ger et al., 2018). Morphology-based features can be used to collect three-dimensional (3D) characteristics which can include the volume, surface areas, and sphericity. Intensity-based features can evaluate the gray-level distribution within the ROI and can characterize the overall variability in intensity (first order) and the local distribution (second order). This is referred to as "texture features." Regarding pathology, advanced texture analysis is rising as a novel medical imaging tool for the analysis of intra-tumoral heterogeneity. Texture analysis can be used to analyze the association between the gray-level intensity of pixels and the location within ROIs. Texture analysis constitutes four steps namely extraction, texture discrimination, texture classification, and shape reconstruction (Nie et al., 2019; Ganeshan et al., 2013). Dynamic features are used to quantify uptake in tumors over time and are derived from dynamic contrast-enhanced CT or MRI and metabolic PET. This can provide information regarding the relationships between prognosis and molecular subclassifications of tumors (Sibille et al., 2019).

5.4.4 Data Analysis

The variables and features collected from extraction can be duplicated and can contain unnecessary information. Data selection or analysis is needed to sift through all the information and only keep the significant data. There are three common selection models namely the filter -, wrapper - and embedded methods. The filter methods can evaluate features without involving the model. The wrapper models involve predictor optimization as part of the selection process (Roffo, 2016). The wrapper models provide improved results. However, the filter methods are less expensive. In the embedded methods, the learning part and the feature selection are joined (Roffo, 2016).

Convolutional neural networks (CNNs) are used for deep learning. CNN combines imaging filters with artificial neural networks through a series of layers (Jun et al., 2019). CNNs use local connections and weights to analyze the input images. This is followed by the pooling operations to obtain spatially invariant features (Liu et al., 2020). After sufficient training data is obtained, algorithms can determine the optimal feature set and the importance of each feature (Xu et al., 2019). After the set is obtained, a prediction model is required to connect the features selected with the genetic information. A radiomics model is used to validate the potential for clinical application. A radiogenomics study can be hypothesis or exploratory driven. In exploratory studies, a multiple hypotheses test is common, where features extracted are assessed against various genomic variables. Using the hypothesis-driven approach, scientists collect sufficient imaging phenotypes and investigate it with a specific hypothesis in mind (Konstantinidis et al., 2014).

AI has allowed the analysis of medical images with the absence of human interference. AI and deep learning have led to automated and consistent medical imaging analysis. It is believed that AI can exceed experienced pathologists in cancer diagnosis and prognosis (Gürsoy Çoruh et al., 2021). Radiogenomics with AI will be able to extract features from an image and link these features with specific phenotypes (Dlamini et al., 2022). The phenotype can be associated with changes on a genomic, transcriptomic, translational, and epigenomic level. This information can be used to improve prognosis, diagnostic, and treatment approaches (Rutman & Kuo, 2009). These image features can also be used as survival indicators and predictors (Dlamini et al., 2022). Using AI with radiogenomics increases accuracy and knowledge for improved diagnosis, prognosis, and treatment.

5.4.5 Current Application of Radiogenomics in Oncology

Radiogenomics uses big data analysis approaches (Incoronato et al., 2017). Radiogenomics also provides an in-depth understanding of tumor biology and imaging biomarkers. These approaches have been validated in a variety of tumors (Pinker et al., 2018). Evidence has showed that there is an association between imaging and genomic characteristics of cancers (Gevaert et al., 2017; Nougaret et al., 2017; Vargas et al., 2017; Nougaret et al., 2018; Li et al., 2018; Liu et al., 2017; Mazurowski et al., 2017; Hong et al., 2018; Kickingereder et al., 2016; Cui et al., 2017; Hu et al., 2017; Jamshidi et al., 2013; Hui et al., 2014; Grimm et al., 2015; Mazurowski et al., 2014; Zhu et al., 2019; Yamamoto et al., 2015; Karlo et al., 2013; Li et al., 2019; Kocak et al., 2019; Cen et al., 2019; Shinagare et al., 2015; Kuo et al., 2007; Xia et al., 2018; Miura et al., 2015; Taouli et al., 2017; Sadot et al., 2015; Vlachavas et al., 2019; Lubner et al., 2015; Miles et al., 2014; Chen et al., 2015; Horvat et al., 2019; Lee et al., 2018; Halpenny et al., 2017; Nair et al., 2014; Stoyanova et al., 2016; McCann et al., 2016; Jansen et al., 2018; Zwirner et al., 2019). Radiogenomics in clinical practice must overcome a few challenges. One such challenge is the repeatability and reproducibility of current radiogenomics models (Kang et al., 2018). Shui et al. noted that researchers must consider the variability arising from the use of different equipment, different software, or different clinics (Shui et al., 2020). Standard practice guidelines are crucial to ensure the accuracy and reliability of analytic results in radiogenomics studies (Andreassen et al., 2016). Figure 5.5 shows the combination of radiomics and genomics into Radiogenomics and radiogenomics, DP and AI integration can help advance the future of personalized patient care and precision oncology.

5.5 Limitations

Various AI approaches are ridiculed for not being able to provide information on how the results were obtained. This causes doubt about the prediction accuracy (Sorell et al., 2022). Clinical, legal, and regulatory issues have been a source of

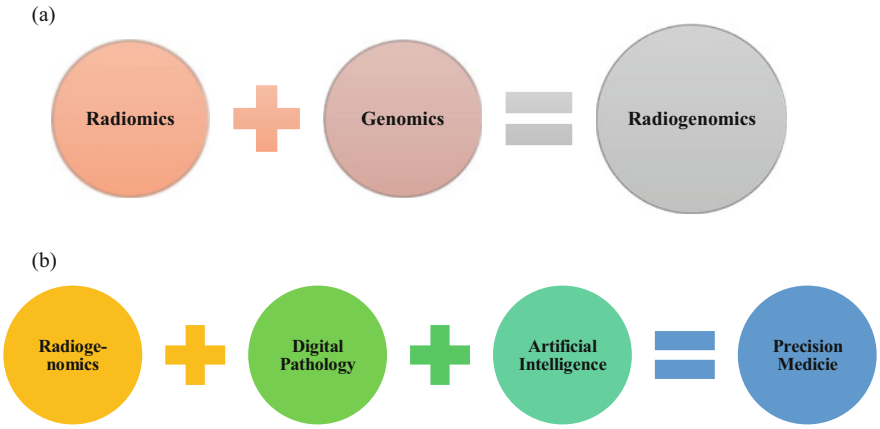


Fig. 5.5 (**a**) The combination of radiomics and genomics into Radiogenomics. The knowledge gained by radiomics and genomics has led to a new field of Radiogenomics. (**b**) Radiogenomics, DP and AI integration can help advance the future of personalized patient care and precision oncology

discussion. Some algorithms are intensive, and research is being conducted to make the algorithms easier to interpret. On the regulatory side, some countries might have restrictions on the use of AI in oncology for example the EU's new General Data Protection Regulation stipulates that "the data subject shall have the right not to be subject to a decision based solely on automated processing" (Consulting, n.d.). AI resulting financial and economic implications are still under review and are speculative. AI development can benefit low-income and middle-income countries positively or negatively. However, AI aims to benefit the pathologists from these countries with faster and more accurate diagnoses. For this reason, there is a need for regulatory control and the devices must be safe and effective. There are developments for example one-shot learning, which uses minute sample amounts for learning. This is useful in pathology where deep learning is challenging due to image size and complexity. In reinforcement learning, algorithms are trained by comparing immediate actions with long-term outcomes to reach a specific goal. Algorithms must be tested to make complex decisions. These algorithms must also be assessed for accuracy. Machine and deep learning by AI must also be assessed or supervised by experts in the field, for example, pathologists, bioinformaticians, and programmers (Pashkov et al., 2020).

AI must be carefully integrated into current clinical practice. Each discipline follows carefully designed patient work-up protocols that include good history taking and a thorough physical examination. This is followed by baseline laboratory blood work and basic radiology (X-ray and ultrasound). An accurate biopsy procedure may require real-time Ultrasound and/or CT/MRI scan guidance. Integrated multidisciplinary clinico-pathologic and radiology correlation leads to well-planned tumor resection and adequate oncology intervention (radiation, chemotherapy or chemo-radiation therapy, with the application of neoadjuvant therapy where necessary).

Some confounding factors that need to be considered when integrating AI into routine cancer management include the following (Table 5.1):

Table 5.1 Factors for consideration when integrating AI in cancer management (Nagahashi et al., 2017; Colomer et al., 2020)

Factors for consideration	Examples of conditions and complexities they may represent
Stage of disease at presentation and first diagnosis	The application of AI protocols and algorithms requires good clinical judgment. Not all patients will benefit from all elements of AI. Patients presenting early, with curable cancers, do not require somatic gene panel analysis. On the opposite end of the spectrum, patients with an advanced, rapidly progressing disease with a short life expectancy, often do not qualify for molecular profiling. These patients are a difficult fit for most AI protocols/algorithms
Composite synchronous and metachronous tumors	AI protocols use specific algorithms based on genomic assays that may not benefit the patient if clinical examination or small, poorly representative biopsies have missed synchronous, metachronous and composite cancers
Cancer occurring in the context of familial syndromes	Clinicians need to select AI protocols that do not miss underlying familial/inherited mutations that patients may not be aware of at presentation
Environmentally induced and enhanced cancers as well as endemic cancers	Exposure to asbestos, ionizing radiation, diet-related carcinogens, smoking, and viral oncogenesis must be carefully woven into AI algorithms. AI algorithms should be designed with endemic cancers in mind
The developing world context	The accuracy of AI algorithms comes with significant financial investment. From instruments to scarce/expensive skills, there are zero margins for error since most developing countries are unable to afford the duplication of highly specialized investigations. In these settings, very good, all-inclusive clinical protocols need to accompany AI algorithms to protect delicate third-world budgets while still benefiting patients
Innate limitations within some genetic level tests. Any discordance at this level will have a negative ripple effect on the entire AI protocol/algorithm that is applied	Static and dynamic tumor heterogeneity. Presence of cancer-associated mutations within the patient's normal tissue Sample age: More than 7 years old tissue (formalin fixed paraffin embedded tissue) is not suitable for NGS analysis (and likely also unsuitable for transcriptomic and proteomic analysis as well)

5.6 Conclusion

Combining digital pathology, radiogenomics and AI can improve workflow and lead to advanced diagnostics. Ultimately, it can lead to detailed and informed cancer diagnosis, prognosis, and treatment. Radiogenomics, DP and AI integration can help advance the future of personalized patient care and precision oncology. Precision medicine aims to offer unique and extraordinary patient care and treatment. Precision medicine would also provide the most accurate information regarding treatment and risk factors. AI can assist in monitoring response, analyzing large data sets, recovery, etc., which will ease the burden on the pathologists and the results will be obtained faster. Although AI has a few limitations, improvements and research are continuously growing to improve AI systems. Precision medicine with AI can improve the prevention of serious diseases.

References

Abernethy, A. P., Etheredge, L. M., Ganz, P. A., Wallace, P., German, R. R., Neti, C., Bach, P. B., & Murphy, S. B. (2010). Rapid-learning system for cancer care. *Journal of Clinical Oncology, 28*, 4268–4274.

Anderson, A. W., Xie, J., Pizzonia, J., Bronen, R. A., Spencer, D. D., & Gore, J. C. (2000). Effects of cell volume fraction changes on apparent diffusion in human cells. *Magnetic Resonance Imaging, 18*, 689–695.

Andreassen, C. N., Schack, L. M. H., Laursen, L. V., & Alsner, J. (2016). Radiogenomics—current status, challenges and future directions. *Cancer Letters, 382*, 127–136.

Apte, A. P., Iyer, A., Crispin-Ortuzar, M., Pandya, R., Van Dijk, L. V., Spezi, E., Thor, M., Um, H., Veeraraghavan, H., Oh, J. H., Shukla-Dave, A., & Deasy, J. O. (2018). Technical note: Extension of cerr for computational radiomics: A comprehensive matlab platform for reproducible radiomics research. *Medical Physics, 45*, 3713–3720.

Au-Ger, R. B., Au-Cardenas, C. E., Au-Anderson, B. M., Au-Yang, J., Au-Mackin, D. S., Au-Zhang, L., & Au-Court, L. E. (2018). Guidelines and experience using imaging biomarker explorer (Ibex) for radiomics. *Jove*, E57132.

Baxi, V., Edwards, R., Montalto, M., & Saha, S. (2022). Digital pathology and artificial intelligence in translational medicine and clinical practice. *Modern Pathology, 35*(1), 23–32. https://doi.org/10.1038/s41379-021-00919-2

Bettinelli, A., Branchini, M., De Monte, F., Scaggion, A., & Paiusco, M. (2020). Technical note: An Ibex adaption toward image biomarker standardization. *Medical Physics, 47*, 1167–1173.

Boellaard, R., O'doherty, M. J., Weber, W. A., Mottaghy, F. M., Lonsdale, M. N., Stroobants, S. G., Oyen, W. J., Kotzerke, J., Hoekstra, O. S., & Pruim, J. (2010). Fdg Pet and Pet/Ct: Eanm procedure guidelines for tumour pet imaging: Version 1.0. *European Journal of Nuclear Medicine and Molecular Imaging, 37*, 181–200.

Caldemeyer, K. S., & Buckwalter, K. A. (1999). The basic principles of computed tomography and magnetic resonance imaging. *Journal of the American Academy of Dermatology, 41*, 768–771.

Cen, D., Xu, L., Zhang, S., Chen, Z., Huang, Y., Li, Z., & Liang, B. (2019). Renal cell carcinoma: Predicting Runx3 methylation level and its consequences on survival with Ct features. *European Radiology, 29*, 5415–5422.

Chen, S.-W., Lin, C.-Y., Ho, C.-M., Chang, Y.-S., Yang, S.-F., Kao, C.-H., & Chang, J.-G. (2015). Genetic alterations in colorectal cancer have different patterns on 18f-Fdg Pet/Ct. *Clinical Nuclear Medicine, 40*.

Chen, Y., Chang, W., Ren, L., Chen, J., Tang, W., Liu, T., Jian, M., Liu, Y., Wei, Y., & Xu, J. (2020). Comprehensive evaluation of relapse risk (Cerr) score for colorectal liver metastases: Development and validation. *The Oncologist, 25*, E1031–E1041.

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., & Prior, F. (2013). The cancer imaging archive (Tcia): Maintaining and operating a public information repository. *Journal of Digital Imaging, 26*, 1045–1057.

Colomer, R., Mondejar, R., Romero-Laorden, N., Alfranca, A., Sanchez-Madrid, F., & Quintela-Fandino, M. (2020). When should we order a next generation sequencing test in a patient with cancer? *Eclinicalmedicine, 25*, 100487.

Consulting, I. (n.d.) *Automated individual decision-making, including profiling.* [Online]. Available: https://Gdpr-Info.Eu [Accessed].

Cui, Y., Ren, S., Tha, K. K., Wu, J., Shirato, H., & Li, R. (2017). Volume of high-risk intratumoral subregions at multi-parametric Mr imaging predicts overall survival and complements molecular analysis of glioblastoma. *European Radiology, 27*, 3583–3592.

Del Giacco, L., & Cattaneo, C. (2012). Introduction to genomics. *Methods in Molecular Biology, 823*, 79–88.

Dlamini, Z., Skepu, A., Kim, N., Mkhabele, M., Khanyile, R., Molefi, T., Mbatha, S., Setlai, B., Mulaudzi, T., Mabongo, M., Bida, M., Kgoebane-Maseko, M., Mathabe, K., Lockhat, Z., Kgokolo, M., Chauke-Malinga, N., Ramagaga, S., & Hull, R. (2022). Ai and precision oncology in clinical cancer genomics: From prevention to targeted cancer therapies-an outcomes based patient care. *Informatics in Medicine Unlocked, 31*, 100965.

Dorador, J., & Rodríguez-Tovar, F. J. (2020). Crossed sequence, a new tool for 3d processing in geosciences using the free software 3dslicer. *Scientific Data, 7*, 270.

Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., & Verweij, J. (2009). New response evaluation criteria in solid tumours: Revised recist guideline (Version 1.1). *European Journal of Cancer, 45*, 228–247.

Farahani, N., Parwani, A. V., & Pantanowitz, L. (2015). Whole slide imaging in pathology: Advantages, limitations, and emerging perspectives. *Pathology and Laboratory Medicine International, 7*, 4321.

Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F., Sonka, M., Buatti, J., Aylward, S., Miller, J. V., Pieper, S., & Kikinis, R. (2012). 3d slicer as an image computing platform for the quantitative imaging network. *Magnetic Resonance Imaging, 30*, 1323–1341.

Fogel, A. L., & Kvedar, J. C. (2018). Artificial intelligence powers digital medicine. *Npj Digital Medicine, 1*, 5.

Ganeshan, B., Goh, V., Mandeville, H. C., Ng, Q. S., Hoskin, P. J., & Miles, K. A. (2013). Non-small cell lung cancer: Histopathologic correlates for texture parameters At Ct. *Radiology, 266*, 326–336.

Gevaert, O., Echegaray, S., Khuong, A., Hoang, C. D., Shrager, J. B., Jensen, K. C., Berry, G. J., Guo, H. H., Lau, C., Plevritis, S. K., Rubin, D. L., Napel, S., & Leung, A. N. (2017). Predictive radiogenomics modeling of Egfr mutation status in lung cancer. *Scientific Reports, 7*, 41674.

Gillies, R. J., Kinahan, P. E., & Hricak, H. (2015). Radiomics: Images are more than pictures, they are data. *Radiology, 278*, 563–577.

Grimm, L. J., Zhang, J., & Mazurowski, M. A. (2015). Computational approach to radiogenomics of breast cancer: Luminal A and luminal B molecular subtypes are associated with imaging features on routine breast mri extracted using computer vision algorithms. *Journal of Magnetic Resonance Imaging, 42*, 902–907.

Gürsoy Çoruh, A., Yenigün, B., Uzun, Ç., Kahya, Y., Büyükceran, E. U., Elhan, A., Orhan, K., & Kayı Cangır, A. (2021). A comparison of the fusion model of deep learning neural networks with human observation for lung nodule detection and classification. *The British Journal of Radiology, 94*, 20210222.

Halpenny, D. F., Plodkowski, A., Riely, G., Zheng, J., Litvak, A., Moscowitz, C., & Ginsberg, M. S. (2017). Radiogenomic evaluation of lung cancer: Are there imaging characteristics associated with lung adenocarcinomas harboring braf mutations? *Clinical Imaging, 42*, 147–151.

Hong, E. K., Choi, S. H., Shin, D. J., Jo, S. W., Yoo, R.-E., Kang, K. M., Yun, T. J., Kim, J.-H., Sohn, C.-H., Park, S.-H., Won, J.-K., Kim, T. M., Park, C.-K., Kim, I. H., & Lee, S. T. (2018). Radiogenomics correlation between Mr imaging features and major genetic profiles in glioblastoma. *European Radiology, 28*, 4350–4361.

Horvat, N., Veeraraghavan, H., Pelossof, R. A., Fernandes, M. C., Arora, A., Khan, M., Marco, M., Cheng, C.-T., Gonen, M., Golia Pernicka, J. S., Gollub, M. J., Garcia-Aguillar, J., & Petkovska, I. (2019). Radiogenomics of rectal adenocarcinoma in the era of precision medicine: A pilot study of associations between qualitative and quantitative Mri imaging features and genetic mutations. *European Journal of Radiology, 113*, 174–181.

Hu, L. S., Ning, S., Eschbacher, J. M., Baxter, L. C., Gaw, N., Ranjbar, S., Plasencia, J., Dueck, A. C., Peng, S., Smith, K. A., Nakaji, P., Karis, J. P., Quarles, C. C., Wu, T., Loftus, J. C., Jenkins, R. B., Sicotte, H., Kollmeyer, T. M., O'neill, B. P., Elmquist, W., Hoxworth, J. M., Frakes, D., Sarkaria, J., Swanson, K. R., Tran, N. L., Li, J., & Mitchell, J. R. (2017). Radiogenomics to characterize regional genetic heterogeneity in glioblastoma. *Neuro-Oncology, 19*, 128–137.

Hui, L., Maryellen, L. G., Li, L., Jyothi, J., & Charlene, A. S. (2014). Comparative analysis of image-based phenotypes of mammographic density and parenchymal patterns in distinguishing between *Brca1/2* Cases, unilateral cancer cases, and controls. *Journal of Medical Imaging, 1*, 1–9.

Incoronato, M., Aiello, M., Infante, T., Cavaliere, C., Grimaldi, A. M., Mirabelli, P., Monti, S., & Salvatore, M. (2017). Radiogenomic analysis of oncological data: A technical survey. *International Journal of Molecular Sciences, 18*.

Jamshidi, N., Diehn, M., Bredel, M., & Kuo, M. D. (2013). Illuminating radiogenomic characteristics of glioblastoma multiforme through integration of Mr imaging, messenger Rna expression, and Dna copy number variation. *Radiology, 270*, 1–2.

Jansen, R. W., De Jong, M. C., Kooi, I. E., Sirin, S., Göricke, S., Brisse, H. J., Maeder, P., Galluzzi, P., Van Der Valk, P., Cloos, J., Eekhout, I., Castelijns, J. A., Moll, A. C., Dorsman, J. C., & De Graaf, P. (2018). Mr imaging features of retinoblastoma: Association with gene expression profiles. *Radiology, 288*, 506–515.

Jubb, A. M., Koeppen, H., & Reis-Filho, J. S. (2014). Pathology in drug discovery and development. *The Journal of Pathology, 232*, 99–102.

Jun, G., Qian, J., Bo, Z., & Daozheng, C. (2019). Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: An overview. *Mathematical Biosciences and Engineering, 16*, 6536–6561.

Kang, J., Rancati, T., Lee, S., Oh, J. H., Kerns, S. L., Scott, J. G., Schwartz, R., Kim, S., & Rosenstein, B. S. (2018). Machine learning and radiogenomics: Lessons learned and future directions. *Frontiers in Oncology, 8*.

Karlo, C. A., Di Paolo, P. L., Chaim, J., Hakimi, A. A., Ostrovnaya, I., Russo, P., Hricak, H., Motzer, R., Hsieh, J. J., & Akin, O. (2013). Radiogenomics of clear cell renal cell carcinoma: Associations between Ct imaging features and mutations. *Radiology, 270*, 464–471.

Katti, G., Ara, S. A., & Shireen, A. (2011). Magnetic resonance imaging (Mri)—a review. *International Journal of Dental Clinics, 3*, 65–70.

Kickingereder, P., Bonekamp, D., Nowosielski, M., Kratz, A., Sill, M., Burth, S., Wick, A., Eidel, O., Schlemmer, H.-P., Radbruch, A., Debus, J., Herold-Mende, C., Unterberg, A., Jones, D., Pfister, S., Wick, W., Von Deimling, A., Bendszus, M., & Capper, D. (2016). Radiogenomics of glioblastoma: Machine learning-based classification of molecular characteristics by using multiparametric and multiregional Mr imaging features. *Radiology, 281*, 907–918.

Kocak, B., Durmaz, E. S., Ates, E., & Ulusan, M. B. (2019). Radiogenomics in clear cell renal cell carcinoma: Machine learning-based high-dimensional quantitative Ct texture analysis in predicting pbrm1 mutation status. *American Journal of Roentgenology, 212*, W55–W63.

König, K., Peifer, M., Fassunke, J., Ihle, M. A., Künstlinger, H., Heydt, C., Stamm, K., Ueckeroth, F., Vollbrecht, C., Bos, M., Gardizi, M., Scheffler, M., Nogova, L., Leenders, F., Albus, K., Meder, L., Becker, K., Florin, A., Rommerscheidt-Fuss, U., Altmüller, J., Kloth, M., Nürnberg, P., Henkel, T., Bikár, S. E., Sos, M. L., Geese, W. J., Strauss, L., Ko, Y. D., Gerigk, U., Odenthal, M., Zander, T., Wolf, J., Merkelbach-Bruse, S., Buettner, R., & Heukamp, L. C. (2015). Implementation of amplicon parallel sequencing leads to improvement of diagnosis and therapy of lung cancer patients. *Journal of Thoracic Oncology, 10*, 1049–1057.

Konstantinidis, I. T., Do, R. K. G., Gultekin, D. H., Gönen, M., Schwartz, L. H., Fong, Y., Allen, P. J., D'angelica, M. I., Dematteo, R. P., Klimstra, D. S., Kemeny, N. E., & Jarnagin, W. R. (2014). Regional chemotherapy for unresectable intrahepatic cholangiocarcinoma: A potential role for dynamic magnetic resonance imaging as an imaging biomarker and a survival update from two prospective clinical trials. *Annals of Surgical Oncology, 21*, 2675–2683.

Kramer, J. A., Sagartz, J. E., & Morris, D. L. (2007). The application of discovery toxicology and pathology towards the design of safer pharmaceutical lead candidates. *Nature Reviews Drug Discovery, 6*, 636–649.

Kuo, M. D., Gollub, J., Sirlin, C. B., Ooi, C., & Chen, X. (2007). Radiogenomic analysis to identify imaging phenotypes associated with drug response gene expression programs in hepatocellular carcinoma. *Journal of Vascular and Interventional Radiology, 18*, 821–830.

Lee, J., Cui, Y., Sun, X., Li, B., Wu, J., Li, D., Gensheimer, M. F., Loo, B. W., Diehn, M., & Li, R. (2018). Prognostic value and molecular correlates of a Ct image-based quantitative pleural contact index in early stage Nsclc. *European Radiology, 28*, 736–746.

Li, Y., Qian, Z., Xu, K., Wang, K., Fan, X., Li, S., Jiang, T., Liu, X., & Wang, Y. (2018). Mri features predict P53 status in lower-grade gliomas via a machine-learning approach. *Neuroimage: Clinical, 17*, 306–311.

Li, Z.-C., Zhai, G., Zhang, J., Wang, Z., Liu, G., Wu, G.-Y., Liang, D., & Zheng, H. (2019). Differentiation of clear cell and non-clear cell renal cell carcinomas by all-relevant radiomics features from multiphase Ct: A Vhl mutation perspective. *European Radiology, 29*, 3996–4007.

Liu, B., Chi, W., Li, X., Li, P., Liang, W., Liu, H., Wang, W., & He, J. (2020). Evolving the pulmonary nodules diagnosis from classical approaches to deep learning-aided decision support: Three decades' development course and future prospect. *Journal of Cancer Research and Clinical Oncology, 146*, 153–185.

Liu, Q., & Hu, P. (2022). Extendable and explainable deep learning for pan-cancer radiogenomics research. *Current Opinion in Chemical Biology, 66*, 102111.

Liu, Q., Jiang, P., Jiang, Y., Ge, H., Li, S., Jin, H., & Li, Y. (2019). Prediction of aneurysm stability using a machine learning model based on pyradiomics-derived morphological features. *Stroke, 50*, 2314–2321.

Liu, X., Mangla, R., Tian, W., Qiu, X., Li, D., Walter, K. A., Ekholm, S., & Johnson, M. D. (2017). The preliminary radiogenomics association between Mr perfusion imaging parameters and genomic biomarkers, and their predictive performance of overall survival in patients with glioblastoma. *Journal of Neuro-Oncology, 135*, 553–560.

Lubner, M. G., Stabo, N., Lubner, S. J., Del Rio, A. M., Song, C., Halberg, R. B., & Pickhardt, P. J. (2015). Ct textural analysis of hepatic metastatic colorectal cancer: Pre-treatment tumor heterogeneity correlates with pathology and clinical outcomes. *Abdominal Imaging, 40*, 2331–2337.

Matthew, J. N., Fei, Y., Darrin, B., Stephen, R. B., George, A. S., & Paul, E. K. (2015). Quantitative radiomics: Impact of stochastic effects on textural feature analysis implies the need for standards. *Journal of Medical Imaging, 2*, 1–13.

Mazurowski, M. A., Clark, K., Czarnek, N. M., Shamsesfandabadi, P., Peters, K. B., & Saha, A. (2017). Radiogenomics of lower-grade glioma: Algorithmically-assessed tumor shape is associated with tumor genomic subtypes and patient outcomes in a multi-institutional study with the cancer genome atlas data. *Journal of Neuro-Oncology, 133*, 27–35.

Mazurowski, M. A., Zhang, J., Grimm, L. J., Yoon, S. C., & Silber, J. I. (2014). Radiogenomic analysis of breast cancer: Luminal B molecular subtype is associated with enhancement dynamics at Mr imaging. *Radiology, 273*, 365–372.

Mccann, S. M., Jiang, Y., Fan, X., Wang, J., Antic, T., Prior, F., Vanderweele, D., & Oto, A. (2016). Quantitative multiparametric Mri features and Pten expression of peripheral zone prostate cancer: A pilot study. *American Journal of Roentgenology, 206*, 559–565.

Meyer, H.-J., Wienke, A., & Surov, A. (2019). Associations between Glut expression and Suv Values derived from Fdg-pet in different tumors—a systematic review and meta analysis. *PLoS One, 14*, E0217781.

Miles, K. A., Ganeshan, B., Rodriguez-Justo, M., Goh, V. J., Ziauddin, Z., Engledow, A., Meagher, M., Endozo, R., Taylor, S. A., Halligan, S., Ell, P. J., & Groves, A. M. (2014). Multifunctional imaging signature for V-Ki-Ras2 Kirsten rat sarcoma viral oncogene homolog (Kras) mutations in colorectal cancer. *Journal of Nuclear Medicine, 55*, 386.

Miura, T., Ban, D., Tanaka, S., Mogushi, K., Kudo, A., Matsumura, S., Mitsunori, Y., Ochiai, T., Tanaka, H., & Tanabe, M. (2015). Distinct clinicopathological phenotype of hepatocellular carcinoma with ethoxybenzyl-magnetic resonance imaging hyperintensity: Association with gene expression signature. *The American Journal of Surgery, 210*, 561–569.

Mouawad, M., Biernaski, H., Brackstone, M., Lock, M., Kornecki, A., Shmuilovich, O., Ben-Nachum, I., Prato, F. S., Thompson, R. T., Gaede, S., & Gelman, N. (2020). The effect of registration on voxel-wise tofts model parameters and uncertainties from Dce-Mri of early-stage breast cancer patients using 3dslicer. *Journal of Digital Imaging, 33*, 1065–1072.

Mroz, P., Parwani, A. V., & Kulesza, P. (2013). Central pathology review for phase III clinical trials: The enabling effect of virtual microscopy. *Archives of Pathology & Laboratory Medicine, 137*(4), 492–495. https://doi.org/10.5858/arpa.2012-0093-RA

Mu, W., Tunali, I., Gray, J. E., Qi, J., Schabath, M. B., & Gillies, R. J. (2020). Radiomics of 18f-Fdg Pet/Ct images predicts clinical benefit of advanced Nsclc patients to checkpoint blockade immunotherapy. *European Journal of Nuclear Medicine and Molecular Imaging, 47*, 1168–1182.

Nagahashi, M., Shimada, Y., Ichikawa, H., Nakagawa, S., Sato, N., Kaneko, K., Homma, K., Kawasaki, T., Kodama, K., Lyle, S., Takabe, K., & Wakai, T. (2017). Formalin-fixed paraffin-embedded sample conditions for deep next generation sequencing. *The Journal of Surgical Research, 220*, 125–132.

Nair, V. S., Gevaert, O., Davidzon, G., Plevritis, S. K., & West, R. (2014). Nf-kB protein expression associates with ^{18}F-Fdg pet tumor uptake in non-small cell lung cancer: A radiogenomics validation study to understand tumor metabolism. *Lung Cancer, 83*, 189–196.

Niazi, M. K. K., Parwani, A. V., & Gurcan, M. N. (2019). Digital pathology and artificial intelligence. *The Lancet Oncology, 20*, E253–E261.

Nie, K., Al-Hallaq, H., Li, X. A., Benedict, S. H., Sohn, J. W., Moran, J. M., Fan, Y., Huang, M., Knopp, M. V., Michalski, J. M., Monroe, J., Obcemea, C., Tsien, C. I., Solberg, T., Wu, J., Xia, P., Xiao, Y., & El Naqa, I. (2019). Nctn assessment on current applications of radiomics in oncology. *International Journal of Radiation Oncology, Biology, Physics, 104*, 302–315.

Nougaret, S., Lakhman, Y., Gönen, M., Goldman, D. A., Miccò, M., D'anastasi, M., Johnson, S. A., Juluru, K., Arnold, A. G., Sosa, R. E., Soslow, R. A., Vargas, H. A., Hricak, H., Kauff, N. D., & Sala, E. (2017). High-grade serous ovarian cancer: Associations between Brca mutation Status, Ct imaging phenotypes, and clinical outcomes. *Radiology, 285*, 472–481.

Nougaret, S., Lakhman, Y., Molinari, N., Feier, D., Scelzo, C., Vargas, H. A., Sosa, R. E., Hricak, H., Soslow, R. A., Grisham, R. N., & Sala, E. (2018). Ct features of ovarian tumors: Defining key differences between serous borderline tumors and low-grade serous carcinomas. *American Journal of Roentgenology, 210*, 918–926.

Parmar, C., Rios Velazquez, E., Leijenaar, R., Jermoumi, M., Carvalho, S., Mak, R. H., Mitra, S., Shankar, B. U., Kikinis, R., Haibe-Kains, B., Lambin, P., & Aerts, H. J. W. L. (2014). Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS One, 9*, E102107.

Pashkov, V. M., Harkusha, A. O., & Harkusha, Y. O. (2020). Artificial intelligence in medical practice: Regulative issues and perspectives.

Pekar, J. J. (2006). A brief introduction to functional Mri. *IEEE Engineering in Medicine and Biology Magazine, 25*, 24–26.

Pinker, K., Shitano, F., Sala, E., Do, R. K., Young, R. J., Wibmer, A. G., Hricak, H., Sutton, E. J., & Morris, E. A. (2018). Background, current role, and potential applications of radiogenomics. *Journal of Magnetic Resonance Imaging, 47*, 604–620.

Rivas, M. A., Beaudoin, M., Gardet, A., Stevens, C., Sharma, Y., Zhang, C. K., Boucher, G., Ripke, S., Ellinghaus, D., Burtt, N., Fennell, T., Kirby, A., Latiano, A., Goyette, P., Green, T., Halfvarson, J., Haritunians, T., Korn, J. M., Kuruvilla, F., Lagacé, C., Neale, B., Lo, K. S., Schumm, P., Törkvist, L., Dubinsky, M. C., Brant, S. R., Silverberg, M. S., Duerr, R. H., Altshuler, D., Gabriel, S., Lettre, G., Franke, A., D'amato, M., Mcgovern, D. P., Cho, J. H., Rioux, J. D., Xavier, R. J., & Daly, M. J. (2011). Deep resequencing of Gwas loci identifies independent rare variants associated with inflammatory bowel disease. *Nature Genetics, 43*, 1066–1073.

Roffo, G. (2016). *Feature selection library* (*Matlab toolbox*).

Rutman, A. M., & Kuo, M. D. (2009). Radiogenomics: Creating a link between molecular diagnostics and diagnostic imaging. *European Journal of Radiology, 70*, 232–241.

Sadot, E., Simpson, A. L., Do, R. K. G., Gonen, M., Shia, J., Allen, P. J., D'angelica, M. I., Dematteo, R. P., Kingham, T. P., & Jarnagin, W. R. (2015). Cholangiocarcinoma: Correlation between molecular profiling and imaging phenotypes. *PLoS One, 10*, E0132953.

Sensakovic, W. F., Armato Iii, S. G., Straus, C., Roberts, R. Y., Caligiuri, P., Starkey, A., & Kindler, H. L. (2011). Computerized segmentation and measurement of malignant pleural mesothelioma. *Medical Physics, 38*, 238–244.

Shinagare, A. B., Vikram, R., Jaffe, C., Akin, O., Kirby, J., Huang, E., Freymann, J., Sainani, N. I., Sadow, C. A., Bathala, T. K., Rubin, D. L., Oto, A., Heller, M. T., Surabhi, V. R., Katabathina, V., & Silverman, S. G. (2015). Radiogenomics of clear cell renal cell carcinoma: Preliminary findings of the cancer genome atlas–renal cell carcinoma (Tcga–Rcc) imaging research group. *Abdominal Imaging, 40*, 1684–1692.

Shui, L., Ren, H., Yang, X., Li, J., Chen, Z., Yi, C., Zhu, H., & Shui, P. (2020). The era of radiogenomics in precision medicine: An emerging approach to support diagnosis, treatment decisions, and prognostication in oncology. *Frontiers in Oncology, 10*, 570465.

Sibille, L., Seifert, R., Avramovic, N., Vehren, T., Spottiswoode, B., Zuehlsdorff, S., & Schäfers, M. (2019). 18f-Fdg Pet/Ct uptake classification in lymphoma and lung cancer by using deep convolutional neural networks. *Radiology, 294*, 445–452.

Sorell, T., Rajpoot, N., & Verrill, C. (2022). Ethical issues in computational pathology. *Journal of Medical Ethics, 48*, 278–284.

Stoyanova, R., Pollack, A., Takhar, M., Lynne, C., Parra, N., Lam, L. L. C., Alshalalfa, M., Buerki, C., Castillo, R., Jorda, M., Ashab, H. A.-D., Kryvenko, O. N., Punnen, S., Parekh, D. J., Abramowitz, M. C., Gillies, R. J., Davicioni, E., Erho, N., & Ishkanian, A. (2016). Association of multiparametric Mri quantitative imaging features with prostate cancer gene expression in Mri-targeted prostate biopsies. *Oncotarget, 7*(33).

Taouli, B., Hoshida, Y., Kakite, S., Chen, X., Tan, P. S., Sun, X., Kihira, S., Kojima, K., Toffanin, S., Fiel, M. I., Hirschfield, H., Wagner, M., & Llovet, J. M. (2017). Imaging-based surrogate markers of transcriptome subclasses and signatures in hepatocellular carcinoma: Preliminary results. *European Radiology, 27*, 4472–4481.

Tixier, F., Cheze-Le-Rest, C., Schick, U., Simon, B., Dufour, X., Key, S., Pradier, O., Aubry, M., Hatt, M., Corcos, L., & Visvikis, D. (2020). Transcriptomics in cancer revealed by positron emission tomography radiomics. *Scientific Reports, 10*, 5660.

Tixier, F., Um, H., Young, R. J., & Veeraraghavan, H. (2019). Reliability of tumor segmentation in glioblastoma: Impact on the robustness of Mri-radiomic features. *Medical Physics, 46*, 3582–3591.

Tizhoosh, H. R., & Pantanowitz, L. (2018). Artificial intelligence and digital pathology: Challenges and opportunities. *Journal of Pathology Informatics, 9*, 38.

Tomczak, K., Czerwińska, P., & Wiznerowicz, M. (2015). Reviewthe cancer genome atlas (Tcga): An immeasurable source of knowledge. *Contemporary Oncology/Współczesna Onkologia*, 68–77.

Um, H., Tixier, F., Bermudez, D., Deasy, J. O., Young, R. J., & Veeraraghavan, H. (2019). Impact of image preprocessing on the scanner dependence of multi-parametric Mri radiomic features and covariate shift in multi-institutional glioblastoma datasets. *Physics in Medicine & Biology, 64*, 165011.

Vailati Riboni, M., Palombo, V. & Loor, J. (2017). *What are omics sciences*?

Vargas, H. A., Veeraraghavan, H., Micco, M., Nougaret, S., Lakhman, Y., Meier, A. A., Sosa, R., Soslow, R. A., Levine, D. A., Weigelt, B., Aghajanian, C., Hricak, H., Deasy, J., Snyder, A., & Sala, E. (2017). A novel representation of inter-site tumour heterogeneity from pre-treatment computed tomography textures classifies ovarian cancers by clinical outcome. *European Radiology, 27*, 3991–4001.

Veeraraghavan, H., Dashevsky, B. Z., Onishi, N., Sadinski, M., Morris, E., Deasy, J. O., & Sutton, E. J. (2018). Appearance constrained semi-automatic segmentation from Dce-Mri is reproducible and feasible for breast cancer radiomics: A feasibility study. *Scientific Reports, 8*, 4838.

Velazquez, E. R., Parmar, C., Jermoumi, M., Mak, R. H., Van Baardwijk, A., Fennessy, F. M., Lewis, J. H., De Ruysscher, D., Kikinis, R., Lambin, P., & Aerts, H. J. W. L. (2013). Volumetric Ct-based segmentation of Nsclc using 3d-slicer. *Scientific Reports, 3*, 3529.

Visvikis, D., Cheze Le Rest, C., Jaouen, V., & Hatt, M. (2019). Artificial intelligence, machine (Deep) learning and Radio(Geno)Mics: Definitions and nuclear medicine imaging applications. *European Journal of Nuclear Medicine and Molecular Imaging, 46*, 2630–2637.

Vlachavas, E. I., Pilalis, E., Papadodima, O., Koczan, D., Willis, S., Klippel, S., Cheng, C., Pan, L., Sachpekidis, C., Pintzas, A., Gregoriou, V., Dimitrakopoulou-Strauss, A., & Chatziioannou, A. (2019). Radiogenomic analysis of F-18-fluorodeoxyglucose positron emission tomography and gene expression data elucidates the epidemiological complexity of colorectal cancer landscape. *Computational and Structural Biotechnology Journal, 17*, 177–185.

Xia, W., Chen, Y., Zhang, R., Yan, Z., Zhou, X., Zhang, B., & Gao, X. (2018). Radiogenomics of hepatocellular carcinoma: Multiregion analysis-based identification of prognostic imaging biomarkers by integrating gene data—a preliminary study. *Physics in Medicine & Biology, 63*, 035044.

Xu, Y., Hosny, A., Zeleznik, R., Parmar, C., Coroller, T., Franco, I., Mak, R. H., & Aerts, H. J. W. L. (2019). Deep learning predicts lung cancer treatment response from serial medical imaging. *Clinical Cancer Research, 25*, 3266–3275.

Yamamoto, S., Han, W., Kim, Y., Du, L., Jamshidi, N., Huang, D., Kim, J. H., & Kuo, M. D. (2015). Breast cancer: Radiogenomic biomarker reveals associations among dynamic contrast-enhanced Mr imaging, long noncoding Rna, and metastasis. *Radiology, 275*, 384–392.

Yankeelov, T. E., Abramson, R. G., & Quarles, C. C. (2014). Quantitative multimodality imaging in cancer research and therapy. *Nature Reviews. Clinical Oncology, 11*, 670–680.

Yip, S. S. F., Parmar, C., Blezek, D., Estepar, R. S. J., Pieper, S., Kim, J., & Aerts, H. J. W. L. (2017). Application of the 3d slicer chest imaging platform segmentation algorithm for large lung nodule delineation. *PLoS One, 12*, E0178944.

Zarella, M. D., Bowman, D., Aeffner, F., Farahani, N., Xthona, A., Absar, S. F., Parwani, A., Bui, M., & Hartman, D. J. (2019). A practical guide to whole slide imaging: A white paper from the digital pathology association. *Archives of Pathology & Laboratory Medicine, 143*, 222–234.

Zhu, Z., Albadawy, E., Saha, A., Zhang, J., Harowicz, M. R., & Mazurowski, M. A. (2019). Deep learning for identifying radiogenomic associations in breast cancer. *Computers in Biology and Medicine, 109*, 85–90.

Zwirner, K., Hilke, F. J., Demidov, G., Socarras Fernandez, J., Ossowski, S., Gani, C., Thorwarth, D., Riess, O., Zips, D., Schroeder, C., & Welz, S. (2019). Radiogenomics in head and neck cancer: Correlation of radiomic heterogeneity and somatic mutations in Tp53, Fat1 and Kmt2d. *Strahlentherapie und Onkologie, 195*, 771–779.

Part II
Artificial Intelligence and Omics in Precision Oncology

Chapter 6
Epigenetics Analysis Using Artificial Intelligence in the Era of Precision Oncology

Rodney Hull, Serwalo Ramagaga, Nomsa Nkosi, Rahaba Marina, Rosemary I. Kabahuma, and Zodwa Dlamini

Abstract It has become a common practice? to use omics data to characterise and diagnose cancer. This has mainly been in the form of genomic mutations, transcriptome and proteomic profiles that are unique or more common in cancer. However, it has become obvious that for complex diseases like cancer, it is important and incredibly useful to examine the genetic changes that occur to alter gene expression that do not rely on changes to the genetic sequence. This is the epigenome and its ability to change according to and to reflect an individual's environment and lifestyle means that it can be directly linked to risk factors for cancer. Epigenome changes have been reported in various cancers. The epigenome also provides data concerning the predisposition of an individual to develop cancer, at the earlier stages of the disease. Like all omics data, epigenomics is considered 'big data' and its analysis and interpretation can only be realistically undertaken through the use of AI and machine learning. However, it is in the combined analysis of genomic, epigenomic and expression data that omics data can give true insight into the underlying molecular basis of cancer. The integration and analysis of these different

R. Hull (✉) · R. Marina · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: rodney.hull@up.ac.za; rahaba.marima@up.ac.za; zodwa.dlamini@up.ac.za

S. Ramagaga
Department of Head and Neck, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: serwalo.ramagaga@up.ac.za

N. Nkosi
Department of Colorectal Surgery, University of the Witwatersrand, Johannesburg, South Africa
e-mail: nomsankosi@dgcolorectal.co.za

R. I. Kabahuma
Department of Otorhinolaryngology, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: rosemary.kabahuma@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_6

data requires the use of specialised AI as machine learning models. This chapter will cover various AI applications in epigenomics-driven precision oncology. Challenges and opportunities in AI-enhanced cancer epigenomics will be discussed.

Keywords Epigenetics · Analysis · Artificial intelligence · Precision oncology · Omics · Cancer · Environment · Lifestyle

6.1 Introduction

Personalised medicine relies on the selection, analysis and integration of information from different 'omics' sources utilised alongside patient and medical data (Rauschert et al., 2020). The ability to integrate and utilise these large data sets would not be possible using conventional analysis. It would take too long to be of any use in any clinical setting. However, Artificial intelligence (AI) algorithms can be designed and trained to accomplish these tasks rapidly, efficiently while addressing bias challenges (Toh et al., 2019). In this way, AI allows clinicians and researchers to make use of and manage big data or large sets of complex biological data (Rauschert et al., 2020). Traditionally, when discussing omics data, the three sources that come to mind are genetic or DNA sequence data (genomics), RNA sequence and expression data (transcriptomics) and protein structure and expression data (proteomics). However, changes in gene expression that occurs at the level of DNA but do not involve changes in the DNA sequence, is known as epigenetics, or above the genome/ Epigenetics is a vital part of an individual's molecular biology and gives rise to another source of omics data, epigenomics. Epigenetic modifications fall can be categorised into: DNA methylation, histone post-translational modifications and variants, nuclear organisation and finally, regulation by small non-coding RNAs (Wen et al., 2017).

Epigenetic modifications of DNA are the result of an individual environment interacting and altering DNA in such a way as to alter gene expression. As such it plays a vital role in the development and progression of diseases such as cancer (Romanowska & Joshi, 2019) (Rauschert et al., 2020). Since these epigenetic modifications serve as a link between environmental contributions to a disease through genetic alterations, epigenetic modifications can generally be identified in the early stages of cancer and thus be useful tools in early cancer detection. They are found in the coding as well as non-coding regions of the DNA (García-Giménez et al., 2012). Because of this, the ability to detect these epigenetic modifications could be used as a biomarker for the diagnosis and prognosis of cancers in the early stages of the disease. Other advantages to the use of epigenetic markers as biomarkers are their stability in fluids such as blood and their presence in cell-free Nucleic Acids (cfNAs). The role of liquid biopsies in early cancer detection, diagnosis and accurate prognosis is an emerging key area towards precision oncology (Wen et al., 2017).

6.2 Types of Epigenetic Modifications

6.2.1 DNA Methylation

DNA methylation represses gene expression. It has been known for some time that changes in DNA methylation occur at cytosine-phosphate-guanine (CpG) sites, often referred to as CpG islands, during the development and progression of cancer (Heyn & Esteller, 2012) (Fig. 6.1). It also occurs in repetitive genomic regions (satellite DNA) and is often found in parasitic elements (long or short interspersed transposable elements or LINES and SINES) (Robertson, 2005). DNA methyltransferases (DNMTs) transfer a methyl group to the 5′ carbon of cytosine nucleotides adjacent to guanines (Reyngold & Chan, 2018). Methylation stops gene transcription by either masking the DNA, preventing transcription factors from binding or by recruiting methyl-CpG binding proteins (MBPs) that silence genes by binding to methylated CpG sites, and recruiting chromatin remodelling molecules (Bird, 2002). Cancer can be caused by the incorrect methylation and resulting in silencing of genes such as tumour suppressor and DNA repair genes, as well as genes that control the cell cycle and genes that regulate genome integrity (Paulsen & Ferguson-Smith, 2001). Cancers have been shown to have a drastically altered methylation pattern and generally lower methylation rates but with hypermethylation occurring in genes that are normally unmethylated (Reyngold &

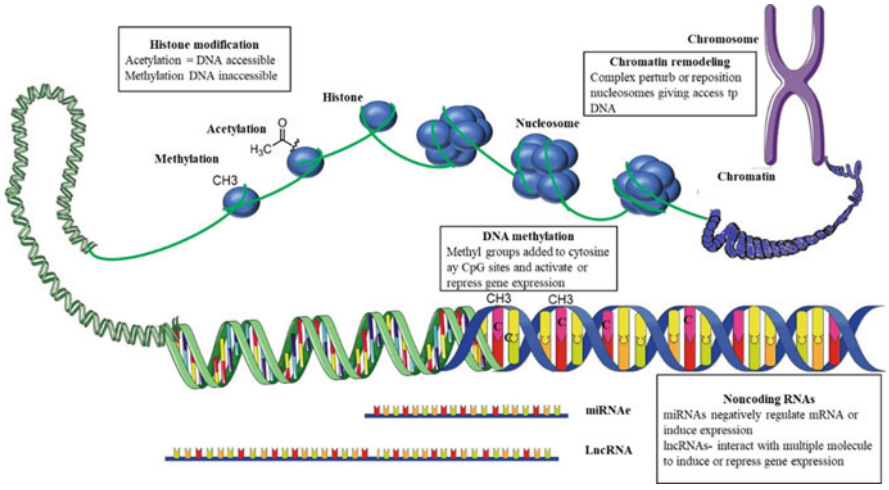


Fig. 6.1 Epigenetic modifications. A representation of epigenetic changes and what role they play in altered gene expression. Chromatin remodelling changes the structure of chromatin by altering the position or stricture of nucleosomes. This changes the accessibility of genes for transcription. Histone modification describes the covalent addition of groups or whole proteins to the tail or core of the histone protein. These groups alter the affinity the histone has for DNA and can either increase or repress gene expression. DNA methylation describes the addition of methyl groups to cytosine bases on DNA inducing or repressing gene expression

Chan, 2018; Laget & Defossez, 2008). Changes in DNA methylation have until recently been measured using microarray assays. The most popular of which is the Illumina HumanMethylation Infinium BeadArray (Kurdyukov & Bullock, 2016). Another method involves the use of methylation-sensitive PCR. These techniques are now being superseded by next-generation sequencing (NGS) in the form of deep-amplicon bisulphite sequencing, which can measure DNA methylation at defined genomic loci. These changes in methylation are then quantified using numerous methods such as affinity enrichment strategies and methods involving bisulphite conversion (Singer, 2019).

DNA methylation is also known to change due to environmental exposure to certain carcinogens. This allows for better risk evaluation in screening for cancers, as well as improved diagnosis and prognosis (How Kit et al., 2012). These changes in DNA methylation can also be targeted for the development of new therapies and opens up another avenue for personalised medicine in cancer management (Jones et al., 2016). Based on this, changes in DNA methylation have been successfully used as an epigenetic biomarker for colorectal cancer. The biomarker, mSEPT9, is a region of the *Septin 9* gene that was found to be methylated in approximately 90% of all colorectal cancer patients. This biomarker has the added advantages that it can be used to diagnose colorectal cancer from blood plasma (Payne, 2010), because DNA methylation is chemically and biologically stable for an extended period of time (How Kit et al., 2012). However, relying only on blood samples for certain types of cancer may not be an effective diagnostic test as many methylation patterns are tissue-specific (Sina et al., 2018).

Large amounts of training data exist for learning algorithms that need to be taught to recognise and assign meaning to methylation profiles. This data can be found in large-scale, data-rich repositories. These include The Cancer Genome Atlas (TCGA), ENCODE, and the BLUEPRINT (Aryee et al., 2014; Jaffe et al., 2012; Leung et al., 2015).

In addition to DNA, mRNA transcripts can also be methylated. Most commonly at an adenosine, the N^6 methyladenosine (m^6A), which are located at the sequences GAC or AAC (Csepany et al., 1990). In the 3′ untranslated regions (3′UTRs). These regions regulate RNA stability, subcellular localisation and translation (Meyer et al., 2012). These modifications can affect alternative splicing, translation, translocation, and degradation (Wang et al., 2016). Many of the proteins that regulate the methylation of mRNA are associated with several cancers where their increased activity silences the expression of tumour suppressor genes and pro-apoptotic genes (Zhong et al., 2019).

6.2.2 RNA Regulation

Non-coding RNAs include pseudogenes, transposon elements, repeated non-coding sequences, regulatory elements, non-coding genes. About 98% of genetic transcripts encode non-coding RNAs (ncRNAs) that play an important role in regulating gene

expression (Comfort, 2015). NcRNAs include microRNAs (miRNA), transcribed ultraconservative regions (T-UCR), small nucleolar RNA (sno-RNA), PIWI-interacting RNA (piRNA), large intergenic non-coding RNA (lincRNA), and long non-coding RNA (lncRNA) (Esteller, 2011). Of these, miRNAs and lncRNAs are the most studied (Fig. 6.1). MicroRNAs (miRNAs) regulate gene expression post-transcriptionally. They do this by repressing translational, degrading mRNA and by targeting promoter sequences to activate gene expression (Place et al., 2008). In this way, miRNAs are involved in the positive and negative regulation of cancer pathways such as proliferation, differentiation, cell cycle regulation, apoptosis, development and stress response (Piletič & Kunej, 2016). They act as part of the epigenetic machinery by binding to specific promoter sites or by recruiting other epigenetic regulators. They themselves can also be regulated by other epigenetic factors. For instance, their expression can be activated or repressed by epigenetic mechanisms such as histone modifications, DNA methylation or RNA methylation (Piletič & Kunej, 2016). Deregulation of miRNA expression is known to contribute to the development of cancer, since they can act as a tumour suppressor or as an oncogene. This means that their silencing or overexpression can either promote or inhibit cancer development and progression (Piletič & Kunej, 2016).

The transcriptional activators of the SWI/SNF complex control the synthesis of long non-coding RNAs (lncRNAs) through the activity of RNA polymerase II. They have multiple roles and can interact with many proteins, DNA and RNA molecules. They are thought to make up most genome transcripts. LncRNAs play an important role in many physiological processes and like miRNA, many of these are related to cancer. These processes include development, differentiation, and proliferation. They are also part of the epigenetic machinery, playing a role in chromatin remodelling, transcriptional and post-transcriptional regulation, splicing regulation, X chromosome inactivation and genomic imprinting (Bhat et al., 2016; Romero-Barrios et al., 2018). Like miRNAs their deregulation is implicated in the growth and development of cancer. As previously stated, lncRNAs play a role in epigenetic modification by recruiting the complexes responsible for chromatin remodelling to specific chromatin loci (Chang et al., 2006). LncRNas can also modify the activity of transcription factors by acting as cofactors. They can also recruit specific RNA-binding proteins to specific gene promoters (Bhat et al., 2016). They play a role in post-transcriptional regulation by controlling translation by binding to and blocking mRNAs and bind to and sequester multiple miRNAs (Filipowicz et al., 2008).

It is known that lncRNAs can silence and activate genes to contribute to the development and progressions of breast cancer. As such machine learning algorithms were designed to identify immune-related lncRNAs and categorise them according to their relationship with patient survival, disease severity, disease progression. The AI identified 43 lncRNAs which were then used to construct a prognostic array examining non-coding gene signatures of the expression of these lncRNAs. This immune-related lncRNA pair (IRLP) signature was found to be highly sensitive and specific for predicting the survival rates of breast cancer patients as well as being able to classify their molecular subtypes of their cancer. For

instance, low-risk signatures were associated with a longer survival time and were classified as tumours expressing low levels of macrophage M2 and high expression levels of biomarkers indicating immunosuppression. High-risk signatures were associated with a lower survival time and were associated with the activation of the MAPK, Jak-STAT and Erb signalling, pathways (Zhu et al., 2022). Machine learning algorithms were also used to build a prognostic toll for skin melanoma based on constructing a competing endogenous RNA (ceRNA) network. This was based on analysing the expression profiles of lncRNAs, miRNAs and mRNAs in melanomas from the TCGA, the Gene Ontology (GO) database, and Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway databases. The algorithms identified all the differentially expressed lncRNAs, miRNAs and mRNAs from these databases by comparing metastatic melanoma data to normal skin tissue data. The algorithm constructed ceRNAs for metastatic melanoma and these different expression profiles were then correlated to survival data. This integrative analysis identified lncRNA, miRNA and mRNA biomarkers in an active ceRNA network in metastatic melanoma that could be used to classify melanomas based on patient survival outcomes (Wang et al., 2019).

6.2.3 Histone Modifications

Histones are the structural proteins that interact with DNA allowing it to be packaged in the nucleus as chromatin. The histone DNA complex creates repeating units called nucleosomes. The nucleosomes consist of eight histone proteins surrounded by a 147 bp of DNA connected by a short DNA linker (Kouzarides, 2007). The regulation of transcription and gene expression occurs through the regulated unpacking of the DNA from this complex so it can be accessed by the transcription machinery (Kouzarides, 2007). Histones are modified at their N-terminal tails or in their core domain and modification occurs following intrinsic and extrinsic changes throughout the life of the cell. The tail region of histones has a high contingent of basic lys/arg and hydroxyl group-containing Ser/Thr/Tyr. The tails are easily accessible and are therefore 'open' to post-translation modifications through the formation of covalent bonds (Fig. 6.1) on these exposed lysines, arginines, sernes and threonines. These modifications include acetylation, methylation, phosphorylation, ubiquitylation, and SUMOylation (Cheng et al., 2019). Additionally, these modifications can alter the way in which the histones interact with the DNA, which can in turn alter the stability of the nucleosome and any interactions DNA has with any DNA interacting molecule (Bowman & Poirier, 2015). Modifications to the core of the histone happen less often as the core is not readily accessible and the modifications are normally the result of the activity of chromatin remodels (Bowman & Poirier, 2015).

The assessment of histone modification in melanoma cell lines was carried out as part of an assessment of genomic sequence variation and mutation on gene

expression. This also included assessing chromatin accessibility and transcriptome and proteome data to generate phased whole genome with epigenetic and expression data for 10 melanoma cell lines. A specialised AI tool was trained for this task and was named DeepMEL2. This tool was able to use this data to identify signalling pathways, regulatory gene networks, and mutations affecting the epigenetic regulation of these pathways and networks that control melanocytic and mesenchymal-like melanoma cell states. The Ai detected thousands of allele-specific chromatin accessibility variants (ASCAVs) in each cell line's genome, 15–20% of which led to changes in transcription factor binding sites. This AI tool is therefore able to identify the presence o and interpret the functional consequence of mutations that affect chromatin accessibility and gene expression (Atak et al., 2021).

Two families of enzymes regulate histone acetylation, histone acetyl transferases (HAT), and deacetylation, histone deacetylases (HDAC). HATs catalyse the transfer of an acetyl group from acetyl-CoA to the ε-amino group of lysine. This changes the charge of the histone reducing the affinity of the tail for DNA, meaning the histone is more easily displaced, giving molecules like transcription factors easier access to the DNA and increasing gene transcription (Zheng, 2015). Deacetylase proteins containing domains such as Bromodomains and extra terminal domains (BET) recognise acetylated histone-lysine residues and then act to remove these groups. This increases the basic charge of the histone allowing it to interact with the DNA more tightly leading to a decrease in gene transcription (Castelli et al., 2018). The aberrant expression of HDACs and HATs, as well as mutations in the *jat* and *hdac* genes, play an important role in the development and progression of cancer (Figueroa et al., 2013). The resulting changes in gene expression occur as a result of modified expression of these proteins. Therefore, both Hats and HDACs are targets for anti-cancer treatments (Figueroa et al., 2013). HDACa can also deacetylate and alter the function of proteins involved in cancer associated processes such as differentiation, autophagy apoptosis, DNA damage repair, and immune responses (Castelli et al., 2018).

Histone Methyl Transferases (HMTs), Histone Methylation Recognising Proteins and Histone Demethylases (HDMs) act as writers, readers, and erasers in the process of transferring a methyl group to histones. Methylation can occur on arginine and lysine residues. The HMTs transfer the methyl group. These residues can be modified with up to three methyl groups (Cheng et al., 2019). Methylation both promotes and inhibits gene expression and disruption of the methylation of histones can lead to the growth and development of cancer (Martinez-Garcia & Licht, 2010).

6.2.4 Chromosomal Structure

Chromatin remodelling allows the genetic material to be packaged and unpackaged when needed. It is also used to regulate access to DNA. It is therefore able to control access to and use of DNA regulatory elements that control chromosomal processes.

Table 6.1 List of AI tools used to interpret chromatin remodelling data

AI tool	Description	Reference
EpiTenso	High-order tensor decomposition-based algorithm to identify 3D spatial associations of chromatin	(Zhu et al., 2016)
TargetFinder	Machine learning pipeline that integrates data for annotation. Including cap analysis of gene expression (CAGE), ChIP-seq, DNase I hypersensitive sites sequencing (DNase-seq), FAIRE-seq (formaldehyde-assisted isolation of regulatory elements) and DNA methylation	(Whalen et al., 2016)
3DEpiLoop	Supervised learning pipeline using a random forest as a statistical learning algorithm to model 3D chromatin loops using one dimensional data	(Al Bkhetan & Plewczynski, 2018)
HiCPlus	Deep convolutional neural network that infers high-resolution chromosomal conformation interaction matrices using Hi-C data	(Zhang et al., 2018)
DeepTACT)	Bootstrapping deep learning model predict chromatin contacts at individual regulatory element level	(Li et al., 2019)

These elements include enhancers, promoters, and replication origins. In doing so, chromatin remodelling controls gene transcription, DNA replication, repair and recombination (Tsuda et al., 2021). There are four families of protein complexes that remodel chromatin. These are the switching defective/sucrose non-fermenting (SWI/SNF) family, the imitation switch (ISWI) family, chromodomain, the helicase, DNA binding (CHD) family, and the Inositol requiring 80 (INO80) family. All these complexes require energy in the form of ATP to remodel chromatin (Clapier & Cairns, 2009). The families differ in their function and can function to either change the structure of the nucleosome by evicting histone or alter the orientation of the subunit in the octamer. Other families can create DNA loops by altering the surface of nucleosomes or can even slide the nucleosome along the DNA strand. Any mutations in these chromatin re-modellers have been shown to be present in many types of cancers (Tsuda et al., 2021). The SWI/SNF family act by sliding nucleosomes, ejecting histones, or repositioning chromatin and are involved in DNA differentiation, proliferation, and repair (Clapier et al., 2017). The genes for this family of re-modellers has been found to carry mutations in more than 20% of human cancers. (Clapier & Cairns, 2009). The ISWI family of re-modellers are also deregulated in various cancers (Zhong et al., 2019). They catalyse nucleosome sliding and organise spacing and assembly of chromatin (Clapier & Cairns, 2009). Since the interpretation of chromatin remodelling often require that the data be used to infer the three-dimensional position of chromatin loops, AI has been used to turn one-dimensional data into three-dimensional models of chromatin. Many AI tools have been developed for this function and some of these tools are listed in Table 6.1.

6.3 AI in the Analysis of Epigenomics

Machine learning (ML) is the process whereby AI learns and makes predictions based on recognising patterns within the data. The more data the algorithm must analyse and learn from, the greater the accuracy of the predictions. With the current high throughput technologies generating large amounts of epigenomics data, these forms of AI are ideal to assist in the use of this data to identify biomarkers, therapeutic targets and assist in the management of cancer. These processes are already regularly performed using genomics and transcriptomics data, where AI is used to perform gene set enrichment analysis to identify upregulated signalling pathways (Ghanat Bari et al., 2017). Different types of ML are used in the analysis of epigenomic data. The two major different types of learning algorithms are supervised, and unsupervised learning (Fig. 6.2), both of which can make use of deep learning (DL).

Personalised medicine depends on the integrated analysis of big data mostly in the form of 'omics' data. In many cases, epigenetic data must be analysed alongside genetic, transcriptomic, and proteomic data to get a clear picture of the changes in gene expression. For instance, mutations in the genes that code for histone-modifying enzymes can be detected through genomic analysis, while the change in the protein function can be detected in the changes of the acetylation or methylation epigenetic markers. Finally, the changes in gene expression resulting from these genetic and epigenetic changes can be seen by examining the mRNAs that were transcribed and the proteins that were finally expressed (Lee et al., 2018). This may involve integrating data from various sources and interpreting different types of data. To do this AI must be able to undertake multimodal learning operations. In addition to this, the size and complexity of these databases demand that AI undertakes multitask learning (Fig. 6.2). This involves the performance of multiple learning tasks at the same time (Strezoski et al., 2019). This also allows multiple models to be trained at the same time and multiple models to be used to assess the same data increasing the accuracy of the final result. It also results in a faster more efficient analysis (Baxter, 2000). In order for these models to function simultaneously, they must share some underlying features, which also allows them to share data (Zhang & Yang, 2021).

6.3.1 Supervised Learning

In supervised learning, labelled datasets are used to train a dataset. This method requires a human to label the data or categorise the different inputs. Once the algorithm has used the labelled data to learn what the patterns are associated with, for example the diagnosis of an aggressive cancer, it can analyse different datasets in the same way to identify variable, such as the presence of an aggressive cancer (Rajkomar et al., 2019). Some of the different algorithms used include linear or

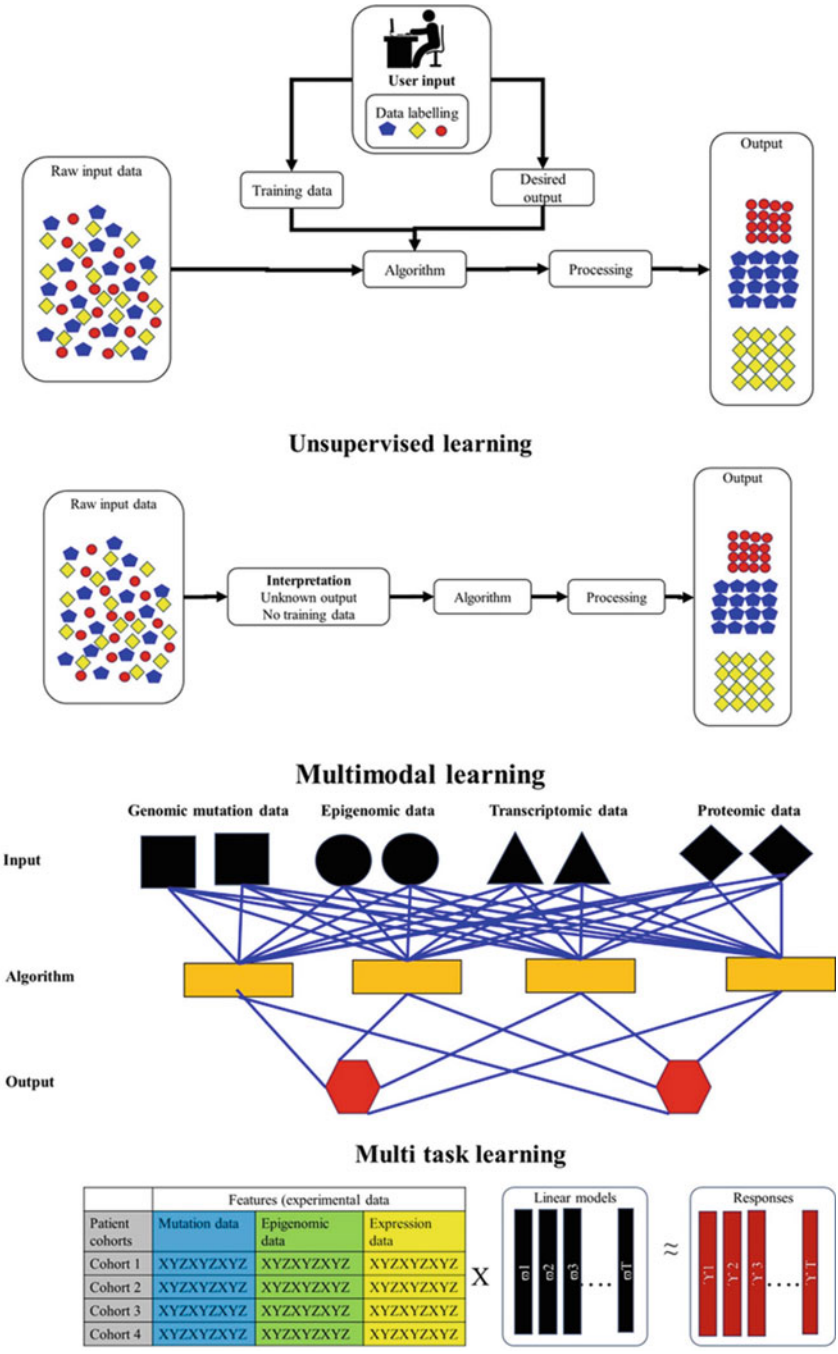


Fig. 6.2 Different learning models used by AI in undertaking machine learning to analyse omics data. Supervised learning is conducted using labelled data and requires user input to label the data. The algorithm then learns to classify data based on these provided labels. Unsupervised learning does not require labelled data and learns to tell data apart based on a classification system it creates

logistic regression, support vector machine, random forest algorithms, and least absolute shrinkage and selection operator regression (LASSO) (Krittanawong et al., 2017). Support vector machine separate classes of labelled data using a hyperplane, a simple line separating two-dimensional space, and the data points nearest to this line are used to classify other unlabelled data (Cristianini et al., 2008). Random Forest decision trees use mutule trees and select an average based on the outcome of the different trees (Cristianini et al., 2008). The LASSO algorithm is a logistic regression model that selects the most important prediction variables from the data using a penalisation model to weigh the feature for their importance (Cristianini et al., 2008).

Generally, supervised learning is used with epigenomic data to perform a predictive role in cancer. It has been used for example to classify metastatic prostate cancer. Here, prostate cancer tissue archives alongside healthy tissue counterparts were analysed for DNA methylation rates and patterns. This information was used for functional assessment, gene-set enrichment and protein interaction analyses, and examination of transcription factor-binding patterns. A LASSO algorithm was used with the data labelled, based on cancerous or healthy samples. Once the classification algorithm was trained, it was validated using many benign and tumour prostate arrays (Aref-Eshghi et al., 2018). Epigenomic analysis of DNA methylation proved to be highly accurate in identifying driver events for the development of prostate cancer. This was related to methylation events disrupting the activity of tumour suppressor genes. The methylation profiles were able to predict the presence of prostate cancer following the identification of as little as four CpGs (Aref-Eshghi et al., 2018). Supervised learning and AI were also used to diagnose brain cancers and asses their metastatic potential using DNA methylation patterns. The DNA methylation profiles of the three most common of brain metastases were determined, these included the spread of melanoma, breast, and lung cancers to form brain cancer. The AI analysed the methylation pattern from normal, primary, and metastatic cancers to create a three-step DNA methylation-based classifier (BrainMETH). This tool classifies brain cancers based on the tissue of origin using their DNA methylation patterns. These predictions matched those given by histopathology examination (Orozco et al., 2018).

6.3.2 Unsupervised Learning

Unsupervised learning uses unlabelled data which allows for a greater measure of correlation between two variables. Unsupervised learning is commonly used in

Fig. 6.2 (continued) based on differences in the data. Multimodal learning is used to combine data from various sources and learns to look for common features on the different data and links them in this way. Multitask learning is the use of AI to analyse multiple related data sets at the same time using different but related models

clustering, which involves assigning data points to various groups based on their inherent characteristics. Unsupervised learning relies on techniques such as *k*-means clustering and hierarchical clustering, principle component analysis, and partial least squares discriminant analysis (Tarca et al., 2007). Unsupervised learning is often used to simplify data. Both principle component analysis, and partial least squares discriminant analysis are used to reduce the dimensionality of data (Meng et al., 2016). Unsupervised algorithms are also more useful for detecting patterns in large datasets with lots of datapoints, which is typical of all omics databases (Nguyen & Rocke, 2002). Unsupervised learning has been used to detect DNA methylation patterns when comparing cancer and non-cancer tissue, or classify cancers based on their metastatic potential. An example of this is the use of unsupervised machine learning AI to perform whole-methylome analysis of primary adrenocortical tumours (pACT) in children. This study aimed to identify biomarkers that could be used for prognosis. The AI model identified two groups based on their DNA methylation profiles. The first group has higher methylation sites in CpG islands related to gene promoter regions. This group had a poorer prognosis with more advanced disease that was recurrent or metastatic (Bueno et al., 2022).

Semi-supervised learning is a technique where a small amount of labelled training data is used for an algorithm to identify features specific to each class, and then, by searching larger data sets, it can classify unlabelled data into one of these classifications based on similarities between the features of the labelled and unlabelled data (Oliver et al., 2018).

6.3.3 Deep Learning

Deep learning is a type of machine learning that constructs a complex hierarchy of analysis using multiple levels of analysis, each using the results of the previous analysis as input. It can use the outputs and inputs from multiple levels of this hierarchy to learn from. In other words, it can learn from features it has selected from higher and lower levels of a selected feature hierarchy that is built upon these lower hierarchy features. Its ability to automatically learn at multiple levels allows it to perform complex problem-solving and analysis (Bengio, 2009). Both supervised and unsupervised learning can make use of deep learning algorithms. Deep learning algorithms are useful in that they can work with very complex data. This includes large-volume, multidimensional data from a variety of sources. This makes them useful for integrating multi-omics data (LeCun et al., 2015). Supervised deep learning has been used to classify gliomas based on mutation and DNA methylation profiles in single cancer cells (Chang et al., 2018). Deep learning models can also be trained using unsupervised learning methods. Unlabelled mammography images were used to train a deep learning model that was trained to assign scores to breast density scores. These scores were shown to be an accurate predictor of breast cancer (Kallenberg et al., 2016).

6.4 The Practical Use of Epigenetic Data and AI in the Management of Cancer

Screening for genetic mutations gives useful information on an individual's predisposition to develop cancer. Epigenomic screening can provide more useful information considering the current status or activity of disease (Leygo et al., 2017). Screening for promoter DNA methylation is simpler since specific regions, promoter regions, are the only areas that need to be examined. By looking for epigenetic markers, involving methylation in promoter regions, numerous early diagnostic techniques for multiple cancer have been successfully pioneered. The genes under the control of these promoters are known to be involved in the initiation of carcinogenic pathways (Leygo et al., 2017). For instance, DNA from stool samples was screened for methylation status to diagnose colorectal cancer. This technique identified patients with cancer even when sequencing of DNA from the colonic mucosa did not indicate the presence of mutations that could predispose an individual to colorectal cancer (Elliott et al., 2013). Cancers of the central nervous system are difficult to diagnose and classify using histopathology, with a large degree of variability between different pathologists (Merve et al., 2019). DNA methylation profiles have been successfully used as biomarkers to classify various CNS tumours (Capper et al., 2018). This was optimised and built upon via the development of a comprehensive machine learning approach that used DNA methylation profiles to classify CNS tumours (Capper et al., 2018). These studies all demonstrate the use of epigenetic markers in the management of various cancers. However, the integrated analysis of epigenetics and other omics data combined with the large databases of epigenetic changes generated by high throughput techniques requires the use of AI and machine learning to fully exploit the potential of epigenomics.

A deep learning-based survival model for hepatocellular carcinoma was designed based on epigenetics data (DNA methylation) in combination with RNA sequencing data (RNA-seq) and microRNA-sequencing data (miRNA-seq). All these datasets were obtained from The Cancer Genome Atlas (TCGA). The model integrated all these datasets using multimodal learning and was able to predict prognosis as well as a model that used genomics and clinical data (Chaudhary et al., 2018). This model used an unsupervised deep learning method known as the autoencoder method which establishes clinical risk based on the level of variations (Chaudhary et al., 2018). By integrating epigenomic data with other omics data using AI, it is possible to clarify the significance of genetic mutations in non-coding regions. Genome-wide association studies have identified single-nucleotide polymorphisms in non-coding DNA regulatory elements (Corces et al., 2018). To do this, an unsupervised learning method was used to construct the genome-wide chromatin accessibility map. This was done using whole genome sequencing and ATAC-sequencing data from 410 tumour samples representing 23 cancer types obtained from the TCGA. The assay for transposase-accessible chromatin using sequencing (ATAC-seq), is an NGS-based technique that constructs the library for sequencing using hyperactive transposase. The adapters for NGS are loaded onto transposase which simultaneous

fragmentation and attaches the adapter to open chromatin regions. This results in a library that is made up of accessible regions of the chromatin (Cui et al., 2021). The unsupervised k-means clustering AI algorithm assisted analysis of this data. It identified 562,709 DNA elements that are transposase-accessible and identified distal enhancers. This data can be further analysed to subtype cancers, generate protein-DNA footprints that can be used to identify transcription factors that drive cancer-related transcription and can be used to, identify gene-regulatory interactions in cancer. This analysis also showed that genetic loci that can predispose individuals to develop cancer are epigenetically active DNA regulatory elements (Corces et al., 2018).

6.5 Limitations of AI-Driven Epigenomics Applications

All three ML methods have their own drawbacks and limitations. Supervised learning requires user input to label the teaching data and so cannot be fully automated. Also, the data must be correctly labelled and therefore the system is sensitive to human error (Krittanawong et al., 2017). Supervised learning may also suffer from the problem known as 'over-fitting'. This is the name given to the phenomenon where the algorithm is able to analyse the training data very well but this does not translate to other data sets as the algorithm is too optimised to the teaching data (Japkowicz & Stephen, 2002). Deep learning requires large amounts of computing power (LeCun et al., 2015). This may be a prohibitive factor in regular diagnostic applications. However, the continuous increase in computing power, along with cloud computing, should offer a solution to this problem. One of the biggest issues with DL is due to the multi-layered complex analysis, it is not always possible for users to discern how the AI arrived at the final solution and this may create trust issues. This black box problem is a common problem with all uses of AI (Rudin, 2019). One of the greatest limitations of unsupervised learning is that it is unable to label the patterns of correlation with a potential biological relevance, meaning that user input is required, and the process cannot be fully automated (Alanazi et al., 2017). This can best be summed up as correlation not implying causation, as the algorithms inability to give the relevance of clustering and/or associations means these methods struggle to assign meaning to the patterns they recognise. This means that care must be taken when interpreting the analysis performed by an AI running an unsupervised learning algorithm (Nguyen & Rocke, 2002). These methods can also be negatively affected by unnecessary, redundant, or irrelevant (noisy) data present in the dataset. This data may lead the algorithm to group the data incorrectly. This can be avoided by the careful pre-processing of data (Nguyen & Rocke, 2002).

A difficulty for all these methods is the high number of variables in epigenetic data, which sometimes exceeds the sample number. Many of the algorithms used struggle with this problem which can only be solved by gathering more samples (Kirpich et al., 2018). It is also important to choose the right algorithm for a specific

task or dataset. Some of the associations with DNA methylation patterns are not linear, with multiple CpGs linked to the same gene able to influence other methylation events and therefore the transcriptome (Deutsch & McIlvane, 2012). Prediction bias is another problem faced by AI in all omics analysis tasks. This is most observed in the bias most models have for the molecular patterns associated with Europeans. This is because most studies are conducted with European samples which means most training data is from Europeans (Phillips et al., 2011). This particular problem can only be solved through the study and collection of data from more non-European populations, especially African populations and populations from low to middle income countries (LMICs) (Char et al., 2018).

6.6 Conclusions

Epigenomic data can be used on its own or integrated with other omics data to provide a clear view of the molecular landscape of signalling pathways involved in the development and progression of cancer. Epigenomic markers offer an advantage in that they are dynamic and change with the patient's lifestyle and environment and are linked to risk factors. Unlike transcriptomic or proteomic analysis, they will provide an earlier indication of the presence of or level of gene expression changes that may lead to cancer. Like all omics analysis, epigenomic analysis requires the analysis of big data sets. This is especially true of integrated analysis across multi-omics data. Practically the best way to accomplish this by using AI. This is especially true as with high throughput technologies like Nest Generation Sequencing. Omics data has entered the era of big data and epigenomics is no exception (Fig. 6.3). The use of AI to analyse epigenomic data carries many of the same problems as all AI does in the field of medicine. These problems are more closely related to omics data involving shortcomings within the actual model used and the data used to teach these models. The data problem should be solved over time as more data is archived and a greater variety of populations and individuals are studied than the traditionally focussed upon European populations. Many of the problems concerning the different learning models can be solved by the correct selection of the appropriate model to perform a task (Fig. 6.3). Also, as more advanced models are created through the efforts of programmers and through trial and error, these problems and shortcomings may become things of the past.

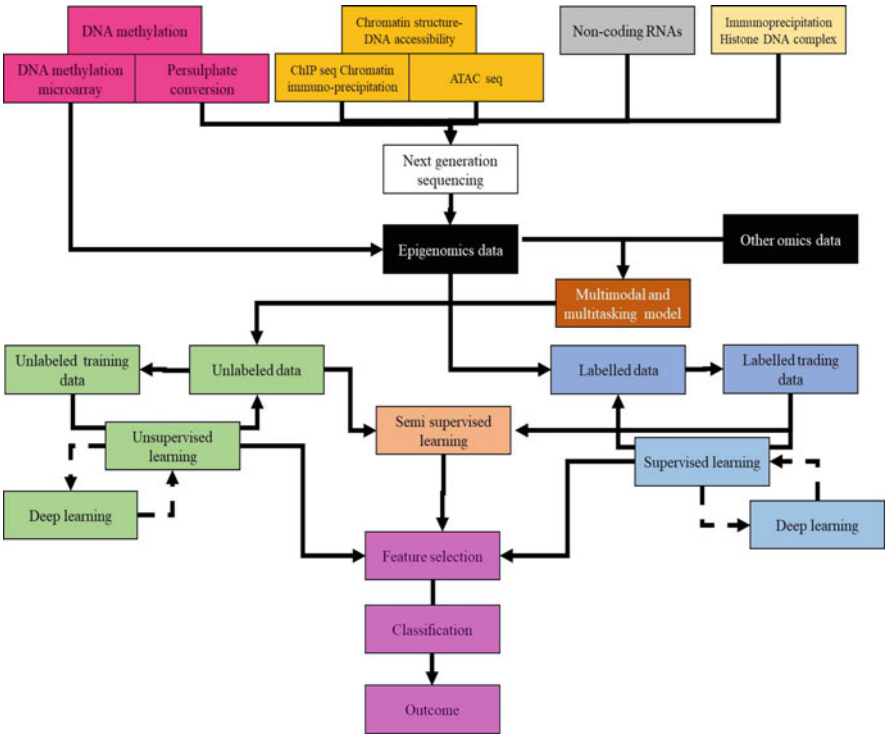


Fig. 6.3 The use of AI to analyse epigenomic data from different sources. Epigenomic data concerning changes in DNA methylation, chromatin structure, histone modification and the expression of non-coding RNAs has changed with the development of Next Generation Sequencing. This has allowed for the acquisition of vast amounts of data and has brought epigenomics into the big data era. To analyse and interpret this large amount of data, which may also require the integration of other data sources, different machine learning models must be implemented by AI algorithms to identify features within the data that can be used to classify the samples for cancer management

References

Al Bkhetan, Z., & Plewczynski, D. (2018). Three-dimensional epigenome statistical model: Genome-wide chromatin looping prediction. *Scientific Reports, 8*, 5217.

Alanazi, H. O., Abdullah, A. H., & Qureshi, K. N. (2017). A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care. *Journal of Medical Systems, 41*, 69.

Aref-Eshghi, E., Schenkel, L. C., Ainsworth, P., Lin, H., Rodenhiser, D. I., Cutz, J. C., & Sadikovic, B. (2018). Genomic Dna methylation-derived algorithm enables accurate detection of malignant prostate tissues. *Frontiers in Oncology, 8*, 100.

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium Dna methylation microarrays. *Bioinformatics, 30*, 1363–1369.

Atak, Z. K., Taskiran, I., Demeulemeester, J., Flerin, C., Mauduit, D., Minnoye, L., Hulselmans, G., Christiaens, V., Ghanem, G. E., Wouters, J., & Aerts, S. (2021). Interpretation of allele-specific chromatin accessibility using cell state-aware deep learning. *Genome Research, 31*, 1082–1096.

Baxter, J. (2000). A model of inductive bias learning. *Journal of Artificial Intelligence Research, 12*, 149–198.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations Trends in Machine Learning, 2*, 1–127.

Bhat, S. A., Ahmad, S. M., Mumtaz, P. T., Malik, A. A., Dar, M. A., Urwat, U., Shah, R. A., & Ganai, N. A. (2016). Long non-coding Rnas: Mechanism of action and functional utility. *Noncoding RNA Research, 1*, 43–50.

Bird, A. (2002). Dna methylation patterns and epigenetic memory. *Genes & Development, 16*, 6–21.

Bowman, G. D., & Poirier, M. G. (2015). Post-translational modifications of histones that influence nucleosome dynamics. *Chemical Reviews, 115*, 2274–2295.

Bueno, A. C., Da Silva, R. M. P., Stecchini, M. F., Marrero-Gutiérrez, J., De Almeida, E. S. D. C., Cardinalli, I., Scrideli, C. A., Junqueira, T., Molina, C. A. F., Ramalho, F. S., Tucci Junior, S., Coeli-Lacchini, F. B., Moreira, A. C., Ramalho, L. N. Z., Brandalise, S., Yunes, J. A., De Castro, M., Vêncio, R. Z. N., & Antonini, S. R. R. (2022). Dna methylation is a comprehensive marker for pediatric adrenocortical tumors. *Endocrine-Related Cancer, 29*, 599.

Capper, D., Jones, D. T. W., Sill, M., Hovestadt, V., Schrimpf, D., Sturm, D., Koelsche, C., Sahm, F., Chavez, L., Reuss, D. E., Kratz, A., Wefers, A. K., Huang, K., Pajtler, K. W., Schweizer, L., Stichel, D., Olar, A., Engel, N. W., Lindenberg, K., Harter, P. N., Braczynski, A. K., Plate, K. H., Dohmen, H., Garvalov, B. K., Coras, R., Hölsken, A., Hewer, E., Bewerunge-Hudler, M., Schick, M., Fischer, R., Beschorner, R., Schittenhelm, J., Staszewski, O., Wani, K., Varlet, P., Pages, M., Temming, P., Lohmann, D., Selt, F., Witt, H., Milde, T., Witt, O., Aronica, E., Giangaspero, F., Rushing, E., Scheurlen, W., Geisenberger, C., Rodriguez, F. J., Becker, A., Preusser, M., Haberler, C., Bjerkvig, R., Cryan, J., Farrell, M., Deckert, M., Hench, J., Frank, S., Serrano, J., Kannan, K., Tsirigos, A., Brück, W., Hofer, S., Brehmer, S., Seiz-Rosenhagen, M., Hänggi, D., Hans, V., Rozsnoki, S., Hansford, J. R., Kohlhof, P., Kristensen, B. W., Lechner, M., Lopes, B., Mawrin, C., Ketter, R., Kulozik, A., Khatib, Z., Heppner, F., Koch, A., Jouvet, A., Keohane, C., Mühleisen, H., Mueller, W., Pohl, U., Prinz, M., Benner, A., Zapatka, M., Gottardo, N. G., Driever, P. H., Kramm, C. M., Müller, H. L., Rutkowski, S., Von Hoff, K., Frühwald, M. C., Gnekow, A., Fleischhack, G., Tippelt, S., Calaminus, G., Monoranu, C. M., Perry, A., Jones, C., et al. (2018). Dna methylation-based classification of central nervous system tumours. *Nature, 555*, 469–474.

Castelli, G., Pelosi, E., & Testa, U. (2018). Targeting histone methyltransferase and demethylase in acute myeloid leukemia therapy. *Oncotargets and Therapy, 11*, 131–155.

Chang, P., Grinband, J., Weinberg, B. D., Bardis, M., Khy, M., Cadena, G., Su, M. Y., Cha, S., Filippi, C. G., Bota, D., Baldi, P., Poisson, L. M., Jain, R., & Chow, D. (2018). Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. *AJNR. American Journal of Neuroradiology, 39*, 1201–1207.

Chang, S. C., Tucker, T., Thorogood, N. P., & Brown, C. J. (2006). Mechanisms of X-chromosome inactivation. *Frontiers in Bioscience, 11*, 852–866.

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care—addressing ethical challenges. *The New England Journal of Medicine, 378*, 981–983.

Chaudhary, K., Poirion, O. B., Lu, L., & Garmire, L. X. (2018). Deep learning-based multi-omics integration robustly predicts survival in liver cancer. *Clinical Cancer Research, 24*, 1248–1259.

Cheng, Y., He, C., Wang, M., Ma, X., Mo, F., Yang, S., Han, J., & Wei, X. (2019). Targeting epigenetic regulators for cancer therapy: Mechanisms and advances in clinical trials. *Signal Transduction and Targeted Therapy, 4*, 62.

Clapier, C. R., & Cairns, B. R. (2009). The biology of chromatin remodeling complexes. *Annual Review of Biochemistry, 78*, 273–304.

Clapier, C. R., Iwasa, J., Cairns, B. R., & Peterson, C. L. (2017). Mechanisms of action and regulation of Atp-dependent chromatin-remodelling complexes. *Nature Reviews. Molecular Cell Biology, 18*, 407–422.

Comfort, N. J. N. (2015). Genetics: We are the 98%. *Nature, 520*, 615–616.

Corces, M. R., Granja, J. M., Shams, S., Louie, B. H., Seoane, J. A., Zhou, W., Silva, T. C., Groeneveld, C., Wong, C. K., Cho, S. W., Satpathy, A. T., Mumbach, M. R., Hoadley, K. A., Robertson, A. G., Sheffield, N. C., Felau, I., Castro, M. A. A., Berman, B. P., Staudt, L. M., Zenklusen, J. C., Laird, P. W., Curtis, C., Greenleaf, W. J., & Chang, H. Y. (2018). The chromatin accessibility landscape of primary human cancers. *Science, 362*.

Cristianini, N., Ricci, E., & Kao, M. (2008). *Encyclopedia of algorithms*. Springer.

Csepany, T., Lin, A., Baldick, C. J., Jr., & Beemon, K. (1990). Sequence specificity of Mrna N6-adenosine methyltransferase. *The Journal of Biological Chemistry, 265*, 20117–20122.

Cui, Y., Li, J. S., & Li, W. (2021). From reads to insights: Integrative pipelines for biological interpretation of Atac-Seq data. *Genomics, Proteomics & Bioinformatics, 19*, 519–521.

Deutsch, C. K., & Mcilvane, W. J. (2012). Non-Mendelian etiologic factors in neuropsychiatric illness: Pleiotropy, epigenetics, and convergence. *The Behavioral and Brain Sciences, 35*, 363–364.

Elliott, G. O., Johnson, I. T., Scarll, J., Dainty, J., Williams, E. A., Garg, D., Coupe, A., Bradburn, D. M., Mathers, J. C., & Belshaw, N. J. (2013). Quantitative profiling of Cpg Island methylation in human stool for colorectal cancer detection. *International Journal of Colorectal Disease, 28*, 35–42.

Esteller, M. (2011). Non-coding Rnas in human disease. *Nature Reviews. Genetics, 12*, 861–874.

Figueroa, M. E., Chen, S. C., Andersson, A. K., Phillips, L. A., Li, Y., Sotzen, J., Kundu, M., Downing, J. R., Melnick, A., & Mullighan, C. G. (2013). Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia. *The Journal of Clinical Investigation, 123*, 3099–3111.

Filipowicz, W., Bhattacharyya, S. N., & Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by Micrornas: Are the answers in sight? *Nature Reviews Genetics, 9*, 102–114.

García-Giménez, J. L., Sanchis-Gomar, F., Lippi, G., Mena, S., Ivars, D., Gomez-Cabrera, M. C., Viña, J., & Pallardó, F. V. (2012). Epigenetic biomarkers: A new perspective in laboratory diagnostics. *Clinica Chimica Acta, 413*, 1576–1582.

Ghanat Bari, M., Ung, C. Y., Zhang, C., Zhu, S., & Li, H. (2017). Machine learning-assisted network inference approach to identify a new class of genes that coordinate the functionality of cancer networks. *Scientific Reports, 7*, 6993.

Heyn, H., & Esteller, M. (2012). Dna methylation profiling in the clinic: Applications and challenges. *Nature Reviews. Genetics, 13*, 679–692.

How Kit, A., Nielsen, H. M., & Tost, J. (2012). Dna methylation based biomarkers: Practical considerations and applications. *Biochimie, 94*, 2314–2337.

Jaffe, A. E., Murakami, P., Lee, H., Leek, J. T., Fallin, M. D., Feinberg, A. P., & Irizarry, R. A. (2012). Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. *International Journal of Epidemiology, 41*, 200–209.

Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis, 6*, 429–449.

Jones, P. A., Issa, J. P., & Baylin, S. (2016). Targeting the cancer epigenome for therapy. *Nature Reviews Genetics, 17*, 630–641.

Kallenberg, M., Petersen, K., Nielsen, M., Ng, A. Y., Pengfei, D., Igel, C., Vachon, C. M., Holland, K., Winkel, R. R., Karssemeijer, N., & Lillholm, M. (2016). Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. *IEEE Transactions on Medical Imaging, 35*, 1322–1331.

Kirpich, A., Ainsworth, E. A., Wedow, J. M., Newman, J. R. B., Michailidis, G., & Mcintyre, L. M. (2018). Variable selection in omics data: A practical evaluation of small sample sizes. *PLoS One, 13*, E0197910.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell, 128*, 693–705.

Krittanawong, C., Zhang, H., Wang, Z., Aydar, M., & Kitai, T. (2017). Artificial intelligence in precision cardiovascular medicine. *Journal of the American College of Cardiology, 69*, 2657–2664.

Kurdyukov, S., & Bullock, M. (2016). Dna methylation analysis: Choosing the right method. *Biology (Basel), 5*.

Laget, S., & Defossez, P. A. (2008). Master and servant: Epigenetic deregulations as a cause and a consequence of cancer. *Medical Science (Paris), 24*, 725–730.

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436–444.

Lee, S. I., Celik, S., Logsdon, B. A., Lundberg, S. M., Martins, T. J., Oehler, V. G., Estey, E. H., Miller, C. P., Chien, S., Dai, J., Saxena, A., Blau, C. A., & Becker, P. S. (2018). A machine learning approach to integrate big data for precision medicine in acute myeloid leukemia. *Nature Communications, 9*, 42.

Leung, M. K., Delong, A., Alipanahi, B., & Frey, B. (2015). Machine learning in genomic medicine: A review of computational problems and data sets. *Proceedings of the IEEE, 104*, 176–197.

Leygo, C., Williams, M., Jin, H. C., Chan, M. W. Y., Chu, W. K., Grusch, M., & Cheng, Y. Y. (2017). Dna methylation as a noninvasive epigenetic biomarker for the detection of cancer. *Disease Markers, 2017*, 3726595.

Li, W., Wong, W. H., & Jiang, R. (2019). Deeptact: Predicting 3d chromatin contacts via bootstrapping deep learning. *Nucleic Acids Research, 47*, E60.

Martinez-Garcia, E., & Licht, J. D. (2010). Deregulation of H3k27 methylation in cancer. *Nature Genetics, 42*, 100–101.

Meng, C., Zeleznik, O. A., Thallinger, G. G., Kuster, B., Gholami, A. M., & Culhane, A. C. (2016). Dimension reduction techniques for the integrative analysis of multi-omics data. *Briefings in Bioinformatics, 17*, 628–641.

Merve, A., Millner, T. O., & Marino, S. (2019). Integrated phenotype-genotype approach in diagnosis and classification of common central nervous system tumours. *Histopathology, 75*, 299–311.

Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., & Jaffrey, S. R. (2012). Comprehensive analysis of Mrna methylation reveals enrichment in 3' Utrs and near stop codons. *Cell, 149*, 1635–1646.

Nguyen, D. V., & Rocke, D. M. (2002). Tumor classification by partial least squares using microarray gene expression data. *Bioinformatics, 18*, 39–50.

Oliver, A., Odena, A., Raffel, C. A., Cubuk, E. D., & Goodfellow, I. (2018). Realistic evaluation of deep semi-supervised learning algorithms. *Advances in Neural Information Processing Systems, 31*.

Orozco, J. I. J., Knijnenburg, T. A., Manughian-Peter, A. O., Salomon, M. P., Barkhoudarian, G., Jalas, J. R., Wilmott, J. S., Hothi, P., Wang, X., Takasumi, Y., Buckland, M. E., Thompson, J. F., Long, G. V., Cobbs, C. S., Shmulevich, I., Kelly, D. F., Scolyer, R. A., Hoon, D. S. B., & Marzese, D. M. (2018). Epigenetic profiling for the molecular classification of metastatic brain tumors. *Nature Communications, 9*, 4627.

Paulsen, M., & Ferguson-Smith, A. C. (2001). Dna methylation in genomic imprinting, development, and disease. *The Journal of Pathology, 195*, 97–110.

Payne, S. R. (2010). From discovery to the clinic: The novel Dna methylation biomarker (M)Sept9 for the detection of colorectal cancer in blood. *Epigenomics, 2*, 575–585.

Phillips, P. J., Jiang, F., Narvekar, A., Ayyad, J., & O'toole, A. (2011). An other-race effect for face recognition algorithms. *ACM Transactions on Applied Perception, 8*, 1–11.

Piletič, K., & Kunej, T. (2016). Microrna epigenetic signatures in human disease. *Archives of Toxicology, 90*, 2405–2419.

Place, R. F., Li, L. C., Pookot, D., Noonan, E. J., & Dahiya, R. (2008). Microrna-373 induces expression of genes with complementary promoter sequences. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 1608–1613.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *The New England Journal of Medicine, 380*, 1347–1358.

Rauschert, S., Raubenheimer, K., Melton, P. E., & Huang, R. C. (2020). Machine learning and clinical epigenetics: A review of challenges for diagnosis and classification. *Clinical Epigenetics, 12*, 51.

Reyngold, M., & Chan, T. (2018). *Molecular oncology: Causes of cancer and targets for treatment*. Cambridge University Press.

Robertson, K. D. (2005). Dna methylation and human disease. *Nature Reviews. Genetics, 6*, 597–610.

Romanowska, J., & Joshi, A. (2019). From genotype to phenotype: Through chromatin. *Genes (Basel), 10*.

Romero-Barrios, N., Legascue, M. F., Benhamed, M., Ariel, F., & Crespi, M. (2018). Splicing regulation by long noncoding Rnas. *Nucleic Acids Research, 46*, 2169–2184.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*, 206–215.

Sina, A. A., Carrascosa, L. G., Liang, Z., Grewal, Y. S., Wardiana, A., Shiddiky, M. J. A., Gardiner, R. A., Samaratunga, H., Gandhi, M. K., Scott, R. J., Korbie, D., & Trau, M. (2018). Epigenetically reprogrammed methylation landscape drives the Dna self-assembly and serves as a universal cancer biomarker. *Nature Communications, 9*, 4915.

Singer, B. D. (2019). A practical guide to the measurement and analysis of Dna methylation. *American Journal of Respiratory Cell and Molecular Biology, 61*, 417–428.

Strezoski, G., Van Noord, N., & Worring, M. (2019). Learning task relatedness in multi-task learning for images in context. Proceedings of the 2019 on international conference on multimedia retrieval. (pp. 78–86).

Tarca, A. L., Carey, V. J., Chen, X. W., Romero, R., & Drăghici, S. (2007). Machine learning and its applications to biology. *PLoS Computational Biology, 3*, E116.

Toh, T. S., Dondelinger, F., & Wang, D. (2019). Looking beyond the hype: Applied Ai and machine learning in translational medicine. *eBioMedicine, 47*, 607–615.

Tsuda, M., Fukuda, A., Kawai, M., Araki, O., & Seno, H. (2021). The role of the Swi/Snf chromatin remodeling complex in pancreatic ductal adenocarcinoma. *Cancer Science, 112*, 490–497.

Wang, L. X., Wan, C., Dong, Z. B., Wang, B. H., Liu, H. Y., & Li, Y. (2019). Integrative analysis of long noncoding Rna (Lncrna), Microrna (Mirna) and Mrna expression and construction of a competing endogenous Rna (Cerna) network in metastatic melanoma. *Medical Science Monitor, 25*, 2896–2907.

Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., Gong, Z., Wang, Q., Huang, J., Tang, C., Zou, T., & Yin, P. (2016). Structural basis of N(6)-adenosine methylation by the Mettl3-Mettl14 complex. *Nature, 534*, 575–578.

Wen, W. S., Hu, S. L., Ai, Z., Mou, L., Lu, J. M., & Li, S. (2017). Methylated of genes behaving as potential biomarkers in evaluating malignant degree of glioblastoma. *Journal of Cellular Physiology, 232*, 3622–3630.

Whalen, S., Truty, R. M., & Pollard, K. S. (2016). Enhancer-promoter interactions are encoded by complex genomic signatures on looping chromatin. *Nature Genetics, 48*, 488–496.

Zhang, Y., An, L., Xu, J., Zhang, B., Zheng, W. J., Hu, M., Tang, J., & Yue, F. (2018). Enhancing hi-C data resolution with deep convolutional neural network Hicplus. *Nature Communications, 9*, 750.

Zhang, Y., & Yang, Q. (2021). A survey on multi-task learning. *IEEE Transactions on Knowledge Data Engineering*.

Zheng, Y.-P. G. (2015). *Epigenetic technological applications*. Academic Press.

Zhong, L., Liao, D., Zhang, M., Zeng, C., Li, X., Zhang, R., Ma, H., & Kang, T. (2019). Ythdf2 suppresses cell proliferation and growth via destabilizing the Egfr Mrna in hepatocellular carcinoma. *Cancer Letters, 442*, 252–261.

Zhu, J. Y., Lyu, A. Q., Wang, Z. T., Chan, W. Y., Qin, T., Miu, K. K., & Yao, H. R. (2022). Machine learning-devised immune-related Lncrna signature panel predicts the prognosis and immune landscape in breast cancer novel Irlp signature in Brca. *Journal of Immunology Research, 2022*, 3704798.

Zhu, Y., Chen, Z., Zhang, K., Wang, M., Medovoy, D., Whitaker, J. W., Ding, B., Li, N., Zheng, L., & Wang, W. (2016). Constructing 3d interaction maps from 1d epigenomes. *Nature Communications, 7*, 10812.

Chapter 7
Association of Metabolomics with AI in Precision Oncology: Emerging Perspectives for More Effective Cancer Care

Langanani Mbodi, Precious Mathebela, and Zodwa Dlamini

Abstract Metabolomics is defined as the comprehensive analysis of metabolites in a biological specimen. It holds a long-awaited promise to inform the practice of precision medicine. Despite its enormous potential that has already been explored widely, metabolomics has been relatively underutilized in all other sectors and so much so in oncology. Metabolites have been used in the recent past to diagnose complex metabolic diseases including disorders such as inborn errors of metabolism. Classifying tumours is essential for determining treatment and prognosis. With next-generation sequencing profiling, the classification of cancers can be conducted using circulating tumour DNA and analysing copy number variation associated with cancers such as small lung cancer as example, with no need for expensive invasive histological classification. The collective efforts of radiomics and deep learning will in the future deliver increased accuracy in diagnostic image analysis. The combined applications of artificial intelligence and machine learning in healthcare will in the future be implemented to improve disease management and provide effective medical care.

L. Mbodi (✉)
Department of Obstetrics and Gynaecology, Gynaecologic Oncology Unit, Charlotte Maxeke Johannesburg Academic Hospital, University of the Witwatersrand, Johannesburg, South Africa
e-mail: langanani.mbodi@wits.ac.za

P. Mathebela
Department of Surgery, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: mathebela.precious@up.ac.za

Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: zodwa.dlamini@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_7

This chapter is intended to assist or further demonstrate the valuable role of metabolomics and AI in precision oncology including specific areas such as radiogenomics for diagnosis, treatment and prognostication of the disease.

Keywords Metabolomics · Precision oncology · Metabolomics in cancer · Artificial intelligence · Cancer care

7.1 Definitions and Broad Applications

7.1.1 Metabolomics

Metabolomics is defined as the comprehensive analysis of metabolites in a biological specimen. As an emerging technology, it holds a long-awaited promise to inform the practice of precision medicine. Metabolites have been used to diagnose complex metabolic diseases including disorders such as inborn errors of metabolism (Clish, 2015; Dettmer et al., 2007). The metabolites are mostly regulated by genetic factors which hold potential as disease treatment targets (Chu et al., 2021).

Metabolomics technologies are capable of precise analysis of up to thousands of metabolites without limiting to standard clinical chemistry techniques. Metabolomics offers a detailed characterization of metabolic phenotypes and allows for easy and practical use in clinical practice, precision medicine approaches such as in disease diagnosis and management including characterization of metabolic derangements that underlie disease, the discovery of novel therapeutic targets and biomarkers that may be used to either diagnose a disease or monitor effect and response to therapeutic agents (Clish, 2015).

In biological specimens, it affords comprehensive detection and quantification of metabolites and small molecules in biological specimens by combining analytical techniques and chemometrics to enable researchers to identify a large proportion of metabolites (the metabolome) present in a sample (Clish, 2015). The metabolites can be in the form of amino acids, sugars, ketones, nucleotides, fatty acids, organic acids, as well as microbial-derived metabolites, and exogenous small molecules (including drugs, food additives, toxins and pesticides) (Chung et al., 2018). Through analysis of extracted data, valuable data is extracted through analysing these products of cellular metabolism and relevant information about an organism's metabolic or physiologic state at the time of sampling is revealed (Chung et al., 2018; Kim et al., 2011). There are other omics which are considered to complement metabolomics and these include genomics, transcriptomics, microbiomics, epigenomics and proteomics. Omics have been used in a wide range of applications such as in environmental analysis, toxicology, nutrition science and systems biology (Robertson et al., 2011; Weckwerth, 2010; Wilmes et al., 2013; Zhao et al., 2017). In agriculture, metabolomics has been used in conjunction with nutrition assessment methods to identify biomarkers that represent diet-related disease risks. In this field, metabolomics technologies have been used to improve commercially significant traits and increase crop yields. However, despite its enormous potential,

metabolomics has been relatively underutilized in veterinary medicine where it could be of great assistance for screening, diagnosing and treatment of spontaneous diseases (Tran et al., 2020).

The most important role of metabolomics in the biomedical field in the current era is to identify and develop novel disease biomarkers while providing new insights into disease pathogenesis and explain complex endogenous and exogenous biochemical pathways to help in disease diagnosis, monitoring of cellular responses to nutrition, drugs, toxins, etc. There are further processes which aid in drug development and cancer diagnosis which are based on the knowledge of cancer being an altered metabolism disease and therefore, metabolomics is relevant to cancer biology research (carcinogenesis, identification of specific biomarkers for diagnosis) (Tran et al., 2020).

7.1.2 Analytical Techniques in Metabolomics

Metabolomic analyses commonly utilize one or more analytical techniques to facilitate the identification and quantification of as many metabolites as possible in a biological sample (Alonso et al., 2015; Sas et al., 2015). The most commonly used are nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS). One-dimensional NMR spectroscopy is commonly used to detect 10 s–100 s of metabolites in biological extracts. One of the advantages of one-dimensional NMR is that it allows the reuse of the sample for other analyses and can quantify metabolites with high accuracy and reproducibility (Ren et al., 2018). Two-dimensional NMR techniques can be used to confirm or elucidate the structures of previously unidentified metabolites, as well as measure the incorporation of stable isotopes in labelling experiments. MS techniques involve the ionization of derivatized or underivatized samples and detection of corresponding charged ions (as mass-to-charge ratios) (Marion, 2013). MS allows greater coverage of metabolites and are more sensitive than the NMR. MS is often coupled with chromatographic separation techniques such as gas or liquid chromatography (GC, LC) or capillary electrophoresis (CE). Chromatographic separation of samples minimizes ion suppression effects associated with complex mixtures thereby increasing sensitivity. This allows great sampling load and provides orthogonal information (retention time prediction) that allows metabolite identification. In recent years there have been advances towards the use of high-resolution, accurate MS instruments such as the orbitrap MS and Fourier-transform ion cyclotron resonance MS (FT-ICR-MS) (Alonso et al., 2015; Sas et al., 2015).

New-generation MS instruments that allow post-chromatographic separation of analytes by ion mobility instruments have the capacity to increase metabolite coverage. This is achieved by allowing the separation of metabolite isomers with the same mass and providing information on the shape (collision cross-section) of molecules that can be used in metabolite identification. The use of multiple

complementary analytical platforms is important given that no single platform offers complete coverage of either polar or apolar metabolites (Tran et al., 2020).

As with proteomics, the profiling of metabolites relies on MS but with the additional use of NMR spectroscopy. Separation-free MS techniques have now been developed which reduce the volume of sample required and variation in the data generated through the analysis without separation or fractionation. The use of hubs such as the Global Natural Product Social Molecular Networking (GNPS) which is a small molecule MS networking hub where researchers can deposit their own MS data for small molecules makes it possible for the data to be available for other users to search and use. Such data can be mined through the use of techniques such as principal component analysis or hierarchical clustering in conjunction with MS. These repositories enhance identification of spectra. These techniques have been used to identify metabolic biomarkers for cancers in different organ systems such as colorectal, pancreatic, lung, breast, gastric, ovarian and prostate (Puchades-Carrasco & Pineda-Lucena, 2017). Figure 7.1 shows the basic workflow in metabolomic studies.

7.1.3 Limitations of Metabolomics

A number of limitations hinder the widespread use of metabolomic techniques regardless of its evolution since its inception. In current practice, most metabolomic studies only identify a minority of metabolites in biological samples, reflecting the complexity of sample analysis, the presence of multiple adducts and isotopes for each species, and the difficulty of validating hundreds of metabolites with suitable standards (Edelstein, 2016). There are some classes of metabolites that are either difficult to detect using current instrumentation or technology or are present below the level at which these techniques are able to detect and this complicates the interpretation of data. When new or unknown metabolites are detected, they present a new opportunity for understanding disease processes and detecting novel disease biomarkers. However, the identification of these unknown or poorly defined metabolites remains one of the biggest challenges for metabolomics (Wang et al., 2010).

When metabolites are identified, their significance with regard to cancer diagnosis as biomarkers, disease prognostic markers, etc. is another challenge. These non-targeted metabolomic analyses generate enormous data sets that may contain a lot of clinically irrelevant or unintended information which may demand a lot of resources to transform these data into valid interpretations and conclusions. It also requires an in-depth understanding of metabolic pathways and the interconnectivity of metabolites and biological systems (Johnson et al., 2016). Hence, in many cases it is difficult to conclude how a change in metabolite steady-state levels translates to changes in metabolic fluxes through one or more associated pathways although this

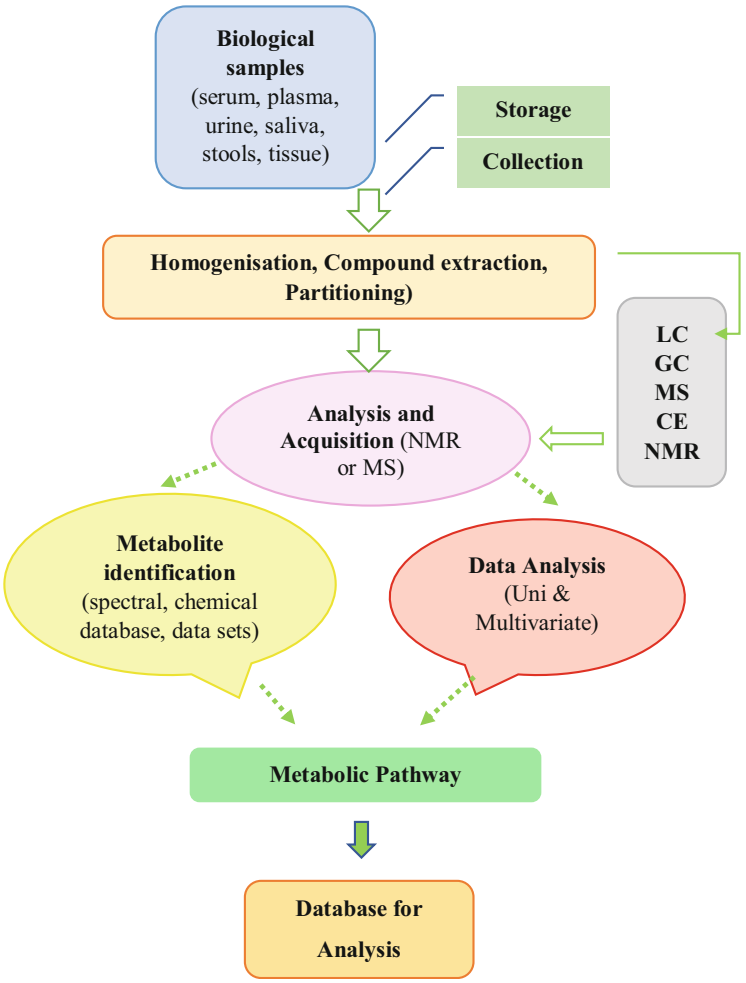


Fig. 7.1 Flow diagram of basic steps to follow when conducting metabolomic studies. *LC* liquid chromatography, *MS* mass spectrometry, *CE* capillary electrophoresis, *GC* gas chromatography, *NMR* nuclear magnetic resonance

is increasingly being addressed by coupling metabolomic approaches with stable isotope labelling. There are metabolites which are only biologically significant in the presence of other metabolites. Another challenge is that pertaining to the optimization of sample collection and storage protocols. This needs to be addressed in order to gain a global understanding of physiologic processes and optimally integrate other omics such as genomics, proteomics, transcriptomics, etc. (Tran et al., 2020).

7.2 Precision Oncology

Precision oncology, also termed personalized medicine in cancer medicine, is an innovative approach to tailoring disease prevention and treatment that considers differences in individuals' genes, environment and lifestyle factors. Its goal is to provide targeted therapy based on the molecular and clinical profiles of each cancer type (Pfohl et al., 2021). The strategy of a "one-size-fits-all" therapeutic approach assumed in most medical treatments which is based on treating the "average patient" may be successful for some patients and not for others. Therefore, it makes sense to eradicate this modality as it is both wasteful and may be associated with poor outcome.

There are powerful discoveries in tailored treatment which are precision medicine driven and some of them are Food and Drug Administration (FDA) approved. These treatments are tailored to characteristics of individuals such as a person's genetic makeup, or the genetic profile of an individual's tumour histological type or subtype. An example is patients with multiple primaries of malignant tumours that may include breast, endometrial and ovarian cancers, who routinely undergo molecular testing as part of patient care. Molecular testing of these tumours enables physicians to select treatments that improve chances of survival and reduce exposure to adverse effects per tumour site instead of an umbrella treatment program (Jones et al., 2019).

7.3 Artificial Intelligence

Artificial intelligence (AI) is a branch of computer science which deals with the simulation of intelligent behaviour in computers (Hamet & Tremblay, 2017). It relies on computers following specific algorithms established by humans or learned by computer methods to support decisions or execute certain tasks. Machine learning (ML) as a subfield of AI represents the process by which a computer is able to improve its own performance by continuously incorporating newly generated data into an existing iterative model. ML has three algorithms which are commonly used: (1) Supervised learning, (2) Unsupervised learning and (3) Reinforcement learning (Hamet & Tremblay, 2017).

A subfile of ML known as deep learning (DL) is where mathematical algorithms are deployed using multi-layered computational units resembling human cognition. These include neural networks (interconnections of computer processors similar to connections in the human brain) with different architecture types such as recurrent neural networks, convolutional, etc. (Hosny et al., 2018; Rajkomar et al., 2019).

At their most basic level, AI techniques are used to update patients on a personalized basis about their upcoming treatment procedures, progress, recovery, therapies used, dietary changes and modifications in lifestyles patterns along with the survival summary of previously recovered cancer patients. This allows patients to be more aware of their diseases and the entire clinical treatment procedures. Noting that

any new technological invention will have its shortfalls and advantages, the days and times for use of AI techniques to provide personalized treatment to cancer patients tailored to their needs in much quicker ways have arrived and we better get ready to embrace it (Dlamini et al., 2020).

7.4 Cancer Management and AI

7.4.1 Diagnosis and Treatment of Cancer

The management of common cancers with high morbidity, such as endometrial cancer, in modern days is moving towards use of AI. It is estimated that 19.3 million new cancer cases were diagnosed each year in the past decade. This figure is increasing each day and will likely continue to increase over the next few decades (Sung et al., 2021).

It is projected that 30.2 million new cancer cases will be diagnosed in 2040 despite substantial improvements in cancer diagnosis and management (Emens et al., 2017). Even though the current modalities in diagnosis and therapy for cancer have resulted in a reduction in cancer mortality over the last two decades, an alarming 10 million cancer-related deaths were recorded in 2020 (Sung et al., 2021). It is therefore imperative to promote bio-technological innovation in healthcare and more especially in cancer care.

The capacity of the human brain to process information is naturally limited. Processing of the modern big data is an urgent need for the implementation of alternative strategies. In addition to the increased availability of data, the augmentation of the storage and computing has boosted the development of data-processing techniques, such as ML and AI. ML and AI are becoming increasingly important tools to tackle complex issues in cancer care. There are more than 90 identified registered clinical trials for AI in cancer diagnosis with majority of trial recorded to have started after the year 2017 (Farina et al., 2022). This chapter provides an overview of the role of AI in oncology, including current applications, future perspectives and highlights some of the limitations of its use in an attempt to stimulate and prepare the reader for incorporation of ML and AI in cancer diagnosis.

Cancer is considered a complex disease biologically and requires that many parameters are addressed and considered correct before a medical decision is made at close to 10,000 parameters. For this extensive process, AI is the obvious solution and helps provide faster and more accurate interpretations of patient genomic and transcriptomic data (Dlamini et al., 2022).

7.4.2 Biomarkers

The discovery of biomarkers is one of the most promising techniques for cancer detection and is considered most effective when using molecular biomarkers. Biomarkers are used for diagnosis, disease prognosticating, predicting patient's response to therapy and estimating patient's overall survival after therapy (Dlamini et al., 2020). There are several roles of biomarkers including but not limited to:

(a) improving cancer treatment and its management.
(b) classify cancers into types or subtypes.
(c) stratify cancers based on the stage of disease. This is important because different types, subtypes or stages require different treatment modalities.
(d) check for drug resistance. Cancer drug resistance is a major obstacle for the successful treatment of cancer. The identification of resistance genes, epigenetic changes and the physiological pathways responsible for the development of such resistance will assist in solving this problem and improve treatment outcomes through direct targeting of the changes and therefore novel drugs can be developed. An example of such use would be for drug-resistant breast cancer due to the ESR1 oestrogen receptor.

The entire process involves determining if the presence of these biomarkers is associated with specific cancers, different stages of these cancers or different patient outcomes (Dlamini et al., 2020; Kumar et al., 2018). Relevant biomarkers have been identified through different omics technologies and the data is analysed using AI. If clinicians or clinician oncologists have to do physical biopsies, it requires tissue to be removed and examined histologically, whereas with the examination of molecular biomarkers via next-generation sequencing (NGS) is non-invasive and performed through blood tests. An example of a molecular biomarker used in most gynaecological malignancies is the circulating cancer antigen 125 (CA 125) used for the detection of ovarian cancer, monitoring of treatment response, prediction of metastasis and also in other epithelial cancers (Wang et al., 2019). However, this biomarker is non-specific as blood is not necessarily representative of the changes occurring in the tumour microenvironment (TME) and can also be elevated in other non-cancer conditions such as pregnancy, endometriosis, diverticulitis, liver cirrhosis and uterine fibroids (Cohen et al., 2014). The identification and characterization of novel biomarkers through NGS can assist through the detection of the presence of molecular or genetic alterations specific to a particular cancer. It also assists through detecting mutation signatures and tumour mutational burden (TMB) (Palmirotta et al., 2018). Using AI, advanced statistical and data analysis is applied to all these changes detected. Through RNA sequencing, information regarding changes in gene expression signatures in cancer is provided as well as the detection of mutations in RNA which can affect gene expression. These changes are all related to the underlying molecular mechanisms of cancer and therefore they can easily and practically be used as biomarkers of different types of cancer as well as for staging. Such can be

used to prognosticate and monitor treatment as previously highlighted (Dlamini et al., 2022).

7.4.3 *Challenges Facing the Application of AI to Cancer Diagnosis, Prognosis and Treatment*

There are several studies that have demonstrated how AI performs better than humans with regard to interpretation of the quantities of data pertaining to a complex disease such as cancer (Pashkov et al., 2020). It is important that we remind clinicians and clinician scientists that AI should always be used to augment human intelligence and not replace it. Any analysis of the medical data or biomarkers performed by AI should be assessed by qualified experts in their respective specialist clinical fields. The use of ML and DL in AI must also be assessed or supervised by experts in bioinformatics and programming. One of the biggest challenges facing the application of AI and DL to cancer diagnosis, prognosis and treatment is lack of knowledge concerning what the AI system is actually doing and how it comes to its final conclusions and therefore humans may blindly apply the report into clinical setting (Sorell et al., 2022). In the future, once AI is fully automated and requires no human intervention, it may become uncertain how an AI is selecting features or making decisions. The fear is that this may create doubt on the accuracy of the predictions made and indirectly force clinicians and researchers to accept these results on "as is" basis without scientific interrogations (Sorell et al., 2022). Although there are AI systems designed to be easily understood by physicians, and whose actions can be understood by the physicians and clinicians to improve their ease of use in clinical settings, there is still a long way before these are well understood by physicians who do not interact with AI on a daily basis or who do not possess IT knowledge (Dlamini et al., 2022).

The use of AI in prostate cancer was recently put in practice from a development using ML to predict if prostate cancer patients could effectively be treated using nerve-sparing radical prostatectomy. The AI did this by predicting whether a tumour could extend beyond the prostate (Kwong et al., 2022). All decisions and analyses produced by the AI could be analysed and explained in layman language using a publicly available web application, Shapley Additive Explanations (SHAP). However, even though this was achieved, DL also requires a large amount of data to learn enough to generate algorithms that it can be applied to new database. This is achieved through a process that requires multiple sampling to act as training data. This may not be accepted by patients. There is also an ethical issue as the use of big data is based on the use of patients' data which may occur without consent of the patient (Dlamini et al., 2022).

The other challenge in adopting AI to the oncology clinic setting is that the AI algorithms are not yet standardized for such a setting (Chua et al., 2021). This is very

difficult to achieve as cancer knowledge continues to expand and this can lead to information disparities.

7.5 The Solution to the Challenges of AI Applications

The challenges facing the implementation of AI and the use of big data are not insurmountable. Ethical problems related to the use of big data can be solved through individual country's policy makers intervention and the implementation of simple rules and guidelines governing its use (Dlamini et al., 2022).

The problem of AI being a mysterious black box that clinicians and oncologists would be uncomfortable using or trusting is being solved with the development and implementation of methods to test the accuracy of the predictions made by the system as well as revealing some of the decision-making processes made by the AI system. The challenge of the availability of training data will autocorrect as more studies that had ethical clearance are performed and data collected and stored. Such data can be used retrospectively to train the AI and as such this challenge will be overcome by the passage of time (Dlamini et al., 2022).

The enormous amounts of data required for storage capability to keep the genomic, transcriptomic, proteomic and medical record data for every patient should not pose a challenge as computer hardware development with "unlimited storage capacity" or use of cloud storage continues to advance.

7.6 Precision Medicine in Cancer Care

7.6.1 Introduction

The clinical trials on precision medicine have shown that drugs' adverse reactions can be avoided by treating cancers with genotype-focused pharmacological agents which are specific to a disease-causing gene mutation. This customized treatment permits for identification of specific genetic variations and associated response to individual drugs (Sicklick et al., 2019). Metabolomics is closely linked to the overall pathophysiological status of an individual. Thus, metabolomics may incorporate the biochemical events of thousands of small molecules in the cells, tissues, organs or biological fluids. The qualitative and quantitative alteration of the metabolite composition observed in disease pathologies or after drug administration translates into complex metabolic signatures (Holmes et al., 2008). The analysis of these signatures can potentially provide useful information for the diagnosis and prognosis of patients as well as for predicting pharmacological responses to specific interventions. Furthermore, specific metabolic signatures that occur after drug treatment provide information from pathways targeted or affected by drug therapy. In modern days, tumour genomic profiling is routinely used to classify tumour types, identify driver

or germline mutations, perform prognostic assessments and make future therapeutic decisions (Aboud & Weiss, 2013). However, the heterogeneity found in the cancer genomes and cancer tissues makes it difficult to determine the underlying causes or ascertain the optimal and best treatment. Moreover, the increased number of mutations and multiple combinations of tumour suppressors and oncogenes make individualized tumour classification or customized therapy almost impossible (Forbes et al., 2015).

7.6.2 A Summary of the Application of AI to Precision Medicine

Multi-omics data can be combined with multiple types of associated data such as published literature, medical images, protein models and gene expression data. AI algorithms use deep neural networks (DNN) and learning algorithms to create, test and improve models to achieve accurate predictions. The results of the networks and algorithms will be in association with biomarkers for diagnosis, prognosis estimation, treatment modification, prediction of survival outcomes and offer opportunities in fine-tuning and individualizing treatment plan. Where mutations are identified, such can be used to predict changes in protein structure (Dlamini et al., 2022). DNN can also be used in conjunction with histopathology, scans, MRI, mammograms, etc. to amplify the accuracy of the prediction model (Bhinder et al., 2021; Coudray et al., 2018). Both Figs. 7.2 and 7.3 depict the application of AI in precision oncology and genomics and how such have an impact on the pathological and pharmacological processes, as well as clinical biomarkers.

7.7 Application of AI to Metabolomics

7.7.1 Application in Therapeutics

In the management of colorectal carcinoma, the use of AI and Metabolomics helps clinicians choose the most efficient treatment with less adverse effect. AI technology is able to predict chemotherapy and radiotherapy response and assist the clinician in selecting the treatment modality that is likely to achieve the best therapeutic benefit. There are several ML models that have been developed to predict the toxicity of chemotherapy through collection of biomarkers at different stages of therapy (Qiu et al., 2022).

Omics data provide information such as prognostic signatures [Complexity Index in SARComas (CINSARC) and Genomic Grade Index (GGI)], transcriptome biomarkers which help establish biological pathways that are essential for understanding the disease mechanisms and treatment of tumours that would otherwise have

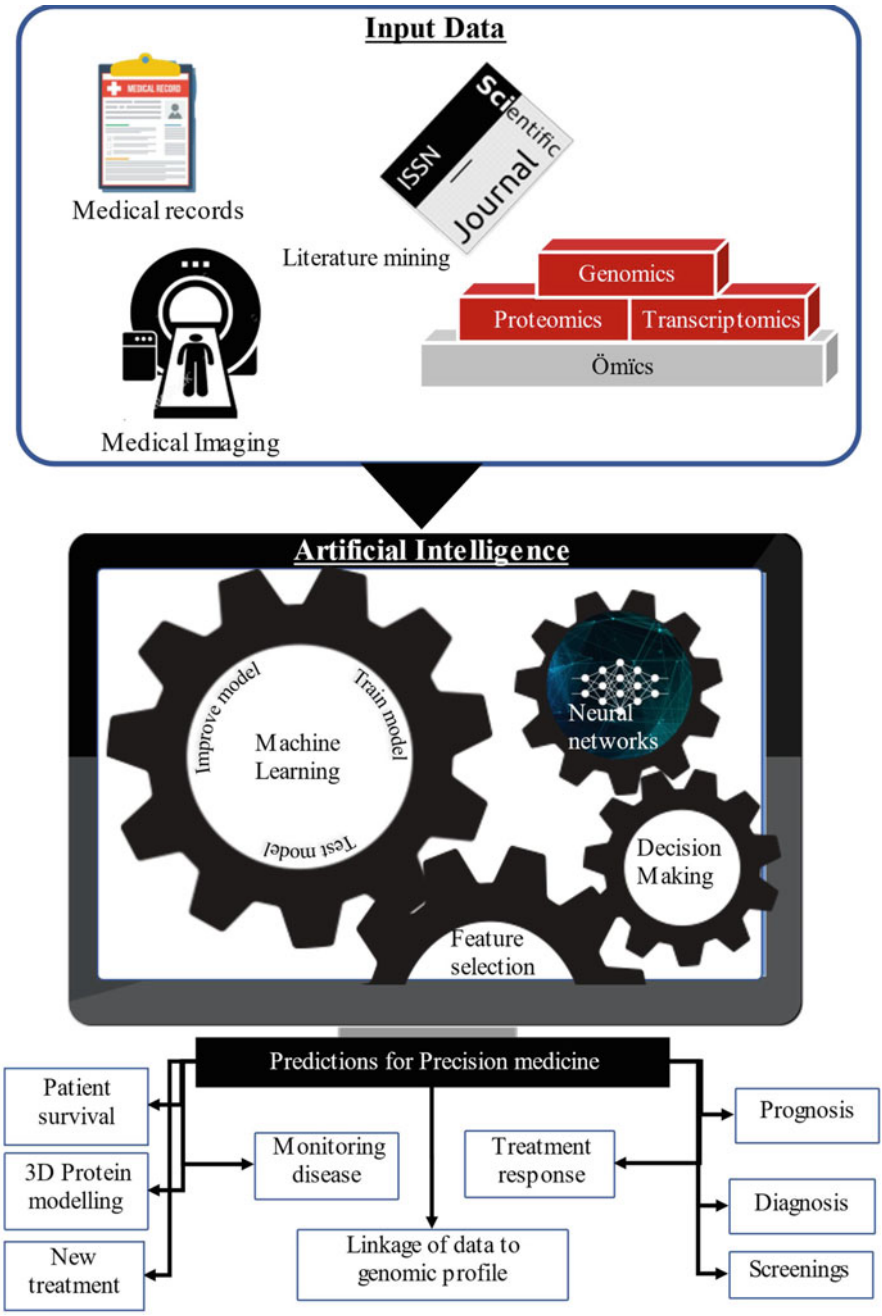


Fig. 7.2 Diagrammatic representation of the application of AI

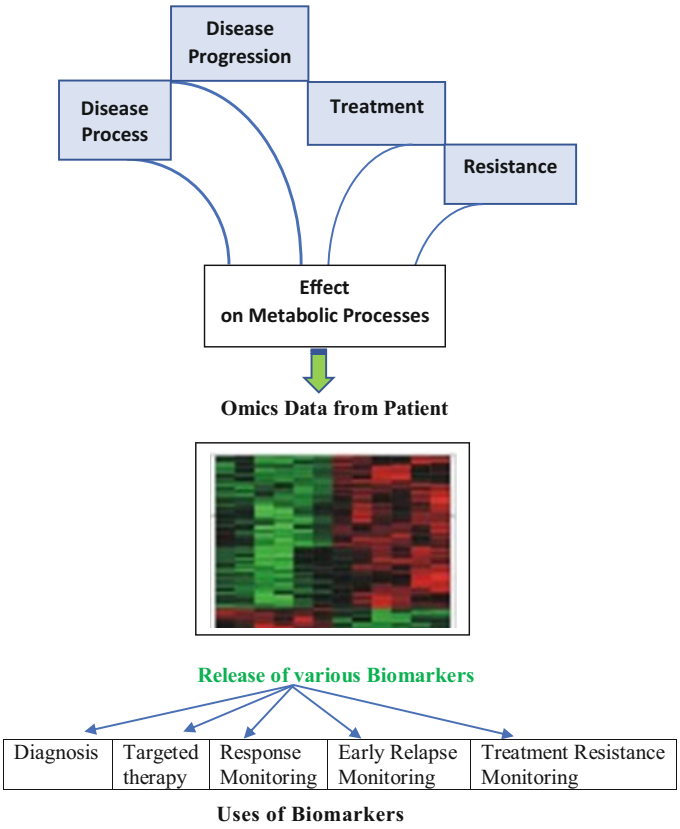


Fig. 7.3 Schematic diagram summarizing the impact of different pathological and pharmacological processes in the metabolism and its application to the identification of clinical biomarkers for precision medicine

poor prognosis such as sarcoma. The omics approach to sarcoma treatment also focuses on development of targeted therapy such as PARP inhibitors through novel therapeutic biomarkers (Zou et al., 2022).

7.8 Application in Imaging Genomics (Radiomics/Radiogenomics)

Imaging genomics, also known as radiomics or radiogenomics, describes the association of features of a tumour identified through tumour imaging with genomic data such as mutations, copy number variation (CNV) and gene expression profiles (Bodalal et al., 2019). The features of tumours such as structures, shapes, lines, points, colours, boundaries or the area of the image closely associated with one of

these features are identifiable. When applied to clinical medicine, these features aid to distinguish tumour tissue and normal tissue. The ability to make these distinctions has traditionally only been able to be performed by a human operator and therefore, subject to bias and interpretation and interpersonal variations in reported findings (Gürsoy Çoruh et al., 2021).

AI can analyse medical images without the need for human interference and with automation and consistency. In its current technological developmental state, AI is thought to outperform skilled and experienced pathologists in the use of medical imaging to diagnose cancer or make prognostic predictions. Imaging genomics relies on the use of AI to extract features identified on an image and link these features with phenotypes (Gürsoy Çoruh et al., 2021). The phenotype reflects protein expression, which is then associated with genomic, transcriptomic and epigenomic changes. This association can then be used to improve diagnostic and prognostic approaches. Therefore, medical imaging can be used to infer that these genetic changes are present within the tumour being imaged, allowing the use of image features as predictors of survival or indicators of the effects of treatment and as an accurate diagnostic tool than the conventional medical imaging modalities (Berger & Mardis, 2018). The most recent work by Yin and colleagues done in 2022 is a good example of the clinical application of the metabolomics, AI and precision medicine. They used an AI brain metastasis detection system which was more sensitive than three experienced and three junior radiologists (Dlamini et al., 2022).

Another application in clinical settings is its use to stratify the risk of mantle cell lymphoma (MCL) using CT-derived 3D images (Lisson et al., 2022). This is more reliable in predicting if the MCL patient would have a poor outcome, when compared to the use of traditional size measurements of the tumour as a prognostic tool. Because of lack of standardization, different clinical teams or units may use different feature selection process (Dlamini et al., 2022).

7.9 The Future of Cancer Care

AI and precision oncology in healthcare with the advancement in technology will revolutionize and transform the future of healthcare through the generation of big digital datasets acquired by means of NGS, use of algorithms for image processing, patient-related health records, data arising from large clinical trials and disease predictions (Dlamini et al., 2020). Oncology has been in the forefront of reaping the benefits of AI for universal cancer management in the form of early detection, tailored or targeted therapy by obtaining genetic information of the patient and predictions of future outcomes. AI's capabilities of pattern recognition and complex algorithms can be employed to gain relevant clinical information that will decrease errors related to diagnostics and therapy. The collective efforts of radiomics and DL will in the future deliver increased accuracy in diagnostic image analysis. The combined applications of AI and ML in healthcare will in the future be implemented

to improve disease management and provide effective medical care (Dlamini et al., 2020).

Cancer management will be improved by identifying clinically relevant biomarkers for early detection of disease and predicting prognosis for effective treatment. With AI and metabolomics, novel molecular biomarkers for different cancers can be deciphered by identifying germline mutations in DNA and whole transcriptome analysis by RNA sequencing. In the future, oncologists and oncology clinicians will use studies such as The Cancer Genome Atlas (TCGA) for RNA sequencing in biomarker identification for diagnosis and as a prognostic predictor. This data from the TCGA showed that despite differences in tumour biology, there was an overlap of molecular features in some tumour types and therefore revealed biomarkers that can predict the overall survival, disease-free survival and progression-free survival, which are essential endpoints in cancer management (Liu et al., 2018; Hutter & Zenklusen, 2018). A recent study utilized shallow RNA sequencing for predicting disease outcome and reduced costs of sequencing without compromising the biological data obtained for attaining accurate clinical insights (Milanez-Almeida et al., 2020). This can easily be applied in genome diagnostics in neuroblastoma, breast and lung cancers. Another recent study on the use of metabolomics in oncology gave hope of future utilization of circulating cell-free DNA (cfDNA) to analyse CNV as a "cost-effective, non-invasive, rapid, robust and sensitive alternative" for predicting the prognosis of malignancies such as neuroblastoma using a sequencing method (Van Roy et al., 2017). Classification of tumours is essential for treatment and prognosis. With NGS profiling, the classification of cancers can be conducted using circulating tumour DNA (ctDNA) and analysing CNV associated with cancers such as small lung cancer with no need for expensive invasive histological classification (Dlamini et al., 2020).

7.10 Conclusion

This chapter is intended to assist to further demonstrate the valuable role of specific areas of AI such as DL, ML and radiogenomics for diagnosis, treatment and prognostication. Metabolomics holds an indubitable potential in bringing solutions as far as cancer diagnostics, therapeutics and management are concerned. Integrating metabolomics technologies with AI indeed offers more effective cancer care strategies as it aids in overcoming challenges that either technique cannot overcome on its own. Further development in the AI technology and the algorithms for a personalized, clinically obtained genomics data gives hope of more targeted therapy for better cancer care.

References

Aboud, O. A., & Weiss, R. H. (2013). New opportunities from the cancer metabolome. *Clinical Chemistry, 59*, 138–146.

Alonso, A., Marsal, S., & Julià, A. (2015). Analytical methods in untargeted metabolomics: State of the art in 2015. *Frontiers in Bioengineering and Biotechnology, 3*, 23.

Berger, M. F., & Mardis, E. R. (2018). The emerging clinical relevance of genomics in cancer medicine. *Nature Reviews Clinical Oncology, 15*, 353–365.

Bhinder, B., Gilvary, C., Madhukar, N. S., & Elemento, O. (2021). Artificial intelligence in cancer research and precision medicine. *Cancer Discovery, 11*, 900–915.

Bodalal, Z., Trebeschi, S., Nguyen-Kim, T. D. L., Schats, W., & Beets-Tan, R. (2019). Radiogenomics: Bridging imaging and genomics. *Abdominal Radiology, 44*, 1960–1984.

Chu, X., Jaeger, M., Beumer, J., Bakker, O. B., Aguirre-Gamboa, R., Oosting, M., Smeekens, S. P., Moorlag, S., Mourits, V. P., & Koeken, V. A. (2021). Integration of metabolomics, genomics, and immune phenotypes reveals the causal roles of metabolites in disease. *Genome Biology, 22*, 1–22.

Chua, I. S., Gaziel-Yablowitz, M., Korach, Z. T., Kehl, K. L., Levitan, N. A., Arriaga, Y. E., Jackson, G. P., Bates, D. W., & Hassett, M. (2021). Artificial intelligence in oncology: Path to implementation. *Cancer Medicine, 10*, 4138–4149.

Chung, H.-J., Sim, J.-H., Min, T.-S., & Choi, H.-K. (2018). Metabolomics and lipidomics approaches in the science of probiotics: A review. *Journal of Medicinal Food, 21*, 1086–1095.

Clish, C. (2015). Metabolomics: An emerging but powerful tool for precision medicine. *Cold Spring Harbor Molecular Case Studies, 1*, A000588. https://Bit.Ly/3gox3is

Cohen, J. G., White, M., Cruz, A., & Farias-Eisner, R. (2014). In 2014, can we do better than Ca125 in the early detection of ovarian cancer? *World Journal of Biological Chemistry, 5*, 286.

Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A. L., Razavian, N., & Tsirigos, A. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature Medicine, 24*, 1559–1567.

Dettmer, K., Aronov, P. A., & Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews, 26*, 51–78.

Dlamini, Z., Francies, F. Z., Hull, R., & Marima, R. (2020). Artificial intelligence (Ai) and big data in cancer and precision oncology. *Computational and Structural Biotechnology Journal, 18*, 2300–2311.

Dlamini, Z., Skepu, A., Kim, N., Mkhabele, M., Khanyile, R., Molefi, T., Mbatha, S., Setlai, B., Mulaudzi, T., & Mabongo, M. (2022). Ai and precision oncology in clinical cancer genomics: From prevention to targeted cancer therapies-an outcomes based patient care. *Informatics in Medicine Unlocked*, 100965.

Edelstein, C. L. (2016). *Biomarkers of kidney disease*. Academic Press.

Emens, L. A., Ascierto, P. A., Darcy, P. K., Demaria, S., Eggermont, A. M., Redmond, W. L., Seliger, B., & Marincola, F. M. (2017). Cancer immunotherapy: Opportunities and challenges in the rapidly evolving clinical landscape. *European Journal of Cancer, 81*, 116–129.

Farina, E., Nabhen, J. J., Dacoregio, M. I., Batalini, F., & Moraes, F. Y. (2022). An overview of artificial intelligence in oncology. *Future Science OA, 8*, Fso787.

Forbes, S. A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., & Ward, S. (2015). Cosmic: Exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Research, 43*, D805–D811.

Gürsoy Çoruh, A., Yenigün, B., Uzun, Ç., Kahya, Y., Büyükceran, E. U., Elhan, A., Orhan, K., & Kayı Cangır, A. (2021). A comparison of the fusion model of deep learning neural networks with human observation for lung nodule detection and classification. *The British Journal of Radiology, 94*, 20210222.

Hamet, P., & Tremblay, J. (2017). Artificial intelligence in medicine. *Metabolism, 69*, S36–S40.

Holmes, E., Wilson, I. D., & Nicholson, J. K. (2008). Metabolic phenotyping in health and disease. *Cell, 134*, 714–717.

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer, 18*, 500–510.

Hutter, C., & Zenklusen, J. C. (2018). The cancer genome atlas: Creating lasting value beyond its data. *Cell, 173*, 283–285.

Johnson, C. H., Ivanisevic, J., & Siuzdak, G. (2016). Metabolomics: Beyond biomarkers and towards mechanisms. *Nature Reviews Molecular Cell Biology, 17*, 451–459.

Jones, D. T., Banito, A., Grünewald, T. G., Haber, M., Jäger, N., Kool, M., Milde, T., Molenaar, J. J., Nabbi, A., & Pugh, T. J. (2019). Molecular characteristics and therapeutic vulnerabilities across paediatric solid tumours. *Nature Reviews Cancer, 19*, 420–438.

Kim, H. K., Choi, Y. H., & Verpoorte, R. (2011). Nmr-based plant metabolomics: Where do we stand, where do we go? *Trends in Biotechnology, 29*, 267–275.

Kumar, S., Lindsay, D., Chen, Q. B., Garrett, A. L., Tan, X. M., Anders, C. K., Carey, L. A., & Gupta, G. P. (2018). Tracking plasma Dna mutation dynamics in estrogen receptor positive metastatic breast cancer with Dpcr-Seq. *NPJ Breast Cancer, 4*, 1–5.

Kwong, J. C., Khondker, A., Tran, C., Evans, E., Cozma, A. I., Javidan, A., Ali, A., Jamal, M., Short, T., & Papanikolaou, F. (2022). Explainable artificial intelligence to predict the risk of side-specific extraprostatic extension in pre-prostatectomy patients. *Canadian Urological Association Journal, 16*, 213.

Lisson, C. S., Lisson, C. G., Achilles, S., Mezger, M. F., Wolf, D., Schmidt, S. A., Thaiss, W. M., Bloehdorn, J., Beer, A. J., & Stilgenbauer, S. (2022). Longitudinal Ct imaging to explore the predictive power of 3d radiomic tumour heterogeneity in precise imaging of mantle cell lymphoma (Mcl). *Cancers, 14*, 393.

Liu, J., Lichtenberg, T., Hoadley, K. A., Poisson, L. M., Lazar, A. J., Cherniack, A. D., Kovatich, A. J., Benz, C. C., Levine, D. A., & Lee, A. V. (2018). An integrated Tcga pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell, 173*, 400–416.

Marion, D. (2013). An introduction to biological Nmr spectroscopy. *Molecular & Cellular Proteomics, 12*, 3006–3025.

Milanez-Almeida, P., Martins, A. J., Germain, R. N., & Tsang, J. S. (2020). Cancer prognosis with shallow tumor Rna sequencing. *Nature Medicine, 26*, 188–192.

Palmirotta, R., Lovero, D., Cafforio, P., Felici, C., Mannavola, F., Pellè, E., Quaresmini, D., Tucci, M., & Silvestris, F. (2018). Liquid biopsy of cancer: A multimodal diagnostic tool in clinical oncology. *Therapeutic Advances in Medical Oncology, 10*, 1758835918794630.

Pashkov, V. M., Harkusha, A. O., & Harkusha, Y. O. (2020). Artificial intelligence in medical practice: Regulative issues and perspectives.

Pfohl, U., Pflaume, A., Regenbrecht, M., Finkler, S., Graf Adelmann, Q., Reinhard, C., Regenbrecht, C. R., & Wedeken, L. (2021). Precision oncology beyond genomics: The future is here—it is just not evenly distributed. *Cell, 10*, 928.

Puchades-Carrasco, L., & Pineda-Lucena, A. (2017). Metabolomics applications in precision medicine: An oncological perspective. *Current Topics in Medicinal Chemistry, 17*, 2740–2751.

Qiu, H., Ding, S., Liu, J., Wang, L., & Wang, X. (2022). Applications of artificial intelligence in screening, diagnosis, treatment, and prognosis of colorectal cancer. *Current Oncology, 29*, 1773–1795.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*, 1347–1358.

Ren, J.-L., Zhang, A.-H., Kong, L., & Wang, X.-J. (2018). Advances in mass spectrometry-based metabolomics for investigation of metabolites. *RSC Advances, 8*, 22335–22350.

Robertson, D. G., Watkins, P. B., & Reily, M. D. (2011). Metabolomics in toxicology: Preclinical and clinical applications. *Toxicological Sciences, 120*, S146–S170.

Sas, K. M., Karnovsky, A., Michailidis, G., & Pennathur, S. (2015). Metabolomics and diabetes: Analytical and computational approaches. *Diabetes, 64*, 718–732.

Sicklick, J. K., Kato, S., Okamura, R., Schwaederle, M., Hahn, M. E., Williams, C. B., De, P., Krie, A., Piccioni, D. E., & Miller, V. A. (2019). Molecular profiling of cancer patients enables personalized combination therapy: The I-predict study. *Nature Medicine, 25*, 744–750.

Sorell, T., Rajpoot, N., & Verrill, C. (2022). Ethical issues in computational pathology. *Journal of Medical Ethics, 48*, 278–284.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a Cancer Journal for Clinicians, 71*, 209–249.

Tran, H., Mcconville, M., & Loukopoulos, P. (2020). Metabolomics in the study of spontaneous animal diseases. *Journal of Veterinary Diagnostic Investigation, 32*, 635–647.

Van Roy, N., Van Der Linden, M., Menten, B., Dheedene, A., Vandeputte, C., Van Dorpe, J., Laureys, G., Renard, M., Sante, T., & Lammens, T. (2017). Shallow whole genome sequencing on circulating cell-free Dna allows reliable noninvasive copy-number profiling in neuroblastoma patients noninvasive copy-number profiling using shallow Wgs. *Clinical Cancer Research, 23*, 6305–6314.

Wang, J., Dean, D. C., Hornicek, F. J., Shi, H., & Duan, Z. (2019). Rna sequencing (Rna-Seq) and its application in ovarian cancer. *Gynecologic Oncology, 152*, 194–201.

Wang, J. H., Byun, J., & Pennathur, S. (2010). Analytical approaches to metabolomics and applications to systems biology. *Seminars in Nephrology, 30*(5), 500–511.

Weckwerth, W. (2010). Metabolomics: An integral technique in systems biology. *Bioanalysis, 2*, 829–836.

Wilmes, A., Limonciel, A., Aschauer, L., Moenks, K., Bielow, C., Leonard, M. O., Hamon, J., Carpi, D., Ruzek, S., & Handler, A. (2013). Application of integrated transcriptomic, proteomic and metabolomic profiling for the delineation of mechanisms of drug induced cell stress. *Journal of Proteomics, 79*, 180–194.

Zhao, P., Li, J., Li, Y., Tian, Y., Yang, L., & Li, S. (2017). Integrating transcriptomics, proteomics, and metabolomics profiling with system pharmacology for the delineation of long-term therapeutic mechanisms of Bufei Jianpi formula in treating COPD. *Biomed Research International, 2017*, 7091087.

Zou, Z., Sun, W., Xu, Y., Liu, W., Zhong, J., Lin, X., & Chen, Y. (2022). Application of multi-omics approach in Sarcomas: A tool for studying mechanism, biomarkers, and therapeutic targets. *Frontiers in Oncology, 12*, 946022.

Chapter 8
Artificial Intelligence Application to Microbiomics Data for Improved Clinical Decision Making in Precision Oncology

Sikhumbuzo Z. Mbatha, Zilungile Lynette Mkhize-Kwitshana, Thanyani Victor Mulaudzi, Rodney Hull, and Zodwa Dlamini

Abstract The microbiome consists of the total number of micro-organisms and their genomes occupying a specific environment with distinct physico-chemical properties, at a specific time. Because of the sheer number of micro-organisms occupying the human body the microbiome offers substantially more genetic diversity, and hence more flexibility, than the human genome.

Disturbances in the microbiome have been linked to the development of various diseases including cancer. Microbes play a significant role in the development and progression of cancer by acting either directly or indirectly through a variety of mechanisms. It has been increasingly recognized that no two patients' cancers are exactly the same, and some of this heterogeneity can be directly attributable to the microbiome. Several studies have demonstrated links between specific microbes to specific cancers like *Helicobacter pylori* and gastric cancer, human papillomavirus and cervical squamous cell carcinoma, *Fusobacterium nucleatum* and colorectal cancer, etc.

The use of informatics, machine learning, and artificial intelligence assists with the interpretation of data generated from multi-omics technologies and real-world data, enhancing decision-support systems in precision oncology. A more in-depth

S. Z. Mbatha (✉)
Department of Surgery, University of Pretoria, Pretoria, South Africa
e-mail: sikhumbuzo.mbatha@up.ac.za

Z. L. Mkhize-Kwitshana
Department of Medical Microbiology, School of Laboratory Medicine & Medical Sciences University of KwaZulu-Natal, Durban, South Africa

SAMRC Research Capacity Development Division, South African Medical Research Council, Tygerberg, Cape Town, South Africa
e-mail: MkhizeKwitshanaZ@ukzn.ac.za

T. V. Mulaudzi · R. Hull · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: thanyani.mulaudzi@up.ac.za; rodney.hull@up.ac.za; zodwa.dlamini@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_8

understanding of the mechanisms of microbiota-mediated immunomodulation and identification of precise immune-stimulatory and immune-inhibiting bacterial strains or pathways could lead to increased precision in cancer therapy.

Keywords Microbiome · Microbiomics · Machine learning · Big data · Precision oncology · Artificial intelligence

8.1 Introduction

Big data, information technology, and artificial intelligence (AI) promise to change the world. Big data analysis is applied in many fields of human activities and must also be productively employed in the field of medicine. To come to grips with the extreme biological complexity of our organism and each of our organs, the completeness of enormous amounts of data is of extraordinary worth if assessed holistically with the -omics disciplines (Vassilios, 2016). The introduction of AI technology with machine learning and deep learning into the medical field including the field of oncology is inevitable. More than 60 AI-equipped medical devices have been approved by the FDA in the USA with several of them already being used for clinical applications, especially in radiology (Hamamoto et al., 2020). The role of the microbiome is one of the most topical aspects and concepts that the medical fraternity is grappling with trying to find more answers about the cause, effect, and possible treatment of a variety of diseases including cancer. The microbiome consists of all micro-organisms (living and non-living with their genome and theater of activity) which interact with each other, live in the same habitat, and form their ecological niche together (Moss, 2017).

Compared to many other fields of multi-omic studies, microbiomes are dynamic ecosystems with active host regulation. As a result, the analysis and interpretation of generated data tends to be more complex and challenging (Nayfach et al., 2019; Topçuoğlu et al., 2020). Based on current observations, the degree of personalization of the human microbiome vastly exceeds that of the host genome with no two individuals showing an identical overlap in the microbial species of their microbiome. This degree of personalization is so high that it may even have forensic applications (Gilbert et al., 2018). Over the past decade more and more microbiome researchers have begun using machine learning algorithms, as they increasingly become aware of the ability of these models to incorporate and assess the interpersonal microbiome variations and ecology, and the ability of artificial intelligence to assess each microbial population together with its co-existent microbial population rather than in isolation. Machine learning offers a holistic view incorporating the structure of the microbial communities as a whole and identifies linkages between the microbes and disease state (Topçuoğlu et al., 2020). The application of machine learning algorithms has proven to be valuable in identifying predictive characteristics of a microbial signature (Gilbert et al., 2018). Changes in the microbial population of the microbiome result in dysbiosis with subsequent onset, flare-up, and resistance to therapy of disease. The ability to manage these changes can represent

the medicine of the future, by acting on modifiable entities within the microbiome, as opposed to the human genome which is not modifiable. Data from microbiomics integrated with patient metadata and other omics data may provide advanced decision-support systems allowing us to further advance precision medicine and oncology (Putignani et al., 2019).

8.2 The Human Microbiome

The human microbiome can be defined as the total number of micro-organisms and their genomes that exist in the human body, occupying a specific environment or region with distinct physiological and chemical properties, at a specific time. As such the microbiome does not only define the microbiota but also their genomic characteristics and theater of activity (Kambouris, 2020). These micro-organisms include viruses, fungi, protozoa, archaea, and bacteria. The microbiome can be characterized as consisting of a core microbiome and a variable microbiome. The core human microbiome is defined as any set of microbial taxa and their associated genomic or functional attributes characteristic of the human host and present in a specific body region in all or large majority of human beings (Turnbaugh et al., 2007; Neu et al., 2021). The number of micro-organisms that exist and coexist in the human body is enormous with approximately 500–1000 species of just bacteria found in the human body at any given time and it is estimated that the microbiota outnumbers human somatic and germ cells by a factor of ten (Turnbaugh et al., 2007; Gilbert et al., 2018). The Human Microbiome Project (HMP) was initiated in 2008 to explore the microbial communities of the human host and characterize their role in human health and disease (Creasy et al., 2021). One of the primary objectives of the HMP was to understand the microbiome and the factors that influence its distribution and evolution over time and its impact on human genetic and physiological diversity (Turnbaugh et al., 2007). The human microbiome is not homogenous within a single individual, with each region of the body having its unique microbiome, largely as a result of the early phases of human development where the microbiome concurrently develops in a regional as well as body site-specific direction leading to each body site having its specific biogeography (Gilbert et al., 2018; Hull et al., 2021).

The skin, for example, shows dramatic variation in microbiome composition and structure across its different sites (e.g. elbow and face) (Gilbert et al., 2018). The human microbiome is a living ecosystem that does not operate in isolation but interacts closely with and is influenced by the environmental microbiome found in the temporo-spatial environment that an individual occupies (Ramaboli et al., 2022). A long co-evolutionary process has enabled the development of mutually beneficial interaction between the human host and the trillions of microbiota living within and on the human, with humans demonstrating traits that they did not evolve on their own but through the direct influence of their co-existent microbial genomes (Rajendhran & Gunasekaran, 2010). Disturbances in the microbiome have been linked to the development of various diseases, for example, the state of the adult

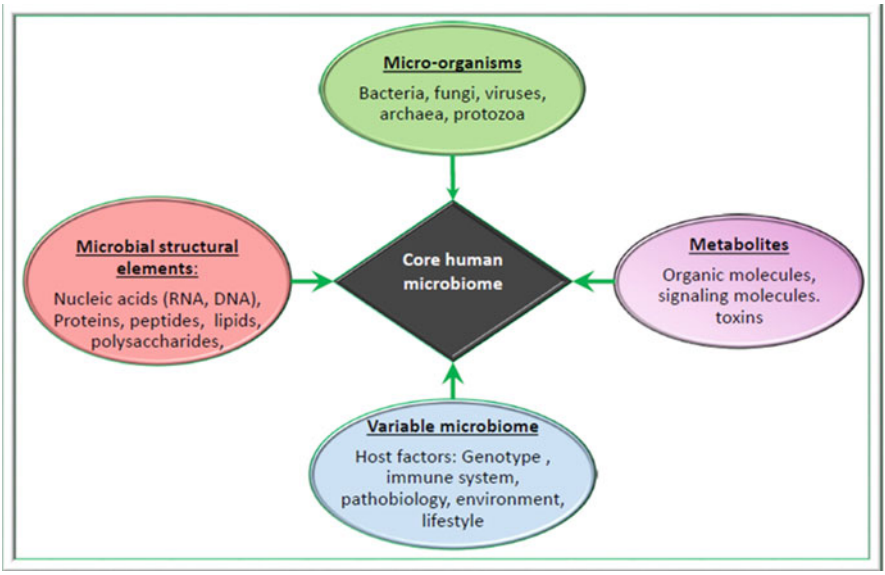


Fig. 8.1 The human microbiome. The core human microbiome is the part of microbiome present in a given habitat in all or the vast majority of humans, where habitat can be the whole human or a specific region, e.g., gut. Micro-organism is the total number of micro-organisms including living (bacteria) and non-living (viruses) organism. Variable microbiome is variation as a result from a combination of host specific factors, e.g., genotype, host physiological status, immune system, pathobiology (disease status), lifestyle (including diet), host environment (at home and/or work) and the presence of transient populations of micro-organisms that cannot persistently colonize a habitat. Microbial structural elements and metabolites constitute the theater of activity of the micro-organisms

human gut, which can harbor up to ten times more microbial cells than the total number of somatic and germ cells, has been linked to the health status of patients such as cancer development (Ramaboli et al., 2022; Shukla et al., 2015). Explorative analysis of the human microbiome should be coupled with an assessment of the environmental exposure including exposure to pathogens, nutrients, drugs, and pollution (Putignani et al., 2019). A growing number of studies have shown that it is subsets of microbiomes that cause health differences among individuals and it is rare for a single bacterial species to be associated with a disease (Topçuoğlu et al., 2020). Microbiome research has significantly improved recently with the advent of high-throughput technologies generating large amounts of data and metadata with the ability to not only inexpensively and easily detect microbiota but also measure their metabolic activity (Gaitanis, 2020). Figure 8.1 below captures the concept of the human microbiome.

8.3 The Microbiome and Cancer

Over the past decade, it has been increasingly recognized that no two patients' cancers are exactly the same, and some of this heterogeneity can be directly attributable to the microbiome (Rodriguez et al., 2021). The gut microbiome composition gradually shifts throughout an individual's lifetime and varies widely between individuals which impact carcinogenesis, disease development, antitumor immunity and clinical response to therapy including immunotherapy (Matson et al., 2021). Disruption of the microbiome (dysbiosis) results in immune dysregulation and impaired ability to control tumorigenesis. Dysbiosis can be characterized by alteration in proportions of certain phyla, increased or decreased abundance, increase in species that promote chronic inflammation (and oncogenesis) or reduction of species that down-modulate chronic inflammation or with antitumor effects (Vivarelli et al., 2019). See Fig. 8.2 below.

Microbiota play a significant role in the development and progression of cancer by acting either directly or indirectly through a variety of mechanisms including the inactivation of some chemotherapeutic agents, production of mutagens, and the promotion of inflammation by reactive oxygen species inducing DNA damage, stimulation of cancer-promoting signaling pathways such as the e-cadherin–Wnt–b-catenin signaling pathway. The e-cadherin–β catenin complex regulates cellular adhesion, thus any disruption leads to unchecked cell migration, invasion, and metastasis (Hull et al., 2021). Some of the breakthrough studies confirming the cancer causality by microbes were studies demonstrating that H. pylori greatly increases the risk of non-cardia gastric cancer, the discovery of oncogenic HPV strains and their strong association with squamous cell cervical carcinoma and the association of hepatocellular carcinoma with chronic viral hepatitis caused by hepatitis B and C viruses (HBV/HCV) (Gaitanis, 2020). A study by Ponziani et al. (2019) investigated gut microbiota features associated with hepatocellular carcinoma (HCC) in patients with complex phenotypes, such as cirrhosis and nonalcoholic fatty liver disease (NAFLD), and found that patients with liver cirrhosis and hepatocellular carcinoma had, amongst other things, deficiency of beneficial bacteria, namely *Akkermansia* and *Bifidobacterium* and that this combined deficiency enhances intestinal and liver inflammation, influencing the initiation and/or the progression of hepato-carcinogenesis. They concluded that in patients with cirrhosis and NAFLD the gut microbiota profile and systemic inflammation are significantly correlated and can concur in hepato-carcinogenesis processes (Ponziani et al., 2019). Persistence of *Helicobacter pylori* infection has emerged as a cause of gastric cancer, and H. pylori is recognized as a type I human carcinogen, with 90% of gastric cancer cases worldwide considered to be associated with *H. pylori* infection and approximately 10% to *Epstein-Barr* virus infection. *H. pylori* drives carcinogenesis by chronic inflammation leading to the Correa cascade of chronic atrophic gastritis, gastric intestinal metaplasia, dysplasia, and adenocarcinoma. H. pylori induces the recruitment of neutrophils and macrophages, which, in turn, produce reactive oxygen species (ROS) and nitrogen species (RNS). Inflammation-mediated ROS/RNS can

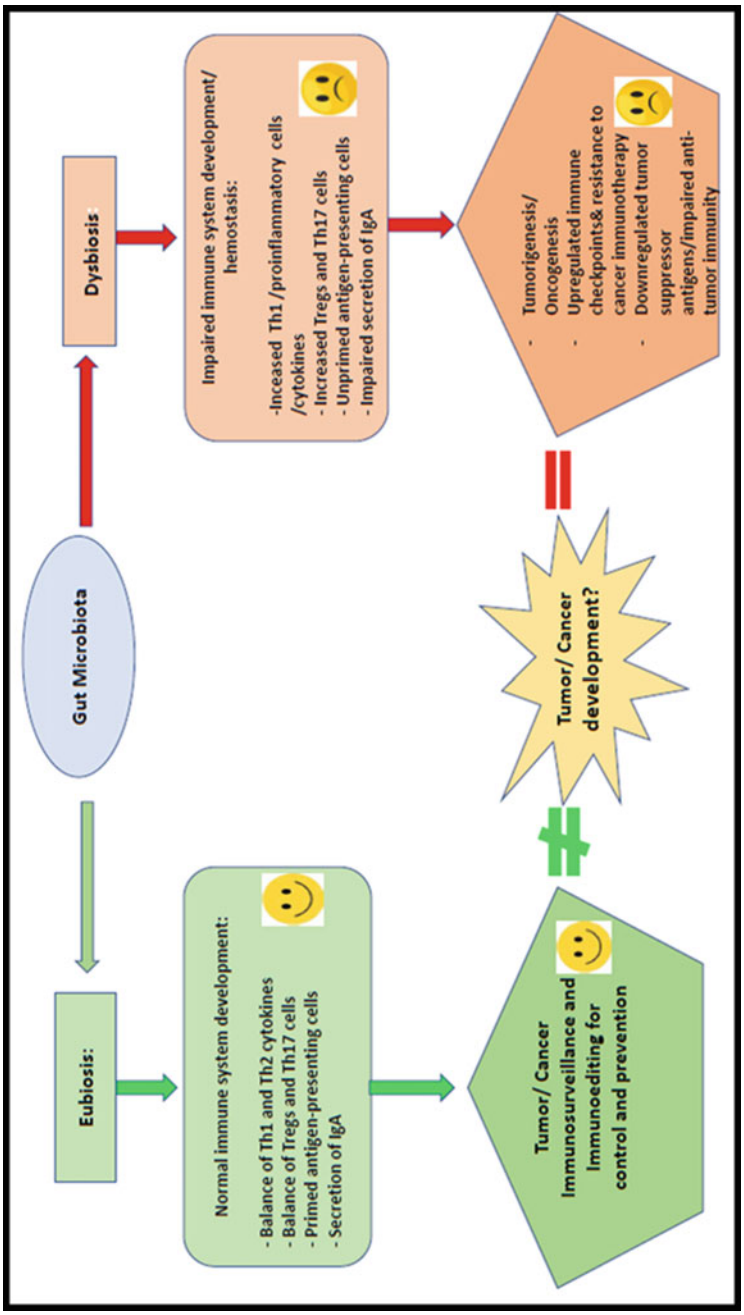


Fig. 8.2 Schematic overview of some immunological interactions between microbiome balance and cancer prevention (≠ green) and dysbiosis and cancer promotion (= red). Microbiome acquired in utero during birth through infancy co-evolve with and help in the development of a balanced immune system that is able to respond appropriately to antigens including cancer antigens; constant interaction with the balanced microbiome throughout life maintains the immune system homeostasis

directly trigger single-strand DNA breaks (SSB) and/or induce the NF-*B* pro-inflammatory pathway that can trigger double-strand DNA breaks (DSB) (Knippel et al., 2021). Some data suggest that successful eradication of H. pylori reduces the risk and rate of gastric cancer (rates reduced from 3% to 1.6% in a meta-analysis and from 13.4% to 7.2% in another study) (Gaitanis, 2020; Knippel et al., 2021). *Fusobacterium nucleatum* has been studied extensively and mechanistically linked as a bacterial driver of tumorigenesis in colorectal cancer (CRC) development via activation of β-catenin signaling and by driving inflammatory responses. High levels of intra-tumoral *F. nucleatum* in CRC are associated with low survival rates, chemo-resistance, and evasion of immunity (Matson et al., 2021).

A study analyzing the microbiota profile and metabolites profile from fecal samples collected from 50 CRC patients and 50 healthy controls observed a three-fold increase in *Proteobacteria*, a 60-fold increase in *Fusobacteria*, and a 0.5-fold decrease in *Firmicutes* in CRC patients compared with the healthy control group and microbes were much less abundant in the CRC group, indicating a significant shift in the gut microbiome of CRC patients (Yang et al., 2019). However, the microbiome does not only lead to the development and progression of cancer but there also exist several favorable microbiotas that are associated with decreased risk of cancer, improved survival, and better response to therapy including immunotherapy. Hayes et al. in a prospective study nested in 2 large US cohorts assessed the oral microbiota with high-throughput sequencing of the 16S ribosomal RNA (16S rRNA) gene in pre-diagnostic oral samples from 129 Head and Neck Squamous Cell Carcinoma (HNSCC) cases and 254 controls, and they found that the greater abundance of the commensal bacterial genera, *Corynebacterium* and *Kingella,* was associated with reduced risk of HNSCC (Hayes et al., 2018). Riquelme et al. in their study found that patients with high alpha diversity of the pancreatic ductal carcinoma (PDAC) tumor microbiome had better overall survival compared to those with low alpha diversity, where alpha diversity is defined as the number of species present within the tumor specimen. They found enrichment of *Saccharopolyspora, Pseudoxanthomonas, Streptomyces*, and *Bacillus clausii* within the PDAC specimens of the long-term survivor cohort and demonstrated that this enrichment is strongly predictive of prolonged survival in resected PDAC (Riquelme et al., 2019). A number of organisms have been positively linked with cancer but only 10 have been classified as true oncogenic as per the International Agency for Cancer Research (IACR), see Table 8.1 below.

Of the trillions of microbes inhabiting the human body only ten microbes have been classified as proper oncogenic, leading to cancer (seven viruses, three parasites, and one bacteria). The other cancer associated microbes are regarded as complicit micro-organisms (i.e., known to have some impact in cancer initiation, progression, and metastasis but not necessarily proven to be the main causative event).

Table 8.1 Micro-organism linked to cancer

Bacteria associated with cancer			
Micro-organism	Cancer	Mechanism	Reference
Helicobacter pylori	Intestinal type gastric cancer	DNA methylation, cyclin E over-expression, p27 degradation, mutation in *TP53 KRAS, APC*. Induces Correa cascade.	Moss (2017), Matson et al. (2021), Eguchi et al. (2004)
Fusobacterium nucleatum	Colorectal cancer	Induces a pro-inflammatory micro-environment and suppression of host immunity. Binds E-cadherin, activates the β-catenin pathway, and induces the expression LEF/T-cell factor (TCF).	Matson et al. (2021), Wu et al. (2019)
Bacteroides fragilis	Colorectal cancer	Activates host colonic epithelial cell (CEC) NF-κB and STAT3 pathways and CEC Wnt signaling. E-cadherin cleavage, induction of c-Myc expression, upregulation of CEACAM6, downregulation of MUC2.	Knippel et al. (2021), Shariati et al. (2021)
Akkermansia and *Bifidobacterium*	Hepatocellular carcinoma	Combined deficiency enhances intestinal and liver inflammation.	Ponziani et al. (2019)
Acidovorax	Lung SCC	Abundance → promotes tumorigenesis. TP53 mutations.	Knippel et al. (2021), Greathouse et al. (2018)
Helicobacter spp., S. typhi	Gallbladder cancer	Mucosal alterations, inflammation, weakening and mucosal dysplasia.	Allegra et al. (2019)
Streptococcus mitis, Neisseria elongata	Pancreatic cancer	Mucosal alterations, inflammation, weakening and mucosal dysplasia.	Allegra et al. (2019)
A. vaginae, Porphyromonas sp.	Endometrial cancer	Increased vaginal pH.	Allegra et al. (2019)
Confirmed oncogenic microbes as per International Agency for Cancer Research (IACR)			
Epstein-Barr Virus	Burkitt lymphoma, B-cell and NK-cell lymphoma	Immortalizes B lymphocytes→ expression of multiple viral proteins, increased proliferation of infected cells, blocking of apoptosis, cell migration and inducing genomic instability.	Cmrečak et al. (2021), Allegra et al. (2019)

(continued)

Table 8.1 (continued)

Bacteria associated with cancer			
Micro-organism	Cancer	Mechanism	Reference
Hepatitis B and C virus (HBV, HCV)	Hepatocellular carcinoma	Vicious cycle of hepatocyte regeneration and necrosis → mutation accumulation, telomerase reactivation → chronic inflammation. Interference with RAF/MAPK/ERK and Wnt/β-catenin signaling pathways and blocking TNF-α mediated apoptosis.	Cmrečak et al. (2021), Allegra et al. (2019)
Human papillomavirus (HPV)	Cervical SCC. Head and neck SCC	E6 and E7 onco-proteins proliferation E6 targets p53 interfering with apoptosis. E7 targets tumor suppressor protein Rb → proliferation and cell differentiation disruption.	Cmrečak et al. (2021), Allegra et al. (2019)
Kaposi sarcoma herpes virus (KSHV)	Kaposi sarcoma primary effusion lymphoma	Latent viral proteins → carcinogenesis by stimulating cell proliferation, anti-apoptotic activity, deregulation of the cell cycle, avoidance, and modulation of immune response.	Allegra et al. (2019), Cmrečak et al. (2021)
Human T-lymphotropic virus (HTLV)	Adult T-cell leukemia/lymphoma (ATL). Non-Hodgkin's lymphoma	Abnormal DNA replication due to increase of B lymphocyte growth. Oxidative stress. Oncogenic activation.	Cmrečak et al. (2021), Allegra et al. (2019)
Human immunodeficiency virus (HIV)		Indirectly increases the risk of cancer by immune-suppression → reactivation of other cancer-related viruses.	Cmrečak et al. (2021), Allegra et al. (2019)
Helicobacter pylori	As above		Garrett (2015), Cmrečak et al. (2021)
Parasites with oncogenic potential			
Schistosoma haematobium	Bladder cancer	Schistosoma eggs in bladder → accumulation of inflammatory cells and production of oxygen-derived free radicals.	Cmrečak et al. (2021), Allegra et al. (2019)
Clonorchis sinensis *Opisthorchis viverrini*	Cholangiocarcinoma	Chronic inflammation in intrahepatic bile ducts Reduced apoptosis, upregulation of Bcl-2, downregulation of p27, augmented cell invasion.	Allegra et al. (2019), Cmrečak et al. (2021)

8.4 Microbiomics

Microbiomics, which is the science of collective characterization and quantification of molecules responsible for the structure, function, and dynamics of a microbial community (Kumar, 2021), is by its nature and character particularly prone to the generation of big data. The human microbiome is a complex entity that is a living ecosystem affected by temporal and spatial influences with each of its constituents undergoing fluctuations in growth rate and survival (Gilbert et al., 2018). Human beings are regarded as "superorganisms" with trillions of associated micro-organisms, co-existing in a symbiotic relationship where each is vital for survival (Rajendhran & Gunasekaran, 2010). The Human Microbiome project revealed that commensal microbial genes exceed the total number of human genes by 100:1 and an adult human gut can harbor up to 100 trillion microbial cells; ten times higher than the total number of somatic and germ cells (Shukla et al., 2015). Each bacterial strain has a genome containing thousands of genes, offering substantially more genetic diversity, and hence more flexibility, than the human genome (Gilbert et al., 2018). This wide genetic diversity of human microbiota results in a wide range of metabolic activities, which are crucial for understanding the fundamental mechanisms of host-microbial crosstalk (Ramaboli et al., 2022). As of 2020 the Human Microbiome Project Data Coordination Center (HMPDACC) Data Portal contained 48 TB of data composed of raw and processed data from both host and microbiome generated from different cohorts from the initial and integrated phases of the HMP. This data is associated with more than 31,000 samples (Creasy et al., 2021). The Human Microbiome Project and systems biology approaches to study intestinal microbiome have generated multitudes of omics data (big data) and the commensurate powerful analytical tools for efficient description of the human microbiome (Putignani et al., 2019). Microbiomics enables the study of microbial communities in their natural environment with their natural partners obviating the need to have to culture them in order to study them in isolation. It has the unique capacity to simultaneously provide a telescopic, big picture view of the dynamics of an entire community and a microscopic view of the behavior of a single gene, protein, or metabolite across large populations. This ability to study microbial communities as a whole and in their natural environments enabled the discovery not only of new species, but also attributed novel metabolic pathways, interactions, and behaviors to them (Kumar, 2021). Compared to many other fields of multi-omic studies, microbiomes are dynamic ecosystems with active host regulation. This adds interesting new dimensions and complexity to the analyses and interpretation of data. Some of the common data types used in microbiome research include amplicon data, shotgun metagenomics data, meta-transcriptome data, and other-omics data such as metabolomics and meta-proteomics data (Nayfach et al., 2019; Topçuoğlu et al., 2020).

8.4.1 Techniques in Microbiomics

8.4.1.1 Quantitative Microbial Profiling Methods

Human beings are regarded as super-organism or holo-bionts because of the micro-organisms that exist within their bodies in a symbiotic relationship, and the genomes of these micro-organisms are responsible for a host of important human metabolic activities (Kumar, 2021; Rajendhran & Gunasekaran, 2010). Therefore, to fully comprehend the whole human genome we must also characterize the human microbiomes. There are two main methods for microbiome analysis that do not rely on growing organisms in pure culture: microbial community profiling by conducting amplicon gene sequences of small-subunit ribosomal RNA (rRNA): the 16S rRNA (for archaea and bacteria) or the 18S rRNA (for eukaryotes), and metagenomics (analysis of all genomes in microbiome ecosystem) studies where community microbial DNA is subjected to whole genome shotgun sequencing (Hamady & Knight, 2009). The genome sequence data obtained through sequencing of microbial samples can then be processed using programs that are specific to the microbiome such as QIIME or Mothur. The 16S rRNA gene sequences are clustered into operational taxonomic unit (OTU), representing a particular microbial taxon abundance (Namkung, 2020; Marcos-Zambrano et al., 2021).

8.4.1.2 Multi-omics Technologies

Multi-omics technologies that probe the gene expression of the microbial genomes which include meta-transcriptomics, meta-proteomics, and metabolomics provide more detailed information on microbial activities in the environment; giving us greater insight into functional potential and the expression profile of microbiome-derived bioactive molecules (Moss, 2017; Hamady & Knight, 2009; Kashyap et al., 2017). In this sense, meta-transcriptomics is the characterization of gene expression from all microbes in an ecosystem, metabolomics is the characterization of all small molecule metabolites in an ecosystem, and meta-proteomics is the characterization of all proteins in an ecosystem (Kashyap et al., 2017).

8.5 Artificial Intelligence: Big Data and Machine Learning

8.5.1 Big Data

Big data in the field of medicine constitutes multiple types of data from the individual patient's data derived from their psychosocioeconomic circumstances, demographics, clinico-laboratory information, radiological studies, histopathology, and the -omics studies data generated by high-throughput technologies to population

data generated from population-wide genomic studies and electronic health records of large patient populations. Machine learning is a fundamental technology that will allow us to integrate and meaningfully process vast amounts of critical data that is beyond the capacity of the human brain to analyze and comprehend within a reasonable space of time (Rajkomar et al., 2019). The molecular technique of next-generation sequencing (NGS) generates large amounts of data related to mutated genes that are collated into gigabytes, terabytes, or petabytes of data, and as such analysis of this massive data requires the use of robust computational approaches to exploit the information effectively (Dlamini et al., 2020; Mirsadeghi et al., 2021). The use of informatics and artificial intelligence is crucial to assist with the interpretation of the huge amounts of data generated from the multi-omics technologies and real-world data that need to be collected, categorized, and analyzed to enhance decision-support systems (De Maria Marchiano et al., 2021). The machine learning algorithms can be trained with data from countless patients whereas it is too difficult for human physicians and biologists to gain such experience in an entire career or their research (Mirsadeghi et al., 2021). Sequencing a single genome, for instance, will produce over 100 gigabytes of data (Topçuoğlu et al., 2020). The AI models give experts the ability to take appropriate clinical decisions ensuring that every diagnosis, management decision, and therapy is personalized based on all known information about a patient, in real-time, while incorporating lessons from a collective experience of the whole (Rajkomar et al., 2019; Mirsadeghi et al., 2021).

8.5.2 Machine Learning in Microbiomics

Machine learning involves models making inferences from available data to identify, classify, and predict patterns and to learn new tasks (Marcos-Zambrano et al., 2021). The application of machine learning algorithms has proven to be valuable in identifying predictive characteristics of a microbial signature (Gilbert et al., 2018). Large-scale microbiome studies resulting in large-scale datasets have been rapidly increasing as sequencing technology has become less costly and has advanced. This has allowed the use of machine learning algorithms to characterize the functional relationships between specific microbiome and the ecosystem properties (Namkung, 2020). Machine learning in microbiomics can be either supervised (use of labeled input datasets to create an algorithm that accurately predicts outcome or classify data) or unsupervised (model learns by analysis and clustering of unlabeled data, identifying patterns and new associations from the presented unlabeled datasets). Different approaches of machine learning can be applied to extrapolate input data, classify, interpret, and predict associated cancer susceptibility and/or treatment outcomes (Table 8.2).

A variety of machine learning algorithms can be used in microbiome studies from supervised learning algorithms that deal with classification issues and uncovering

Table 8.2 Examples of machine learning models used in microbiome studies

Model	Description	Study where model used	References
Supervised models			
Logistic regression	Solves classification by fitting linear model to data assessing the probability of an outcome for a binary variable.	Beck & foster: used LR to uncover possible microbial interactions associated with bacterial vaginosis diagnosis.	Marcos-Zambrano et al. (2021), Zhou and Gallins (2019)
Linear discriminant analysis (LDA)	Finds a linear combination of microbial features in the training data that models the multivariate mean differences between classes.	Segatta *et al*: used LDA effect size (LEfSe) algorithm to detect bacterial organisms and functional characteristics differentially abundant between two or more microbial environments.	Segata et al. (2011), Marcos-Zambrano et al. (2021), Zhou and Gallins (2019)
k-nearest neighbors (k-NN)	Uses proximity to assign a new sample to a class with majority of characteristics nearest to individual data point.	Asher & basher: predicted species' abundance profiles based on their presence/absence configuration using the kNN regression algorithm.	Asher and Bashan (2022), Marcos-Zambrano et al. (2021)
Naïve Bayes classifiers	Probabilistic machine learning algorithm based on the Bayes Theorem used for classification tasks.	Werner *et al* compared naïve Bayesian taxonomic classification results using training sets built from three different reference databases of varying diversity and overall taxonomic structure.	Werner et al. (2012), Marcos-Zambrano et al. (2021)
Support vector machines (SVM)	Determines a linear or nonlinear separating line/decision boundary in the given dataset to make the largest distance or margin to the nearest training data points of any classes.	Liu *et al* used SVM for the quantitative prediction of medium-chain carboxylate production in two continuous anaerobic bioreactors from 16S rRNA gene dynamics in enriched communities**.**	Liu et al. (2022), Zhou and Gallins (2019), Namkung (2020)
Artificial neural networks	Machine learning based on mimicking the functioning of the human brain's neural network, learning from complex nonlinear relationships.	Lo & Marculescu: proposed a new neural network-based pipeline that is suitable for classifying metagenomic datasets.	Lo and Marculescu (2019), Marcos-Zambrano et al. (2021)
Deep Learning	Uses artificial neural networks (ANNs) with deep architectures, i.e., multiple hidden layers, yielding a higher level of abstraction.	Reiman et al.: developed a CNN model for classification of a microbiome sample based on its microbial taxonomic abundance profile**.**	Reiman et al. (2017)

(continued)

Table 8.2 (continued)

Model	Description	Study where model used	References
Ensemble Methods Combines multiple classifiers to obtain better performance than a single classifier:			
Random forests (RF)	Ensemble learning in which a complex model is made by combining many simple models.	Liu et al. showed that RF performs better producing more consistent results when using 16S rRNA genes to predict n-caproate and n-caprylate productivities.	Liu et al. (2022), Marcos-Zambrano et al. (2021)
Multiple decision trees	Decision trees are created from set of training examples for which the class labels are known and they used to classify previously unseen examples.	Wang et al. used decision trees in RF and gradient boosting models.	Wang and Liu (2020), Kingsford and Salzberg (2008)
Gradient boosting (GB)	Gradient boosting: a process of ensemble modeling by averaging weak predictions from decision trees (learners) of fixed size.	Wang examined the classification performance of RF, XGBoost compared to the elastic net (ENET) and support vector machine (SVM) in 29 benchmark human microbiome datasets.	Wang and Liu (2020), Marcos-Zambrano et al. (2021), Zhou and Gallins (2019)
Unsupervised Learning Methods			
Clustering	Aims to group datasets with shared attributes building a hierarchy of nested clusters in order to extrapolate algorithmic relationships based on different metrics. In microbiome data analysis, clustering is often used to identify naturally occurring clusters, which can then be assessed for associations with characteristics of interest.	Shi et al.: systematically compared methods for clustering microbiome observations from four published studies with either geographical or seasonal variables as the true cluster label, which enables biological interpretation of the group separation.	Shi et al. (2022), Marcos-Zambrano et al. (2021)
Nonnegative matrix factorization (NMF)	Model extracts hidden patterns from a series of high-dimensional vectors or nonnegative datasets.	Ko et al.: developed a network-based method using NMF to identify functional meta-microbial features that better discriminate specific environmental conditions of samples using microbiome data.	Ko et al. (2021), Marcos-Zambrano et al. (2021)

associations of possible microbial interactions with disease states, to unsupervised learning for clustering microbiomes and assessing functional microbial features.

8.6 Advancing Precision Oncology

Given that cancer is the most genetically heterogeneous of all diseases, the key challenge is how to use the knowledge of the molecular composition of the patient/tumor to choose the treatment(s) that provides patients with the highest likelihood of a curative outcome and minimal likelihood of drug resistance and toxicity (Soldatos et al., 2019). Precision oncology seeks to optimize individual patient care by extracting important information from medical big data, especially genomic information to improve the clinical decision-making process enabling clinicians to select optimal treatment protocols that improve outcomes and increase the quality of life of cancer patients (Hamamoto et al., 2020; Lassen et al., 2021). Simply put precision oncology seeks to deliver the right cancer treatment to the right patient at the right dose and at the right time (Schwartzberg et al., 2017). Whereas genomics gives us the understanding of cancer's input codes in a form of genes, the output codes which are proteins are needed to fully capture the informational state of a tumor and provide a more complete and precise picture of how to understand and treat the underlying molecular pathology more especially since it has been demonstrated that even when a targeted drug is a good match for a specific mutation, it is not always effective (Rodriguez et al., 2021). Three major classes of predictive biomarkers are currently driving the emerging practice of precision oncology; the dependency of tumor cells on cancer driver mutations which disrupt the cellular control mechanism leading to unregulated cell growth and survival, secondly the biomarkers that act by influencing the sensitivity of a tumor to immune recognition, e.g. deficient mismatch repair (DMMR), tumor mutational burden (TMB), and microsatellite instability (MSI), thirdly its synthetic lethality resulting from the observation that cell death is more efficiently induced by the simultaneous loss of function of two or more key players in cellular signaling pathways (Soldatos et al., 2019). With the increase in the availability of molecular studies the challenge for physicians is that the more molecular outcomes data is generated around biomarkers, the more difficult it becomes to clinically interpret the mutational profile of a patient. Fortunately, with the computational power of AI and machine learning, the more clinic-molecular outcomes data is generated globally, the easier it will become to analyze an individual patient's profile (Lassen et al., 2021). Nonetheless with the combination of computational characterization and experimental validation, it is possible to narrow down the list of markers and assist precision oncologists to design compact targeted panels (Mirsadeghi et al., 2021).

The microbiome plays a role in tumor development and progression and has a significant impact on the host immune system influencing antitumor immunity and response to treatment. Therefore a more in-depth understanding of the mechanisms of microbiota-mediated immunomodulation and identification of precise immune-

stimulatory and immune-inhibiting bacterial strains or pathways could lead to increased precision in cancer therapeutic approaches (Matson et al., 2021). The human microbiome is as unique and specific to each individual as fingerprints are, it develops in first few years of life before it stabilizes for the rest of life, but is influenced by diet, lifestyle, and environmental factors (Marcos-Zambrano et al., 2021). The microbiome is resilient as it tends to resort to its stable form after insults while at the same time demonstrating a level of plasticity as it can be modified by external factors. It is thus more mutable than human cells (Kashyap et al., 2017). AI through machine learning can efficiently process all the big data from microbiomic signatures, patient clinical, genetic, and laboratory data, environmental lifestyle factors as well as tumor profile to compute algorithms that can predict optimal, individual treatment outcomes.

In pursuit of enhancing precision oncology, oncologists must gain easy access to big data that is generated globally to avoid some of the pitfalls associated with machine learning based on human decisions. Because algorithms that learn from human decisions will also learn human mistakes, such as over-testing and over-diagnosis, failing to notice people who lack access to care, under-testing those who cannot pay, and mirroring race or gender biases (Obermeyer & Lee, 2017).

8.7 Targeting the Microbiome in the Treatment of Cancer

The genomic revolution and precision medicine efforts have allowed human genomic screening to identify a spectrum of germline encoded mutations that lead to cancer allowing individual-specific application of preventive and therapeutic strategies as well as the stratification of patients based on response to treatment and development of adverse events. The microbiome's malleability and its propensity to easy manipulation is particularly appealing for developing personalized targeted therapies by using precision microbiome targeting approaches combined with other patient and population data as outlined in Fig. 8.3 (Kashyap et al., 2017).

8.8 Limitations

The study or characterization of the human microbiome is complex due to a number of reasons: the microbiome is variable between human to human and over time within the same human; abundance of a particular microbe in cancer may be directly linked to cancer development and/or progression or could simply be opportunistic with the cancer micro-environment being favorable to the growth of that microbe; for microbial quantification to be meaningful, it must be coupled with metagenomics analysis. Some of the available computational tools may not be sufficiently powered to accurately analyze data for guiding patient-specific therapy strategy.

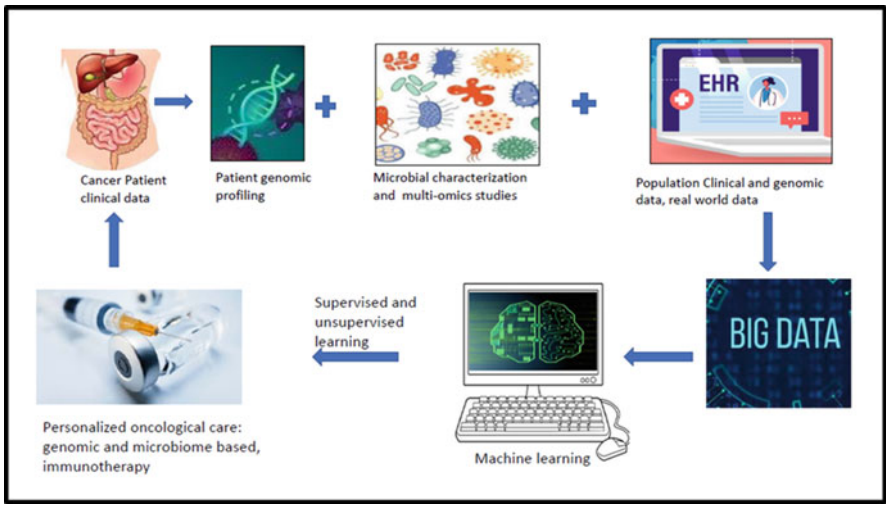


Fig. 8.3 Pathway to precision oncology involving microbiomics: Combination of big data from the real-word data, population data coupled with patients own multi-omics, clinico-laboratory data can be accurately analyzed using computational power of machine learning and AI to advance precision oncology

While microbiome can be targeted for precision medicine, a major limitation is that there is no standardization of methods to develop reliable and reproducible microbiome-based diagnostic and therapeutic strategies. To be able to successfully implement microbiome-based diagnostics and therapeutics, uniform collection, sequencing, and analysis standards that would enhance reproducibility of results across centers and reduce biases in their interpretation need to be developed (Kashyap et al., 2017).

Cancers develop over long time periods, while different micro-organisms contribute to oncogenesis and are relatively abundant in the tumor micro-environment at distinct time points during the neoplastic process. So by the time cancer is detected, the window of opportunity for identifying the inciting microbial agent(s) may have passed, allowing these organisms to remain elusive (Garrett, 2015).

8.9 Conclusions

Disturbances in the natural state of the human microbiome and the prevalence of some microbes in certain body regions play a significant role in the development and progression of cancer acting either directly or indirectly. Cancer is the most heterogeneous disease with no two cancers being the same, much the same way the human microbiome demonstrates the wide diversity from person to person and body region to body region. The microbiome has two major characteristics that make it a suitable

target for precision oncology. Firstly, its uniqueness and specificity to each individual that is as unique as the fingerprints and secondly its malleability and ability to be easily manipulated by external factors like diet, introduction of probiotics to its environment, and pharmaceuticals. The microbiota can be targeted as part of a treatment strategy to successfully manage cancer progression. Given the diversity of both the microbiome and cancer, and recent progress in multi-omics studies it is inevitable that machine learning and AI algorithms must be incorporated as essential tools required for the accurate interpretation of big data generated by microbiomics and cancer genomics to enhance decision-making systems in cancer treatment and advance precision oncology.

References

Allegra, A., Innao, V., Allegra, A. G., Ettari, R., Pugliese, M., Pulvirenti, N., & Musolino, C. (2019). Role of the microbiota in hematologic malignancies. *The Netherlands Journal of Medicine, 77*, 67–80.

Asher, E. E., & Bashan, A. (2022). Model-free prediction of microbiome compositions. *bioRxiv*, 2022.02.04.479107.

Cmrečak, F., Stručić, F., Beketić-Orešković, L., & Andrašek, I. (2021). Microbiome and cancer. *Libri oncologici, 49*, 124–132.

Creasy, H. H., Felix, V., Aluvathingal, J., Crabtree, J., Ifeonu, O., Matsumura, J., Mccracken, C., Nickel, L., Orvis, J., Schor, M., Giglio, M., Mahurkar, A., & White, O. (2021). HMPDACC: A human microbiome project multi-omic data resource. *Nucleic Acids Research, 49*, D734–d742.

De Maria Marchiano, R., Di Sante, G., Piro, G., Carbone, C., Tortora, G., Boldrini, L., Pietragalla, A., Daniele, G., Tredicine, M., Cesario, A., Valentini, V., Gallo, D., Babini, G., D'Oria, M., & Scambia, G. (2021). Translational research in the era of precision medicine: Where we are and where we will go. *Journal of Personalized Medicine, 11*.

Dlamini, Z., Francies, F. Z., Hull, R., & Marima, R. (2020). Artificial intelligence (AI) and big data in cancer and precision oncology. *Computational and Structural Biotechnology Journal, 18*, 2300–2311.

Eguchi, H., Carpentier, S., Kim, S. S., & Moss, S. F. (2004). P27kip1 regulates the apoptotic response of gastric epithelial cells to Helicobacter pylori. *Gut, 53*, 797–804.

Gaitanis, M. L. A. G. (2020). Cancer microbiomatics? In: A, K. M. A. V. (Ed.), *Translational and applied genomics series microbiomics: Dimensions, applications, and translational implications of human and environmental microbiome research.* London: Academic Press.

Garrett, W. S. (2015). Cancer and the microbiota. *Science, 348*, 80–86.

Gilbert, J. A., Blaser, M. J., Caporaso, J. G., Jansson, J. K., Lynch, S. V., & Knight, R. (2018). Current understanding of the human microbiome. *Nature Medicine, 24*, 392–400.

Greathouse, K. L., White, J. R., Vargas, A. J., Bliskovsky, V. V., Beck, J. A., Von Muhlinen, N., Polley, E. C., Bowman, E. D., Khan, M. A., Robles, A. I., Cooks, T., Ryan, B. M., Padgett, N., Dzutsev, A. H., Trinchieri, G., Pineda, M. A., Bilke, S., Meltzer, P. S., Hokenstad, A. N., Stickrod, T. M., Walther-Antonio, M. R., Earl, J. P., Mell, J. C., Krol, J. E., Balashov, S. V., Bhat, A. S., Ehrlich, G. D., Valm, A., Deming, C., Conlan, S., Oh, J., Segre, J. A., & Harris, C. C. (2018). Interaction between the microbiome and TP53 in human lung cancer. *Genome Biology, 19*, 123.

Hamady, M., & Knight, R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Research, 19*, 1141–1152.

Hamamoto, R., Suvarna, K., Yamada, M., Kobayashi, K., Shinkai, N., Miyake, M., Takahashi, M., Jinnai, S., Shimoyama, R., Sakai, A., Takasawa, K., Bolatkan, A., Shozu, K., Dozen, A.,

Machino, H., Takahashi, S., Asada, K., Komatsu, M., Sese, J., & Kaneko, S. (2020). Application of artificial intelligence technology in oncology: Towards the establishment of precision medicine. *Cancers (Basel), 12*.

Hayes, R. B., Ahn, J., Fan, X., Peters, B. A., Ma, Y., Yang, L., Agalliu, I., Burk, R. D., Ganly, I., Purdue, M. P., Freedman, N. D., Gapstur, S. M., & Pei, Z. (2018). Association of oral microbiome with risk for incident head and neck squamous cell cancer. *JAMA Oncology, 4*, 358–365.

Hull, R., Lolas, G., Makrogkikas, S., Jensen, L. D., Syrigos, K. N., Evangelou, G., Padayachy, L., Egbor, C., Mehrotra, R., Makhafola, T. J., Oyomno, M., & Dlamini, Z. (2021). Microbiomics in collusion with the nervous system in carcinogenesis: Diagnosis, Pathogenesis and Treatment. *Microorganisms, 9*.

Kambouris, A. V. A. M. (2020). Introduction: The microbiome as a concept: Vogue or necessity? In: A, K. M. A. V. (Ed.), *Translational and applied genomics series microbiomics: Dimensions, applications, and translational implications of human and environmental microbiome research.* London: Academic Press.

Kashyap, P. C., Chia, N., Nelson, H., Segal, E., & Elinav, E. (2017). Microbiome at the frontier of personalized medicine. *Mayo Clinic Proceedings, 92*, 1855–1864.

Kingsford, C., & Salzberg, S. L. (2008). What are decision trees? *Nature Biotechnology, 26*, 1011–1013.

Knippel, R. J., Drewes, J. L., & Sears, C. L. (2021). The cancer microbiome: Recent highlights and knowledge gaps. *Cancer Discovery, 11*, 2378–2395.

Ko, Y. J., Kim, S., Pan, C. H., & Park, K. (2021). Identification of functional microbial modules through network-based analysis of meta-microbial features using matrix factorization. *IEEE/ACM Trans Comput Biol Bioinform*.

Kumar, P. S. (2021). Microbiomics: Were we all wrong before? *Periodontology 2000, 2000*(85), 8–11.

Lassen, U. N., Makaroff, L. E., Stenzinger, A., Italiano, A., Vassal, G., Garcia-Foncillas, J., & Avouac, B. (2021). Precision oncology: A clinical and patient perspective. *Future Oncology, 17*, 3995–4009.

Liu, B., Sträuber, H., Saraiva, J., Harms, H., Silva, S. G., Kasmanas, J. C., Kleinsteuber, S., & Nunes da Rocha, U. (2022). Machine learning-assisted identification of bioindicators predicts medium-chain carboxylate production performance of an anaerobic mixed culture. *Microbiome, 10*, 48.

Lo, C., & Marculescu, R. (2019). MetaNN: Accurate classification of host phenotypes from metagenomic data using neural networks. *BMC Bioinformatics, 20*, 314.

Marcos-Zambrano, L. J., Karaduzovic-Hadziabdic, K., Loncar Turukalo, T., Przymus, P., Trajkovik, V., Aasmets, O., Berland, M., Gruca, A., Hasic, J., Hron, K., Klammsteiner, T., Kolev, M., Lahti, L., Lopes, M. B., Moreno, V., Naskinova, I., Org, E., Paciência, I., Papoutsoglou, G., Shigdel, R., Stres, B., Vilne, B., Yousef, M., Zdravevski, E., Tsamardinos, I., de Santa, C., Pau, E., Claesson, M. J., Moreno-Indias, I., & Truu, J. (2021). Applications of machine learning in human microbiome studies: A review on feature selection, biomarker identification, Disease prediction and treatment. *Front Microbiol, 12*, 634511.

Matson, V., Chervin, C. S., & Gajewski, T. F. (2021). Cancer and the microbiome-influence of the commensal microbiota on cancer, immune responses, and immunotherapy. *Gastroenterology, 160*, 600–613.

Mirsadeghi, L., Haji Hosseini, R., Banaei-Moghaddam, A. M., & Kavousi, K. (2021). EARN: An ensemble machine learning algorithm to predict driver genes in metastatic breast cancer. *BMC Medical Genomics, 14*, 122.

Moss, S. F. (2017). The clinical evidence linking helicobacter pylori to gastric cancer. *Cellular and Molecular Gastroenterology and Hepatology, 3*, 183–191.

Namkung, J. (2020). Machine learning methods for microbiome studies. *Journal of Microbiology, 58*, 206–216.

Nayfach, S., Shi, Z. J., Seshadri, R., Pollard, K. S., & Kyrpides, N. C. (2019). New insights from uncultivated genomes of the global human gut microbiome. *Nature, 568*, 505–510.

Neu, A. T., Allen, E. E., & Roy, K. (2021). Defining and quantifying the core microbiome: Challenges and prospects. *Proceedings of the National Academy of Sciences of the United States of America, 118*.

Obermeyer, Z., & Lee, T. H. (2017). Lost in thought – The limits of the human mind and the future of medicine. *The New England Journal of Medicine, 377*, 1209–1211.

Ponziani, F. R., Bhoori, S., Castelli, C., Putignani, L., Rivoltini, L., Del Chierico, F., Sanguinetti, M., Morelli, D., Paroni Sterbini, F., Petito, V., Reddel, S., Calvani, R., Camisaschi, C., Picca, A., Tuccitto, A., Gasbarrini, A., Pompili, M., & Mazzaferro, V. (2019). Hepatocellular carcinoma is associated with gut microbiota profile and inflammation in nonalcoholic fatty liver disease. *Hepatology, 69*, 107–120.

Putignani, L., Gasbarrini, A., & Dallapiccola, B. (2019). Potential of multiomics technology in precision medicine. *Current Opinion in Gastroenterology, 35*, 491–498.

Rajendhran, J., & Gunasekaran, P. (2010). Human microbiomics. *Indian Journal of Microbiology, 50*, 109–112.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*, 1347–1358.

Ramaboli, M., Nesengani, L., Katsidzira, L., Haller, D., Kinross, J., Ocvirk, S., & O'Keefe, S. J. D. (2022). Interactions between the environmental and human microbiota in the preservation of health and genesis of disease: Symposium report. *Current Opinion in Gastroenterology, 38*, 146–155.

Reiman, D., Metwally, A., & Yang, D. (2017). Using convolutional neural networks to explore the microbiome. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2017*, 4269–4272.

Riquelme, E., Zhang, Y., Zhang, L., Montiel, M., Zoltan, M., Dong, W., Quesada, P., Sahin, I., Chandra, V., San Lucas, A., Scheet, P., Xu, H., Hanash, S. M., Feng, L., Burks, J. K., Do, K. A., Peterson, C. B., Nejman, D., Tzeng, C. D., Kim, M. P., Sears, C. L., Ajami, N., Petrosino, J., Wood, L. D., Maitra, A., Straussman, R., Katz, M., White, J. R., Jenq, R., Wargo, J., & Mcallister, F. (2019). Tumor microbiome diversity and composition influence pancreatic cancer outcomes. *Cell, 178*, 795–806.e12.

Rodriguez, H., Zenklusen, J. C., Staudt, L. M., Doroshow, J. H., & Lowy, D. R. (2021). The next horizon in precision oncology: Proteogenomics to inform cancer diagnosis and treatment. *Cell, 184*, 1661–1670.

Schwartzberg, L., Kim, E. S., Liu, D., & Schrag, D. (2017). Precision oncology: Who, how, what, when, and when not? *American Society of Clinical Oncology Educational Book***,** 160–169.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., & Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. *Genome Biology, 12*, R60.

Shariati, A., Razavi, S., Ghaznavi-Rad, E., Jahanbin, B., Akbari, A., Norzaee, S., & Darban-Sarokhalil, D. (2021). Association between colorectal cancer and Fusobacterium nucleatum and Bacteroides fragilis bacteria in Iranian patients: A preliminary study. *Infect Agent Cancer, 16*, 41.

Shi, Y., Zhang, L., Peterson, C. B., Do, K.-A., & Jenq, R. R. (2022). Performance determinants of unsupervised clustering methods for microbiome data. *Microbiome, 10*, 25.

Shukla, S. K., Murali, N. S., & Brilliant, M. H. (2015). Personalized medicine going precise: From genomics to microbiomics. *Trends in Molecular Medicine, 21*, 461–462.

Soldatos, T. G., Kaduthanam, S. & Jackson, D. B. (2019). Precision oncology-the quest for evidence. *Journal of Personalized Medicine, 9*.

Topçuoğlu, B. D., Lesniak, N. A., Ruffin, M. T. T., Wiens, J., & Schloss, P. D. (2020). A Framework for effective application of machine learning to microbiome-based classification problems. *mBio, 11*.

Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., & Gordon, J. I. (2007). The human microbiome project. *Nature, 449*, 804–810.

Vassilios, F. (2016). *The "-omics" technologies: Metabolomics and microbiomics personalized medicine from the fetus to the adult.* London Academic Press an Imprint of Elsevier.

Vivarelli, S., Salemi, R., Candido, S., Falzone, L., Santagati, M., Stefani, S., Torino, F., Banna, G. L., Tonini, G., & Libra, M. (2019). Gut microbiota and cancer: From pathogenesis to therapy. *Cancers (Basel), 11*.

Wang, X. W., & Liu, Y. Y. (2020). Comparative study of classifiers for human microbiome data. *Medicine in Microecology, 4*.

Werner, J. J., Koren, O., Hugenholtz, P., Desantis, T. Z., Walters, W. A., Caporaso, J. G., Angenent, L. T., Knight, R., & Ley, R. E. (2012). Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. *The ISME Journal, 6*, 94–103.

Wu, J., Li, Q., & Fu, X. (2019). Fusobacterium nucleatum contributes to the carcinogenesis of colorectal cancer by inducing inflammation and suppressing host immunity. *Translational Oncology, 12*, 846–851.

Yang, Y., Misra, B. B., Liang, L., Bi, D., Weng, W., Wu, W., Cai, S., Qin, H., Goel, A., Li, X., & Ma, Y. (2019). Integrated microbiome and metabolome analysis reveals a novel interplay between commensal bacteria and metabolites in colorectal cancer. *Theranostics, 9*, 4101–4114.

Zhou, Y. H., & Gallins, P. (2019). A review and tutorial of machine learning methods for microbiome host trait prediction. *Frontiers in Genetics, 10*, 579.

Part III
Artificial Intelligence in Cancer Therapy and Clinical Applications

Chapter 9
AI and Nanomedicine in Realizing the Goal of Precision Medicine: Tailoring the Best Treatment for Personalized Cancer Treatment

Amanda Skepu, Boitumelo Phakathi, Malose Makgoka, Zukile Mbita, Botle Precious Damane, Demetra Demetriou, and Zodwa Dlamini

Abstract Nanomedicine and precision medicine are modern concepts in the medical field. Nanomedicine is defined as the nanotechnology application in medical practice and is incorporated into diagnosis and treatment to manage different medical conditions. Precision medicine or personalized medicine aims at individualizing treatment for patients to overcome general treatment that works for some patients. For other patients, the treatment is ineffective or may be toxic to the patient. Novel nanomedicine technologies are used in the treatment of various diseases and can be modified to individual patients according to their genetic profiles, however there are still some limitations to these technologies. This chapter will examine the role of

A. Skepu (✉)
Council for Scientific and Industrial Research (CSIR), Next Generation Health Cluster, Pretoria, South Africa
e-mail: Askepu@csir.co.za

B. Phakathi
Department of Surgery, Faculty of Health Sciences, University of Kwa-Zulu Natal, Durban, South Africa
e-mail: PhakathiB@ukzn.ac.za

M. Makgoka · B. P. Damane
Department of Surgery, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: malose.makgoka@up.ac.za; precious.setlai@up.ac.za

Z. Mbita
Department of Biochemistry, Microbiology and Biotechnology, University of Limpopo, Polokwane, South Africa
e-mail: Zukile.Mbita@ul.ac.za

D. Demetriou · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: dd.demetriou@up.ac.za; zodwa.dlamini@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_9

Artificial Intelligence (AI) and nanotechnology integration in realizing the goal of precision medicine.

Keywords Nanomedicine · Nanotechnologies · Artificial Intelligence · Machine Learning · Precision medicine · Treatment

9.1 Introduction

Precision medicine offers tremendous opportunity to revolutionize the future of healthcare, with great strides already being made in the field of cancer. In precision medicine, it is important to understand the relationship between genetic variability, lifestyle and environment, and patient's health. This relationship has been studied extensively in populations of European and Asian ancestry, but there is not much information about African populations. Extensive knowledge and understanding are therefore needed in this area before precision medicine programs for Africans populations can be effectively implemented. Moreover, this diversity amongst patients is greater in different cancer types, further exacerbating the stage at diagnosis and the treatment thereof. Furthermore, precision medicine aims to provide tailored, individualized treatment approach, ensuring that the correct drug at the right dose is given to each patient (Alghamdi et al., 2022).

Over the recent years, nanotechnology has offered innovative solutions for some of the world's most pressing problems, particularly in the health space. An important application of nanotechnology is in nanomedicine, where nanoscale systems are leveraged for their unique properties to improve the diagnosis and treatment of various diseases, predominantly cancer. Nanomedicine is advancing with novel drug development through multifunctional approaches and sources in diagnosis and therapy. Nanomedicine-based drug treatment is investigated at a fixed dosage where the drug action is dose-dependent, time-dependent, and specific to the patient. AI is used to optimize the dose and drug parameters that will overcome the challenges of nanomedicine trials. Real-time adaptation between AI and nanomedicine can improve patient data analysis and nanomaterial design. For instance, an AI-based pattern analysis and algorithm models can further improve the accuracy of diagnosis, prognosis, and therapy. Predictive computational models could help to accelerate the translation of nanomedicine into clinical practice.

9.2 Nanotechnology Solutions in Precision Medicine

Nanotechnology is a growing branch of science that allows the design of tools and devices of various sizes that can specifically function at the atomic, molecular, and cellular levels (Auffan et al., 2009). Nanomedicine is the application of nanotechnology in biomedical research and clinical practice. The properties of nanomaterials include transformable shape and size, high specific surface area, and tunable

chemical reactivity, which make them perfect for improving early diagnosis and treatment of cancer and neurodegenerative diseases. Moreover, these properties allow them to adapt to any microenvironment which makes them ideal as drug-load carriers and as imaging agents.

Nanomedicine has made significant improvements in diagnosis and therapy developments, as evidenced by improved treatment outcomes and contrast efficiency as a result of imaging agents and nanoparticle-modified drug compounds (Ho et al., 2019). Nanomedicine platforms have been incorporated into clinical practice, for example, with the approval for Abraxane® and many other products (Ho et al., 2019). As the field of nanomedicine continues to grow, multifunctional approaches have been investigated that will allow the integration of diagnostic and therapeutic agents into a single treatment solution. These strategies can refine treatment outcome via targeted, multi-agent delivery that maintains drug synergy (Ho et al., 2019).

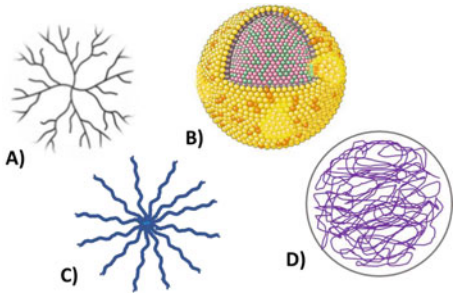
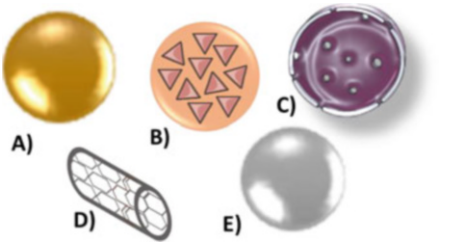
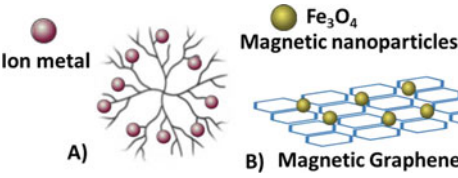
Since the emergence of nanomedicine, an array of nanotechnology-based platforms (including polymer, lipid, natural compound, silicon, metal, and carbon-based platforms) have been investigated for imaging applications and drug delivery (Glinel et al., 2018), for conditions such as cancer, cardiovascular disease, regenerative medicine, and others (Zavaleta et al., 2018; Bowerman et al., 2016). Nanoparticles are classified into organic, non-organic, and composite. In addition to drug delivery systems, these commonly used nanoparticles can also be used to track the intracellular homing of biomolecules. It is important to select the right type of nanoparticle for specific applications as the reliability of the nanoparticle used for medical purposes can depend on the composition and characteristics of the nanomaterial used (Joudeh & Linke, 2022; Gessner & Neundorf, 2020) (Table 9.1).

Nanomedicine has been implicated in the diagnosis, monitoring, prevention, and treatment of diseases. These tools are currently used in current clinical practice (Greish, 2012). Moreover, nanotechnology can be used as a technology to investigate unmet needs, for example with patient-specific and population-specific problems. There are multiple areas where nanomedicine and personalized medicine interlink and these include:

1. Diagnostics, where nanomedicine can be used to investigate pharmacogenetic testing, drug targets, and simultaneous in vitro and in vivo testing (Alghamdi et al., 2022).
2. Therapeutics, where nanomedicine can be tailored for target specificity for each patient improving precision medicine (Ventola, 2012).
3. Targeting capability, where nanomedicine dosages can be tailored based on individual patient conditions (Maeda, 2001).

The intention of nanodrugs was to improve the accuracy and properties of available drugs or diagnostic agents. Current nanodrugs are designed to minimize the side effects of drugs and improve the half-life, bioavailability, and overall pharmacodynamic and pharmacokinetic properties of drugs (Alghamdi et al., 2022). There are several pharmacokinetic advantages that nano-based drugs can offer, as compared to the traditional drug delivery systems, and these include: controlled release possibility, solubility and absorption, drug stability and

Table 9.1 Classification of nanoparticles

Nanoparticles	Examples	Structure
Organic	(A) Dendrimers (B) Liposomes (C) Micelles (D) Hydrogels	
Non-organic	(A) Gold (B) Carbon quantum dots (C) Mesoporous silica (D) Carbon nanotubes (E) Iron oxide	
Composite	(A) Metal (B) Magnetic	

metabolism improvement, improved blood circulation, side effects reduction, and improved targeted delivery (Zhang et al., 2020). Traditionally, the focus of pharmaceutical industries and nanotechnology was on a one-size-fits-all basis, however, nanomedicine is able to personalize the pharmacodynamics and pharmacokinetics, thereby offering personalized therapeutical opportunity (Pereira et al., 2015; Mitchell et al., 2021).

With the advent of nanomedicine, superior strategies of simultaneously functionalizing and co-delivering nanomaterial platforms with multiple therapeutic agents for improved treatment outcomes have emerged (Zhang et al., 2017; Qi et al., 2017; Karp & Peer, 2018). These multilayered approaches have been for treatment administration control, thus allowing efficient delivery-based enhancements (Linden et al., 2016). These capabilities have further allowed the investigation of nanomedicine platforms for combination therapies and co-delivery to address multiple disease ranging from oncology to infectious and regenerative diseases (Elbashir et al., 2001; Tabernero et al., 2013).

9.2.1 *Combining AI and Nanotechnology Solutions in Tailoring the Best Treatment for Cancer Treatment*

Artificial Intelligence is defined as the ability of computers to supplement complex human analysis through machine learning (ML) algorithms (Dlamini et al., 2020). ML algorithms range a host of different complexities and mechanisms including artificial neural networks (ANN) and clustering-based approaches (Deo, 2015). This section explores the synergistic relationship between AI and nanomedicine in the context of cancer precision medicine.

As discussed in the previous section, nanotechnology offers a personalized experience, ranging from tumor diagnosis, targeted drug delivery to organ transplantation, offering healthcare practitioners with novel tools. With the incorporation of AI and ML, these tools will have eyes, hands, and a brain, with the capability of providing advanced personalized care and ability to track the patients' well-being. Figure 9.1 depicts the impact of integrating AI and nanotechnology in personalized medicine.

For targeted therapy, precision medicine relies on probe specificity for a particular molecular target. The development of targeted nano-based therapies or nanodrugs is often complicated by differences in the omic-profile of the patient, lifestyle, and molecular interactions. To improve treatment efficacy and outcomes, reduction of interaction with non-target cells and non-specific binding in a healthy host is needed. Therefore, AI and ML algorithms can predict membrane-bound ligand interaction, surface binding properties and biocompatibility that can improve the field of nanomedicine (Hayat et al., 2021). Currently, there are few published studies

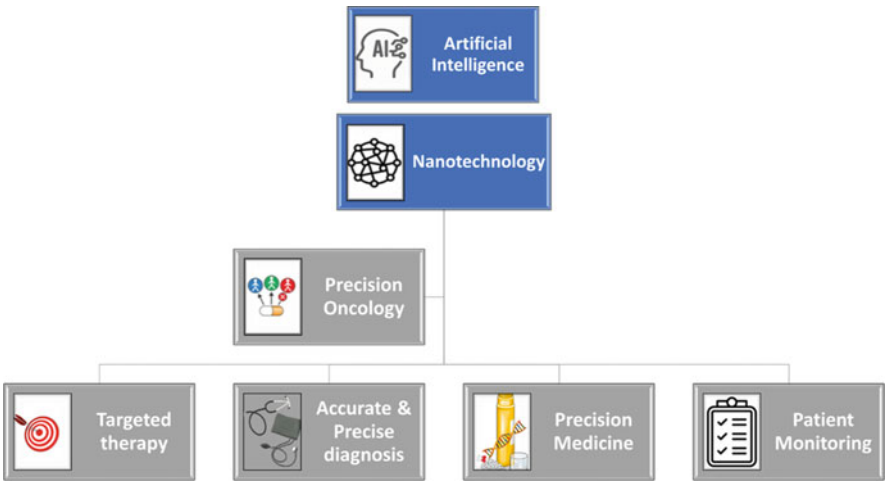


Fig. 9.1 Integration of AI and Nanotechnology for Precision Medicine. Integrated nanomedicine and AI can be exploited or applied across different spectrums within medicine from tailored patient-specific diagnosis, treatment, and patient follow-up

using AI algorithms for nanodrug target selection and genomic screening for optimal design. Examples of specific areas where AI has empowered nanomedicine are highlighted in the next sections.

9.3 Role of AI in Drug Development Optimization

Traditionally, drug development has always relied on novel compound target-based discovery followed by screening and dose studies to determine efficacy and safety. During the screening process, compounds which do not show good efficacy are often eliminated from the screen. However, even if a drug does not show efficacy on its own, it can mediate efficacy when the right drug at the right dose is given. To identify this combination is challenging. AI can play a huge role in enhancing the progress in drug development (Ho et al., 2015; Zarrinpar et al., 2016). This is evident from a study conducted by Weiss et al. (2015), where they assessed anti-angiogenic agents for preclinical ovarian cancer treatment. Drugs that showed no efficacy on its own at specific doses resulted in the decrease in the burden when combined with another drug at the same doses. This showed the role that AI can play in identifying drug combinations that will improve nanomedicine (Weiss et al., 2015, Weiss & Nowak-Sliwinska, 2017).

Optimization of drug combinations can enhance drug synergism to improve cancer treatment efficacy. However, this still remains a challenge as drug optimization requires the right combination of dosage, dosing frequency, and drugs to enhance their efficacy, while decreasing toxicities and side effects. Furthermore, drug combinations may also lead to unexpected toxicities due to biological system complexities. Despite improved therapeutic efficacy through multifunctional nanomedicines, there are still optimization challenges. Combining AI and nanomedicine can help overcome these challenges, thereby improving the efficacy of cancer therapy (Ho et al., 2019).

9.3.1 Role of Artificial Intelligence in Clinical Therapy: Drug Dosing and Therapeutic Efficacy Correlation

It is envisaged that AI will also play a significant role in the optimization of the administration of nanotechnology-unmodified and modified drug combinations. Drug synergy is dose-dependent, time-dependent, and patient-specific, but can change during treatment, resulting in sub-optimal response rates when a fixed-dose treatment is given to a heterogeneous patient population (Ho et al., 2019). Approaches that incorporate big data-driven approaches have been used for drug selection to assist in treatment decision making (Kawamoto et al., 2005; Warken et al., 2018). These strategies can be seen as a crucial initial step toward using

valuable information databases to refine the regimen design process, which may improve broader patient population safety and efficiency (Ho et al., 2019).

In oncology, reducing dose has always been seen as a way to managing or reducing toxicity related to treatment, as opposed to improving drug efficacy in combination therapy. Emerging studies demonstrate that it is feasible to increase drug efficacy while reducing the dosing in combination therapy (Ho et al., 2019). For this to be implemented in practical use, the main hurdle to overcome is to establish a dosing approach that can correlate the correct dose and the best treatment outcome, at a specific time point, the hurdle that can be addressed through AI and ML.

On the other hand, dosing control is not always adequate to personalize the treatment due to the varying response action of patients with different pharmacogenomic profiles. This is where AI can be used to correlate drug dosing and the treatment outcome (Ho et al., 2019). For example, Valdes et al. (2017) developed ANN for constructing personalized radiotherapy treatment for cancer patients according to the treatment goal, the radiation's physical specifications, and the patients' anatomical and physiological parameters. These methods are adapted to predict drug-response relationships based on drug properties, gene expression profiles, and physiological measurements (Linden et al., 2016).

9.3.2 Role of AI in Improved Targeting

Personalized targeting, a concept that was first introduced by Paul Ehrlich in the early 1900s as the "Magic Bullet" theory (Ehrlich, 1960), is one of the advantages of nanomedicine. This technology uses targeted drugs which recognize and activate the disease target site, thereby protecting surrounding healthy tissues. It is done by using a specific ligand such as membrane-bound receptor ligands, antibodies, or other cellular markers to coat the surface of drug-loaded nanoparticles. Despite the promise of targeted nanomedicine, its implementation in the clinic has not yet come to fruition, with only a few formulations currently in clinical studies (Shi et al., 2017). There are several published reviews that are detailing some of the challenges that need to be addressed in order to successfully translate targeted nanomedicine to the clinic (Brannon-Peppas & Blanchette, 2012) and this is where AI can play a role in closing the gaps. Tagging drug-loaded nanoparticles with a targeting moiety does not translate to successful delivery and release at the target site. The effect of the nanomedicine properties on the interactions with the cellular membranes, plasma, and the vasculature endothelium is not easily justified and can be significantly improved with AI. Integration of AI computational modeling in the design of nanoparticles can play a role in increasing the success rate of targeted therapies.

Moreover, computational models can be used in predicting the capability of nanoparticles to cross barriers en route to the target organ. Various model types have been used to predict the ability of nanoparticles to cross the blood–brain barrier and their potential toxicity (Shityakov Roewer et al., 2017). These models can be

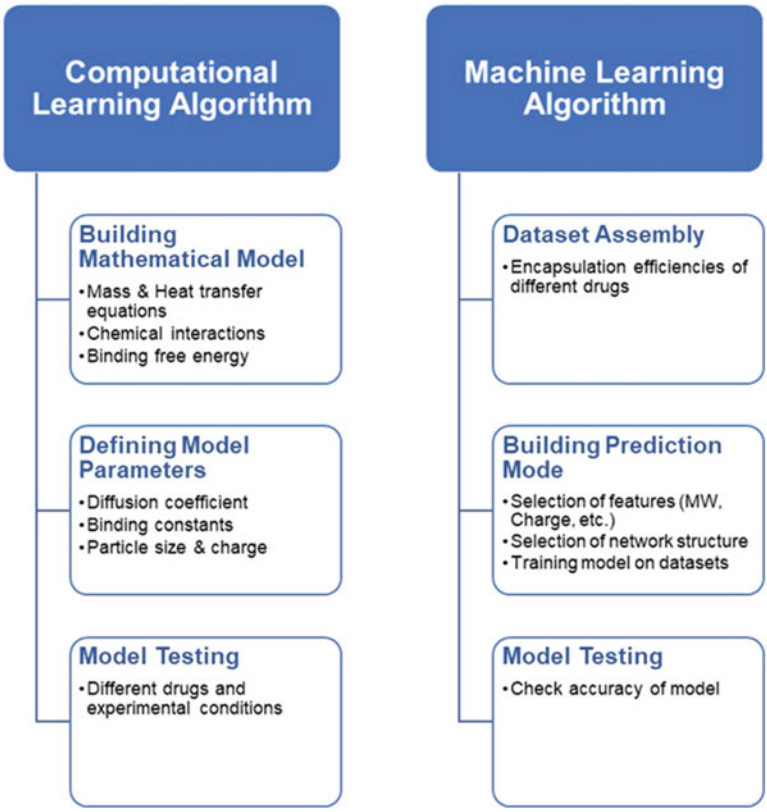


Fig. 9.2 Computational models and machine learning algorithms for prediction of drug encapsulation. Computational models depend on deep understanding of the physical, chemical, and biological processes

used to improve the formulations of brain targeting nanoparticles. However, due to the complexities of the permeation process, these computational models require large computation capabilities. Recently, a ML approach for blood–brain permeability prediction was reported based on the drugs' side effects, indications, and chemical properties (Gao et al., 2016). However, the design of ML methods requires different types of data than used in traditional computational models as shown in Fig. 9.2. For computational models, extensive biochemical and physical knowledge is needed, while, in ML methods, prior understanding is not a prerequisite. The latter utilizes large datasets of experimental results and detects correlations in the data which are then translated into a prediction model.

9.3.3 Role of AI in Gene Therapy

Gene therapy is a form of precision medicine in which patient-specific mutations are inhibited, corrected, or removed. Gene silencing with RNA interference (RNAi) was first demonstrated by Fire and Mello and has been used to target complementary mRNA molecules in cells and lead to their degradation (Elbashir et al., 2001). This mechanism has been explored for precision treatment of various diseases. Tabernero et al. (2013) have shown that silencing of proteases that mediate cell invasion and metastasis, oncogenes, drug resistance genes, and angiogenic factors, has positive therapeutic effects in cancer. However, the utility of RNAi as a precision treatment requires efficient delivery vehicles (Whitehead et al., 2009). These delivery vehicles must be able to cross the endothelial barrier and plasma membranes, avoid phagocytic uptake, release the small interfering RNA (siRNA) in the cytoplasm, and escape the endosome (Kanasty et al., 2013). Various nanoparticles, namely, poly (ethylene imine) polymer nanoparticles, lipid nanoparticles, polysaccharide-based nanoparticles composed of chitosan or cyclodextrin and self-assembled nucleic acid nanoparticles, have been investigated in animal models and were shown to deliver siRNA *in vivo* (Semple et al., 2010). However, their use as delivery tools in cancer remains a challenge, which includes high interstitial fluid pressure in the tumor environment and inter-patient variability in tumor vasculatures. Clinical trials are currently underway using siRNA against different targets in various cancer types (Kim et al., 2016).

Various AI-based models, including neural networks (NN), decision trees, and SVMs were used for classification of ineffective and effective sequences for RNAi in order to recognize the key features in their design (Peek 2007). However, these models do not consider the delivery method but only the efficiency of the siRNA sequence. Instead, laborious experimental scanning of chemical libraries was performed in several studies to evaluate the parameters in the design of carrier systems (Alabi et al., 2013). However, these data can be explored to help discover overlooked design parameters and development of ML algorithms. Moreover, specific modeling of membrane–nanoparticle interactions can provide knowledge of intercellular pathways, particle uptake mechanisms, and the effects of the nanoparticle on these processes (Ding & Ma, 2015). By considering these properties, the transfection efficiency of the nanoparticles can be improved (Adir et al., 2019).

9.4 Challenges with AI Integrated Nanotechnologies

9.4.1 AI-Enabled Nanomedicine

For implementation of precision medicine, computational methods are proving to be key. Studies have demonstrated the value of AI algorithms for screening and classifying patients' drug suitability and for optimizing nanomedicine properties in

precision medicine (Adir et al., 2019). However, for AI algorithms to reach clinical implementation, several challenges must be addressed, including obtaining large datasets that will be used for training the algorithms in order to achieve high accuracy. Furthermore, a multi-disciplinary collaborative approach is required among the experts in the fields of nanomaterials, medicine, and computer science and implementation of computation in all stages of academic and industrial research will help to optimize their performance and clinical relevance (Adir et al., 2019).

9.4.2 Current Nanotechnology Strategies

As compared to conventional therapeutics, nanomedicine boasts several benefits, which include improved dose-response, targeted delivery, precision/personalization, and therapeutic efficacy (Ventola, 2012; Mitchell et al., 2021). However, there are limitations in realizing the applications of nanomedicine into real clinical use. These encompass issues of biocompatibility, insufficient standardized quantification methods for monitoring and analyzing patient response to therapy due to the inability by physicians to perform data analysis (Sanhai et al., 2007). Additionally, the translation of nanomedicine-based therapies is limited by the amounts of genomic data and information required to be decoded when selecting candidate small molecule targets. In order to overcome some of these limitations that inhibit translation of nanomedicine, this is where AI comes in to provide the computational power and throughput necessary for the realization of this technology in theragnostic imaging and cancer nanodrug development (Ho et al., 2019). Some of these approaches are discussed in the following sections of this chapter.

9.5 Conclusion and Perspectives

Despite the successes of nanomedicine-based therapies, nanomedicine is still not yet suitable for every patient or disease including cancer. The key question to ask when it comes to precision medicine treatment is whether nanomedicine can be personalized for every patient. Besides the complicated clinical approval process for personalized nanomedicine, other limitations include the different fabrication techniques and the high costs of nanomedicine development. A combined approach encompassing the use of both precision diagnostic platform and personalized drug-tailoring can improve the patient's treatment outcome. This will improve therapeutic efficacy and overcome drug resistance. AI and other computational models play a key role in the development, design, and implementation of these nanotechnologies (Fig. 9.3).

The development of ML algorithms has fostered extraordinary growth in the application of AI algorithms in nanomedicine. Some future perspectives of nanomedicine with AI integration include simulation and modeling of

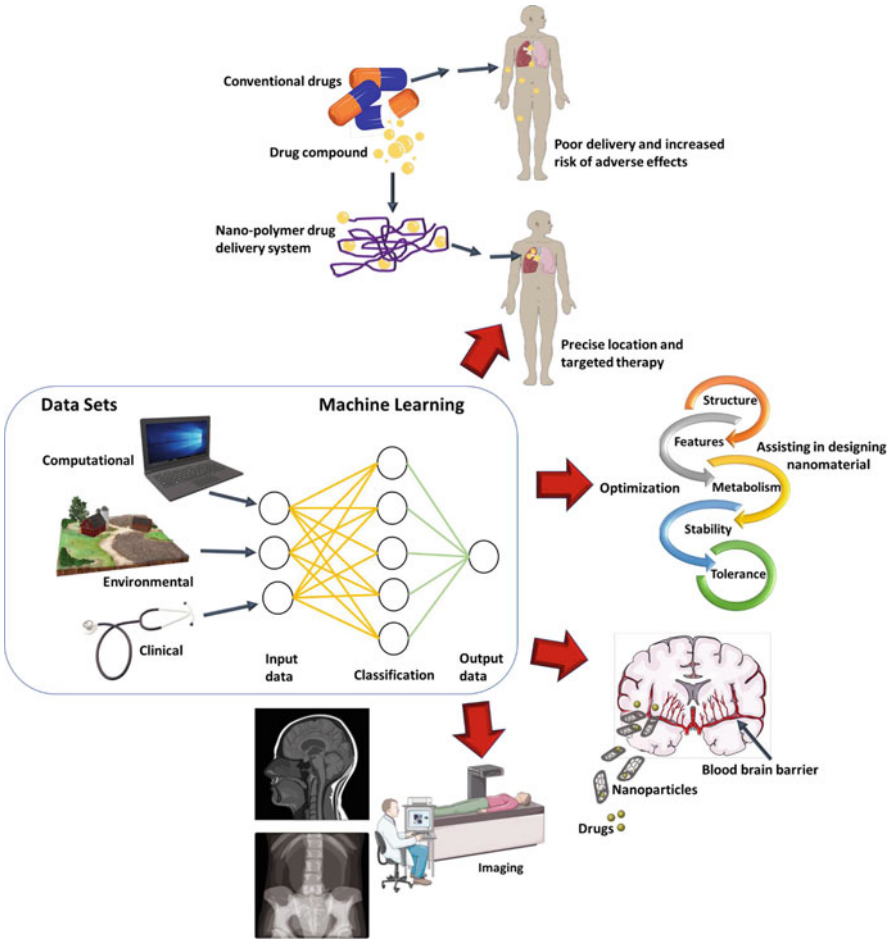


Fig. 9.3 Application of AI technologies for the development of nanomedicines. Machine learning assists with integrating multiple large datasets that would not be possible for the human mind to comprehend. These datasets are processed and used for multiple purposes in nanotechnology. Nanotechnology-developed medical products can be used for the purpose of diagnosis, drug delivery systems for delivery of targeted and site-specific precision medicines, designing of combinatorial therapies, and optimization of nanomaterials

nanotoxicology, optimization of nanodrug doses, and predictions of protein corona formation. A challenge in drug administration is that drug synergy is patient-specific, time-dependent, and dose-dependent at any given point of treatment (Hayat et al., 2021). AI in nanomedicine can assist in overcoming these challenges resulting in the improvement of nanotherapy and precision medicine.

Although AI-based algorithms for nanomedicine can improve care, their effectiveness and safety must still be ensured. Accurate evaluation and optimization of methods should be included in phases of development. Regulatory bodies can be

assigned to address key aspects and to ensure the effectiveness, performance, and safety of ML algorithms at each step (Larson et al., 2021). With all of these efforts, advances in nanomedicine in combination with AI will be a game changer in precision medicine and in the development of healthcare.

References

Adir, O., Poley, M., Chen, G., Froim, S., Krinsky, N., Shklover, J., Shainsky-Roitman, J., Lammers, T., & Schroeder, A. (2019 July). Integrating artificial intelligence and nanotechnology for precision cancer medicine. *Advanced Materials, 32*(13), e1901989.

Alabi, C. A., Love, K. T., Sahay, G., Yin, H., Luly, K. M., Langer, R., & Anderson, D. G. (2013). Multiparametric approach for the evaluation of lipid nanoparticles for siRNA delivery. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 12881.

Alghamdi, M. A., Fallica, A. N., Virzì, N., Kesharwani, P., Pittalà, V., & Greish, K. (2022). The promise of nanotechnology in personalized medicine. *Journal of Personalized Medicine, 12*, 673.

Auffan, M., Rose, J., Bottero, J.-Y., Lowry, G. V., Jolivet, J.-P., & Wiesner, M. R. (2009). Towards a definition of inorganic nanoparticles from an environmental, health and safety perspective. *Journal of Nature Nanotechnology, 4*, 634–641.

Bowerman, C. J., Byrne, J. D., Chu, K. S., Schorzman, A. N., Keeler, A. W., Sherwood, C. A., Perry, J. L., Luft, J. C., Darr, D. B., & Deal, A. M. (2016). Docetaxel-loaded PLGA nanoparticles improve efficacy in taxane-resistant triple-negative breast cancer. *Nano Letters, 17*, 242–248.

Brannon-Peppas, L., & Blanchette, J. O. (2012). Nanoparticle and targeted systems for cancer therapy. *Advanced Drug Delivery Reviews, 64*, 206.

Deo, R. C. (2015). Machine learning in medicine. *Circulation, 132*, 1920–1930.

Ding, H.-M., & Ma, Y.-Q. (2015). Cellular uptake: Theoretical and computational investigations of nanoparticle–biomembrane interactions in cellular delivery. *Small, 11*, 1055.

Dlamini, Z., Francies, F. Z., Hull, R., & Rahaba, M. R. (2020). Artificial intelligence (AI) and big data in cancer and precision oncology. *Computational and Structural Biotechnology, 18*, 2300–2311.

Ehrlich, P. (1960). Experimental researches on specific therapy. In *The collected papers of Paul Ehrlich*. Elsevier, pp. 106–117.

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. T., & Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature, 411*, 494–498.

Gao, Z., Chen, Y., Cai, X., & Xu, R. (2016). Predict drug permeability to blood–brain-barrier from clinical phenotypes: Drug side effects *and* drug indications. *Bioinformatics, 33*, 901.

Gessner, I., & Neundorf, I. (2020). Nanoparticles modified with cell-penetrating peptides: Conjugation mechanisms, physicochemical properties, and application in cancer diagnosis and therapy. *International Journal of Molecular Sciences, 21*.

Glinel, K., Behrens, A., Langer, R. S., Jaklenec, A., & Jonas, A. M. (2018). Nanofibrillar patches of commensal skin bacteria. *Biomacromolecules, 20*(1), 102–108.

Greish, K. (2012). Enhanced permeability and retention effect for selective targeting of anticancer nanomedicine: Are we there yet? *Drug Discovery Today: Technologies, 9*, e161–e166.

Hayat, H., Nukala, A., Nyamira, A., Fan, J., & Wang, P. (2021). A concise review: The synergy between artificial intelligence and biomedical nanomaterials that empowers nanomedicine. *Biomedical Materials, 16*, 052001.

Ho, D., Wang, C.-H. K., & Chow, E. K.-H. (2015). Nanodiamonds: The intersection of nanotechnology, drug development, and personalized medicine. *Science Advances, 1*, e1500439.

Ho, D., Wang, P., & Kee, T. (2019). Artificial intelligence in medicine. *Nanoscale Horizons, 4*, 365–377.

Joudeh, N., & Linke, D. (2022). Nanoparticle classification, physicochemical properties, characterization, and applications: a comprehensive review for biologists. *Journal of Nanobiotechnology, 20*, 262.

Kanasty, R., Dorkin, J. R., Vegas, A., & Anderson, D. (2013). Delivery materials for siRNA therapeutics. *Nature Materials, 12*, 967.

Karp, J. M., & Peer, D. (2018). Focus on RNA interference: From nanoformulations to in vivo delivery. *Nanotechnology, 29*, 4.

Kawamoto, K., Houlihan, C. A., Balas, E. A., et al. (2005). Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *British Medical Journal, 330*(7494), 765–768.

Kim, H. J., Kim, A., Miyata, K., & Kataoka, K. (2016). Recent progress in development of siRNA delivery vehicles for cancer therapy. *Advanced Drug Delivery Reviews, 104*, 61.

Larson, D. B., Harvey, H., Rubin, D. L., Irani, N., Tse, J. R., & Langlotz, C. P. (2021). Regulatory frameworks for development and evaluation of artificial intelligence-based diagnostic imaging algorithms: Summary and recommendations. *Journal of the American College of Radiology, 18*, 413–424.

Linden, A., Yarnold, P. R., & Nallamothu, B. K. (2016). Using machine learning to model dose–response relationships. *Journal of Evaluation in Clinical Practice, 22*, 860.

Maeda, H. (2001). SMANCS and polymer-conjugated macromolecular drugs: Advantages in cancer chemotherapy. *Advanced Drug Delivery Reviews, 46*(1–3), 169–185.

Mitchell, M. J., Billingsley, M. M., Haley, R. M., Wechsler, M. E., Peppas, N. A., & Langer, R. (2021). Engineering precision nanoparticles for drug delivery. *Nature Reviews. Drug Discovery, 20*, 101–124.

Peek, A. S. (2007). Improving model predictions for RNA interference activities that use support vector machine regression by combining and filtering features. *BMC Bioinformatics, 8*, 182.

Pereira, M. C., Reshetnyak, Y. K., & Andreev, O. A. (2015). Advanced targeted nanomedicine. *Journal of Biotechnology, 202*, 88–97.

Qi, S. S., Sun, J. H., Yu, H. H., & Yu, S. Q. (2017). Co-delivery nanoparticles of anti-cancer drugs for improving chemotherapy efficacy. *Drug Delivery, 24*, 1909–1926.

Sanhai, W. R., Spiegel, J., & Ferrari, M. (2007). A critical path approach to advance nanoengineered medical products. *Drug Discovery Today: Technologies, 4*, 35–41.

Semple, S. C., Akinc, A., Chen, J., Sandhu, A. P., Mui, B. L., Cho, C. K., Sah, D. W., Stebbing, D., Crosley, E. J., & Yaworski, E. (2010). Rational design of cationic lipids for siRNA delivery. *Nature Biotechnology, 28*, 172.

Shi, J., Kantoff, P. W., Wooster, R., & Farokhzad, O. C. (2017). Cancer nanomedicine: Progress, challenges and opportunities. *Nature Reviews. Cancer, 17*, 20.

Shityakov Roewer, N., Broscheit, J. A., & Förster, C. (2017). In silico models for nanotoxicity evaluation and prediction at the blood-brain barrier level: A mini-review. *Computational Toxicology, 2*, 20.

Tabernero, J., Shapiro, G. I., LoRusso, P. M., Cervantes, A., Schwartz, K. G., Weiss, J., Paz-Ares, L., Cho, D. C., Infante, J. R., Alsina, M., Gounder, M. M., Falzone, R., Harrop, J., White, A. C. S., Toudjarska, I., Bumcrot, D., Meyers, R. E., Hinkle, G., Svrzikapa, N., Hutabarat, R. M., Clausen, V. A., Cehelsky, J., Nochur, S. V., Gamba-Vitalo, C., Vaishnaw, A. K., Sah, D. W. Y., Gollob, J. A., & Burris, H. A. (2013). First-in-humans trial of an RNA interference therapeutic targeting VEGF and KSP in cancer patients with liver involvement. *Cancer Discovery, 3*, 406.

Valdes, G., Simone, C. B., 2nd, Chen, J., Lin, A., Yom, S. S., Pattison, A. J., Carpenter, C. M., & Solberg, T. D. (2017). Clinical decision support of radiotherapy treatment planning: A data-driven machine learning strategy for patient-specific dosimetric decision making. *Radiotherapy and Oncology, 125*, 392.

Ventola, C. L. (2012). The nanomedicine revolution: Part 1: Emerging concepts. *Pharmacology & Therapeutics, 37*, 512.

Warken, ZN, Smyth, H.D.C., Davis, D.A., Weitman, S, Kuhn, J.G. and Williamm II, R.O. (2018). Personalized medicine in nasal delivery: The use of patient-specific administration parameters to improve nasal drug targeting using 3D-printed nasal replica casts. Molecular Pharmaceutics 15(4).

Weiss, A., Berndsen, R. H., Ding, X., Ho, C. M., Dyson, P. J., van den Bergh, H., Griffioen, A. W., & Nowak-Sliwinska, P. (2015). A streamlined search technology for identification of synergistic drug combinations. *Scientific Reports, 5*, 14508.

Weiss, A., & Nowak-Sliwinska, P. (2017). Current trends in multidrug optimization: An alley of future successful treatment of complex disorders. *SLAS Technology, 22*, 254–275.

Whitehead, K. A., Langer, R., & Anderson, D. G. (2009). Knocking down barriers: Advances in siRNA delivery. *Nature Reviews Drug Discovery, 8*.

Zarrinpar, A., Lee, D.-K., Silva, A., Datta, N., Kee, T., Eriksen, C., Weigle, K., Agopian, V., Kaldas, F., & Farmer, D. (2016). Engineering precision biomaterials for personalized medicine Sci. *Translational Medicine, 8*, 333ra349.

Zavaleta, C., Ho, D., & Chung, E. J. (2018). Special issue: Quantitative imaging in medicine and the life sciences. *SLAS Technology, 23*, 281–293.

Zhang, D., Kong, Y. Y., Sun, J. H., Huo, S. J., Zhou, M., Gui, Y. L., Mu, X., Chen, H., Yu, S. Q., & Xu, Q. D. (2017). Co-delivery nanoparticles with characteristics of intracellular precision release drugs for overcoming multidrug resistance. *International Journal of Nanomedicine, 12*, 2081–2108.

Zhang, A., Meng, K., Liu, Y., Pan, Y., Qu, W., Chen, D., & Xie, S. (2020). Absorption, distribution, metabolism, and excretion of nanocarriers in vivo and their influences. *Advances in Colloid and Interface Science, 284*, 102261.

Chapter 10
Artificial Intelligence-Based Medical Devices Revolution in Cancer Screening: Impact into Clinical Practice

Ahmed Bhayat, Rodney Hull, Nkhensani Chauke-Malinga, and Zodwa Dlamini

Abstract AI-based medical devices promise to revolutionise medical imaging tools and lead to advancement in the diagnosing and management of cancer patients. These devices are by their very nature a form of personalised medicine as they analyse data recorded from individual patients. They may take the form of an actual recording device with AI built-in or merely as a software that can be loaded onto a variety of devices or even operated remotely from the cloud. There are a variety of innovative medical equipment that have been developed recently using Artificial Intelligence (AI) algorithms. These devices aim not only at improving the diagnosis of cancer but also to improve the treatment of a variety of different cancers. The most common cancers that have AI testing devices include lung, breast, central nervous system, and prostate cancers. The AI devices that play a role in the management of these cancers have shown to be able to diagnose, characterise, and image tumours to ensure early treatment and appropriate management. AI devices are also able to be used for cancer screening through endoscopies radiology and medical imaging analysis. While the technology faces problems in the form of regulation, inflated costs, ethical considerations, and a lack of trust in the devices, its promise cannot be ignored.

A. Bhayat (✉)
Department of Community Dentistry, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
e-mail: ahmed.bhayat@up.ac.za

R. Hull · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: rodney.hull@up.ac.za; zodwa.dlamini@up.ac.za

N. Chauke-Malinga
Department of Plastic and Reconstructive Surgery, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: chauke-malinga@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_10

Keywords Artificial Intelligence · Cancer · Oncology · Computer-assisted diagnosis · Personalised medicine · Medical imaging

10.1 Introduction

There is an emerging role for medical devices that use artificial intelligence to screen for and diagnose various cancers in patients. This would allow for more rapid and precise treatments. The most frequently used AI devices were to manage and treat cancers in the breast, lung, and prostate. Many of these devices have already been approved and are currently being used across the globe. The FDA has also approved numerous AI devices and the list of devices continues to grow. The advantage of many of these devices is that they increase access to early cancer detection, prevent invasive surgical treatment options, and indirectly reduce the cost of management of these cancers by reducing surgical intervention. Clinically, the reduction of hospital stays because of the use of these devices reducing surgical procedures assists patients and improves diagnosis due to AI mapping which provides a more detailed diagnosis and hence treatment options. A study by Luchini et al. (2022) reviewed diagnostic tools approved by the FDA for diagnosing cancer (Luchini et al., 2022). This review reported that breast, lung, and prostate cancers had the largest number of AI diagnostic tools. Many of these devices have already been approved. A full list of the devices approved by the US FDA[1] is attached in Appendix I. Hence, there are a variety of AI diagnostic tools available, and more research is required on other AI tools in other types of cancers.

This chapter will discuss the use of these artificial intelligence-based medical devices (AIMDs) in precision oncology. It will cover a brief history of these devices, followed by a discussion of their uses, the basis for their function, their regulation, and finally the challenges and limitations faced by these devices.

10.2 The Definition and Characteristics of an AI Device

A device can be classified as an AU-based device in several ways. At face value, the term is assigned based on the announcements by regulatory agents such as the FDA, or through communications from the company producing the device. More importantly, the device is defined as an AI-based device if it uses machine or deep learning as well as the algorithms used in the operating of the device. Aside from actual physical devices that are used in medical applications or can be implanted in patients, the term medical devices can also cover software as a medical device (SaMD), medical big data common data model (CDM), and digital therapeutics (DTx). SaMDs are just AI-based software that can function as a medical tool and does not require a physical device (Kawtrakul & Praneetpolgrang, 2014). CDMs are tools and programs that are used to standardise and curate medical data and DTx devices are meant to replace existing drugs device (Kawtrakul & Praneetpolgrang,

2014). Software that is a medical device can be divided into those that operate on their own or embedded software which is run off a specific hardware device known as software in a medical device (SiMD) (Moor, 2006).

In addition to these groups, a subset of AI-based medical devices which assist medical professionals in making diagnoses or selecting the appropriate treatment is known as a clinical decision support system (CDSS). These devices typically analyse medical images (Kawtrakul & Praneetpolgrang, 2014). Some of the newer devices even learn and contain algorithms to take the patient's wishes into account, since shared decision-making is beneficial (Nilsson & Nilsson, 1998). These devices accomplish this by using a patient decision support system (PDSS) to assist in making decisions (Kawtrakul & Praneetpolgrang, 2014). Broadly speaking, physical AIMDs can be classified as embedded or independent, depending on whether they are implanted. The AIMD can then be classified by function control, measurement, analysis, diagnosis, data conversion, transmission, reception, and finally display (Kawtrakul & Praneetpolgrang, 2014).

10.3 History of Artificial Intelligence (AI) Devices

At first medical devices were merely capable of making measurements as well as displaying and recording the results. The incorporation of AI has transformed medical devices into advanced tools capable of diagnosing and determining the risk of developing cancer. These devices' names vary worldwide, but this review will use the term artificial intelligence medical devices (AIMDs) (Luger, 2005). As of early 2020 more than 60 medical devices with AI have been approved by the Food and Drug Administration (FDA) in the USA. Many of these devices have been designed for cancer and the majority of them are in the radiology and diagnostic fields (Hamamoto et al., 2020a). The term AI has been used in medicine since around 1950. Studies in different countries have shown that AI has been developing over the past few decades and this renewed interest has helped to overcome one of the challenges faced by researchers, which was the lack of funding (Kawtrakul & Praneetpolgrang, 2014; Hamamoto et al., 2020a). The history of AI can be classified into three time periods: the Birth of AI (First AI Boom), the second AI boom, and the third AI Boom (the era of deep learning) (Hamamoto et al., 2020a).

The first AI Boom started when computers were becoming accessible, and scientists realised their value in medicine. The name "Artificial Intelligence (AI)" was coined in the mid-fifties at a workshop in the States and this initiated the academic field of AI (Moor, 2006). Following this, computer programs were developed albeit the programs were slow and expensive. The costs of testing and implementing these programs were also steep and as a result, funding started decreasing for research to continue. The second AI boom started around the 1980s when companies around the world started investing and developing programs that could answer questions and solve problems using logical rules obtained from expert knowledge (Nilsson & Nilsson, 1998; Luger, 2005). Examples of these included

computers which could identify compounds from a spectrometer and the ability to diagnose infectious diseases in the blood (Lindsay et al., 1980; Shortliffe, 1976). The interest in AI started declining again and by the 1990s again reached challenges with funding. The third AI boom and the era of Deep Learning began in the mid-2000s, AI once again attracted attention through the invention of the computer that was able to perform actions without the assistance of humans (Hinton & Salakhutdinov, 2006). Then in 2010, the term "big data" was proposed due to the explosion of the internet and its multiple uses in data collection, storage, and analysis (Hinton & Salakhutdinov, 2006). Since then AI has been extensively used in medicine through various image analyses (Asada et al., 2020; Weisberg et al., 2020) and tumour screening abilities (Blanc et al., 2020; Jiang et al., 2021).

The earliest medical devices can be said to date back to ancient times with devices such as forceps, knives, scalpels, saws, lancets, needles, and trocars. When discussing radiology and oncology, the earliest medical device was a linear accelerator, which was used to treat cancer. Following this, most of the medical devices consisted of imaging technologies such as CT, MRI, PET mammograms, ultrasound, and endoscopy. These devices allowed for better diagnosis screening and monitoring of cancer. However, it was not until the application of software algorithms to design and manage radiation treatment plans that the first software-based devices were developed for cancer management (Benjamens et al., 2020).

In the 1950s with the first boom in AI, it was thought that clinical expert systems could replace physicians through the creation of a "doctor in a box". At that time the available computing power was insufficient to deal with the large amounts of data required for this. Rule-based expert systems were developed in the 1970s and these were used in simple non-oncology medical devices that were mere impersonations of AIMDs. With the third boom in AI starting in the early 200 s the next logical step in the revolution of medical devices was the addition of intelligence to a medical device. This proved to can be extremely successful (Jaakkola et al., 2019).

10.4 The Basis of AIMDs

AIMDs must be able to answer questions and analyse data in the same way as a human being but faster, without bias and with a high degree of repeatability. Several types of technologies were developed as AI advanced towards a point where the use of AI in medical devices became viable. The basis of these technologies are the decision support systems that they use. These are decision-making systems that try to emulate human decision-making. These terms do not describe the precise algorithm used but the basis of the selection or area the algorithm works within. There are a variety of these systems. Different types of thought-emulating technologies are found on various AIMDs (Gurupur & Wan, 2020). The first requirement for an AIMD is a decision support system (DSS), a computer program that compiles and organises raw data to identify problems that require solutions and prepare the data for further analysis by specific algorithms. The diverse types of DSS are shown in

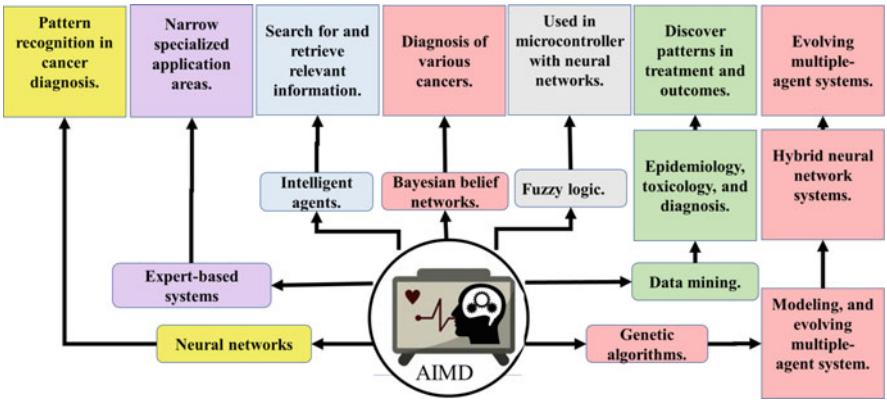


Fig. 10.1 Decision Support systems used by AIMDs and their applications: The above schematic represents the decision support systems that can be used by AIMDs to compile and analyse raw data and identify problems that need to be solved and thereby optimise later decision-making performed by specific algorithms. In general, the rule-based DSSs, both Bayesian and expert based, are only applied to narrow applications as rule construction and writing can become too complex

Fig. 10.1 with details of their use in medicine and oncology. The initial expert systems involved the use of experts in the field to program computers with the knowledge to answer sets of then questions using set rules (Herasevich et al., 2013). These were followed by statistical probability systems also known as Bayesian belief networks. In this process, the computer makes decisions based on statistical analyses and computes probabilities to find the answer to a problem (Maragoudakis et al., 2008). Although these rule-based systems can be used in AIMDs, this technology has the drawback that as the scenarios they were applied to become more complex the system of rules governing them became too unwieldy. Neural networks use small computational units or nodes. These nodes then interact with each other feeding each other the outputs of the analysis they carried out. In this way, the device replicates the thought processes of the human brain (Bhambhvani et al., 2021). The data mining technique involves the use of specific search algorithms to scan large sets of data for patterns. It looks for specific patterns that can be matched with the data it has received from the patient (Wang et al., 2022). The intelligent agent or multiple agent system uses software that is organised into a network of discrete units that act independently that perform autonomous tasks. Genetic algorithms mimic evolution by using natural selection to select those results that are most likely or are preferable in some other way. It continues to work only with these results in further analysis and eliminates the others. Finally, fuzzy logic where the truth of any test result is assigned a value between 0 and 1. These partial truths are dealt with by a superset of conventional logic to select the true outcome.

The DSS is the framework the AI is built upon; however, the actual work is done by the various types of algorithms which an AI can use to conduct data analysis. These include support vector machine (SVM), neural networks (NN), naïve Bayes (NB), K-nearest neighbour (KNN), decision tree (DT), random forest (RF), and

logistic regression (Manickam et al., 2022). SVMs are one of the most used algorithms in medicine. They work by separating data into discrete groups and then finding the midpoint between the two groups of data. This hyperplane or line separating the data is the furthest from the data points. Data falling on each side of the plane is classified appropriately. This algorithm is commonly used for cancer diagnosis due to its accuracy and speed. The algorithm can also easily be scaled to fit higher dimensional data (Sweilam et al., 2010). NNs are the other most used algorithm in medicine. These are efficient, fast, and flexible algorithms. They also do not require specific rules to produce a result and are capable of multitasking (Manickam et al., 2022). These algorithms are also used as DSSs and are discussed above.

10.5 The Practical Use of AI Devices in Cancers

In medicine, AIMDs are most applied to Radiology and Cardiology. These two medical specialities have the most FDA-approved AIMDs. AIMDs have been designed for use in fields as diverse as internal medicine/endocrinology, neurology, ophthalmology, emergency medicine, and of the greatest interest to us, oncology (Benjamens et al., 2020). Currently, AI is used in a variety of settings and in oncology it is used for the detection, characterisation, and monitoring of tumours. AI has been used extensively in the diagnosis of lung, breast, central nervous system, and prostate cancers (Tables 10.1 and 10.2) (Bi et al., 2019). In terms of staging, AI has been used in the staging and diagnosis of various cancers and hence indirectly in

Table 10.1 General use in cancer monitoring

Name of the device	Description of the device and its role	Reference
Kaiku Health (Kaiku Oy)	Outcome monitoring and symptom tracking.	Schmalz et al. (2020)
C the Signs (C the Signs Ltd.)	Assessment of symptoms to support cancer diagnosis.	BA and Bakshi (2021)
Hot Spot APP (Visiopharm A/S)	Hotspot scoring method for various cancer applications.	Hida et al. (2020)
Invasive Tumour Detection APP (Visiopharm A/S)	Distinguish non-invasive and invasive tumours using cytokeratin and p63 markers.	Hida et al. (2020)
SubtlePET (Subtle Medical)	Increased speed and safety scanning exams.	Xu et al. (2020)
RayCare 2.3 (RaySearch Laboratories)	Management of oncology care and follow-up.	Bhalla and Laganà (2022)
FoundationOne Liquid CDx (Foundation Medicine, Inc.)	Diagnostic NGS-based test using circulating cell-free DNA (cfDNA) from cancer patient plasma. Used for treatment with targeted therapies.	Woodhouse et al. (2020)

This table lists the AIMDs that are used for general cancer diagnosis, monitoring, and treatment

Table 10.2 AI-based devices for specific cancers

Name of the device:	Description of the device and its role	Reference
Breast		
ER APP, Breast Cancer (Visiopharm A/S)	Determination of oestrogen receptor expression.	Stålhammar et al. (2016)
PR APP, Breast Cancer (Visiopharm A/S)	Determination of progesterone expression.	Kårsnäs et al. (2015)
QuantX (Quantitative Insights)	Diagnosis.	Scientific (2017)
densitasAI (Densitas Inc.)	Breast density assessment.	Tan et al. (2020)
cmTriage (CureMetrix)	Triage for mammography.	Tartar et al. (2021)
Her2 dual DNA probe cocktail.	Determine HER2 gene amplification status.	Probe (n.d.)
Cervical and endometrial cancer		
AVEC (Automated Visual Evaluation of the Cervix) (MobileODT Ltd.)	Cervical cancer screening and diagnosis.	Xue et al. (2020)
CINtec PLUS cytology (Ventana Medical Systems, Inc.)	Diagnosis using a qualitative p16INK4a and Ki-67 immunocytochemical assay in cervical specimens.	Uijterwaal et al. (2015)
VENTANA MMR RxDx Panel (Ventana Medical Systems, Inc.)	Identifying endometrial cancer patients who may benefit from treatment with dostarlimab-gxly.	Goodpaster (n.d.)
Skin		
DERM (Skin Analytics Ltd.)	Skin cancer diagnosis support.	Phillips and Greenhalgh (2020)
Lymphoma		
Cobas® EZH2 Mutation Test (Roche Molecular System, Inc.)	Identifies follicular lymphoma patients that have an EZH2 mutation and can be treated with TAZVERIK (tazemetostat).	Okosun et al. (2019)
Prostate		
DeepDx-Prostate Connect (Deep Bio Inc.)	Recognition of acinar adenocarcinoma.	Ryu et al. (2019)
Paige Prostate (Paige Inc.)	Diagnosis using prostate needle biopsies.	Raciti et al. (2020)
AI-Pathway Companion Prostate Cancer (Siemens Healthcare GmbH (parent company: Siemens AG)	Diagnostic support.	Henkel et al. (2022)
Galen Prostate (Ibex Medical Analytics Ltd.)	Diagnosis using prostate core needle biopsies.	Pantanowitz et al. (2020)

This table lists AUMDs used to manage specific cancers and by the type of cancer they are used for

the management of these cancers (Bi et al., 2019). A list of cancer monitoring tests and devices is given below:

10.5.1 Radiology and the Analysis of Images for Pathology

Medical images obtained using imaging techniques such as X-rays, CT scans, MRI, PEY, or ultrasound are extremely valuable as a means of diagnosis, screening, disease, and disease monitoring. This makes this form of medical data vital for clinical applications. Imaging also lends itself a wall to analysis and interpretation using AI. As such radiologists and associated applications were one of the first to make use of the applications of digital medicine and AI devices (Pesapane et al., 2018). The diagnosing of pathological lesions is a major concern in many countries including South Africa, where there is a shortage of pathologists (Yoshizawa, 2013). AI is used in Computer-Assisted Diagnosis (CAD) when analysing diagnostic images (Qin et al., 2018). CAD has improved the diagnostic accuracy and reproducibility of image reading and reduced the reading time (Hizukuri et al., 2021). Given this background, the use of AI in diagnosing and drafting pathological reports will go a long way in dealing with the shortage of trained qualified human pathologists and radiologists. Digital pathology has also been useful in the management, diagnosis, and grading of genitourinary cancers as described in Fig. 10.1 (Pai et al., 2020).

AIMDs can use medical image analysis to assist in the diagnosis of cancer. They use variations of a convolutional neural network (CNN)-based deep learning algorithm. The CNN analyses the image by producing a large number of small images, each with a single feature. These features are mapped and the final classification is based on this map (Papa et al., 2020). The ability of AI to detect objects relies on the CNN trained using many images. The ground truth, reference, for the analysis consists of images that have been annotated by experts. The trained AI can then detect and pinpoint objects. AIMDs use three basic methods to analyse medical images. Those are classification, detection, and segmentation. In the classification process, the medical image to a predefined category. In the detection process, features are identified in medical images as being a tumour.

The algorithms used in the detection process are a more advanced convoluted neural networks (CNNs) called faster regions with convolutional neural networks (fR-CNN). These neural networks learn using a region proposal network (RPN). They then extract a region of interest (ROI). Another algorithm that can be used in the decision process is a regression technique called You Only Look Once (YOLO). Finally, the segmentation process involves identifying structures or features within the ROI (Woods, 2007). The use of AINDs to analyse medical images has been applied to multiple different cancers and imaging technologies. These include breast, prostate, lung, brain, and cervical cancer amongst others. Prostate needle core biopsy is the gold standard for the detection of prostate cancer. This involves pathologists analysing the images obtained from the biopsies and assigning a Gleason score, but

the number of tumour-positive cores can give a better idea of the extent and malignancy of cancer. This involves time and resources which may be limited. Additionally, it has been reported that a blinded re-examination of slides improves cancer detection rates (Varma et al., 2018). AI is capable of accurately detecting prostate cancer from digital whole slide images (WSIs) (Litjens et al., 2016). A device named the Paige Prostate Alpha is an AIMD that can detect prostate cancer accurately and more efficiently using core needle biopsies. The authors of this study suggested that due to the speed and accuracy of this device it could be used to carry out the blinded re-examination of slides previously viewed by a pathologist (Raciti et al., 2020).

10.5.2 Endoscopy

Various types of equipment have been developed for the detection of colorectal cancer (Table 10.3). These have improved the diagnosis and early detection of cancers and reduced the number of missed cases (Yamada et al., 2019). A study reported an overall sensitivity and specificity of the AI for colorectal lesions to be 98.8 and 99%, respectively (Yamada et al., 2019). Full blood counts (TBCs), tests to establish the size and number of various blood cells, can be used to diagnose CRC

Table 10.3 AI-based devices used in colorectal cancer

Name of the device	Description of the device and its role	Reference
ColonFlag (Medial EarlySign Inc.)	Pre-symptomatic high-risk patient screening.	Ayling et al. (2021)
GI-Genius (Medtronic Inc. (parent company: Medtronic plc.))	Detection.	Hassan et al. (2020)
Discovery AI (Pentax Medical GmbH (parent company: Pentax Corporation)	Polyp detection.	Boese et al. (2022)
Metastasis Detection App (Visiopharm A/S)	Detects metastasis in lymph nodes for colorectal and breast adenocarcinoma.	Thagaard (2017)
CAD EYE (FUJIFILM Europe GmbH)	Polyp detection and characterisation using colonoscopy.	(Fitting et al., 2022)
NaviCam Capsule Endoscope System with NaviCam Stomach Capsule (AnX Robotica, Inc.)	A magnetically manoeuvred capsule endoscopy system.	Cave et al. (2022)
GI-Genius (Cosmo Artificial Intelligence—AI, LTD)	Detecting colonic mucosal lesions in real time using white-light endoscopy.	Gowda et al. (2021)

This table lists AIMDs used for the diagnosis, management, and monitoring of colorectal cancer, mainly through the use of endoscopy

(Reed et al., 2020). The AIMD ColonFlag (Medial EarlySign, Kfar Malal, Israel) detects FBC changes and converts these into risk scores. It uses AI algorithms, in the form of decision trees, to take the patient's age, sex, and FBC into account to detect CRC. This is based on FBC changes being present before CRC symptoms are noticeable (Kinar et al., 2016). Colonoscopies are currently the gold standard for cRC-diagnosis and screening, but they often miss neoplastic lesions. There is also variability in the interpretation of the resulting images due to human bias (Zhao et al., 2019). Both these problems can be eliminated using AI to interpret the images. The AI can be programmed to recognise and detect polyps on colonoscopy images The GI-Genius, Medtronic AI system was taught using video images obtained from a high-definition white-light colonoscopy of histologically confirmed polyps. It was trained with a large dataset of 1.5 million images in combination with reports associated with these images made by expert endoscopists. In a study on the abilities of this device, it was found that the device performed well or better than 5 expert endoscopists (Hassan et al., 2020) (Table 10.4).

The actual use of an AIMD (Fig. 10.2) involves data acquisition by the device, or data previously acquired being entered into the device. The AI then uses data reduction techniques to simplify the data while only losing the minimum amount of information. This is done to reduce data dimensionality and complexity to allow for further analysis or to reduce the size of the data (de Hond et al., 2022). The decision support system of the device then uses a DSS such as to organise and compile the data to prepare it for further analysis. For example, a data mining DSS compares the patients data against a libraray of medical data and searches for patterns in the patient data that can be matched or intepreted using the stored data or other DSS such as statistical Bayesian analysis, expert systems and neural networks as discussed above (Manickam et al., 2022). The DSS may identify problems which may require further data from the medical practitioner using the device. The request for more information may also allow the device to refine its algorithms. The onboard decision or feature identification algorithms will then analyse the data using a variety of different algorithms. The results of this analysis can then be used by the device to teach itself through a machine or deep learning to improve its algorithms for future use (de Hond et al., 2022).

10.6 The Regulation of AI-based Devices

Due to the unique position occupied by medical devices and AI-based medical devices in medicine, the regulation and approval of these devices have become complicated as no regulatory organisation initially had guidelines in place that could be successfully applied to them. The previous FDA regulations for medical device licensing are strict and these also applied to AIMDs, which involve rigorous testing processes that are long and expensive. These regulations and testing procedures were a major obstacle for the development and use of AIMDs in medicine (Benjamens et al., 2020). Both the Food and Drug Administration (FDA) in the USA and the

Table 10.4 AI-based devices used in imaging

Name of the device	Description of the device and its role	Reference
ClearRead CT (Riverain Technologies LLC.)	Analyse multi-slice CT scans of the chest to detect nodules.	Singh et al. (2021)
Veye Chest (Aidence BV)	Detection of pulmonary nodule using CT scans.	Murchison et al. (2022)
AmCAD-US (AmCad BioMed Corporation)	Quantify ultrasound image data.	Hamamoto et al. (2020b)
DLCExpert (Mirada Medical Ltd.)	Assist in contouring radiation therapy from CT scans.	Min Seo et al. (n.d.)
Arterys MICA (Arterys)	An AI-based platform for analysing medical images such as MRI and CT.	Borgers (2021)
AmCAD-UT (AmCad BioMed Corporation)	Assistance in the analysis of thyroid ultrasound images.	Reverter et al. (2019)
Arterys Oncology DL (Arterys Inc.)	A cloud-based medical imaging AI that measures and tracks lesions and nodules using MRI and CT scans. Can also be used to confirm the absence or presence of lesions.	Wang et al. (2019)
Deep Learning Image Reconstruction (GE Medical Systems)	A deep-learning-based CT image reconstruction technology.	Greffier et al. (2020)
Paige Insight (Paige Inc.)	Digital pathology viewer for diagnosis.	Raciti et al. (2020)
InferRead CT Lung (Beijing Infervision Technology Co. Ltd.)	A tool that uses CT scans for lung cancer screening and management.	Li et al. (2021)
Broncholab (Fluidda Inc.)	Support in diagnosis and documenting pulmonary tissue images using CT scans.	Wang et al. (2022)
JPC-01 K (JLK Inspection Inc.)	Detection for diagnostic support from MRI images.	Turkbey and Haider (2022)
Transpara (ScreenPoint Medical BV)	Interpretation of mammograms for screening.	Sasaki et al. (2020)
QVCAD (QView Medical Inc.)	Aid to detect mammography-occult lesions.	Xu et al. (2018)
HealthMammo (Zebra Medical Vision Inc.)	Analyse mammograms.	Hu and Giger (2021)
Mia-Mammography Intelligent Assessment (KheironMedical Technologies Ltd.)	Breast cancer detection using mammograms.	Harvey et al. (2019)
Syngo-Breast Care (Siemens Healthcare GmbH (parent company: Siemens AG)	Diagnosis using mammograms.	Baptist et al. (2017)
ProFound AI for 2D Mammography (iCAD Inc.)	Detection from 2D mammograms.	Stephens (2021)
ProFound AI for Digital Breast Tomosynthesis (iCAD Inc.)	Detection and diagnosis using digital breast tomosynthesis (DBT) exams.	Overman (2022)
Breast-SlimView (Hera-MI SAS)	Diagnosis by analysing mammograms.	Vijayalakshmi et al. (2021)

(continued)

Table 10.4 (continued)

Name of the device	Description of the device and its role	Reference
Vara (Merantix Healthcare GmbH)	Screening support and triaging from mammograms.	Gassner and Juknat (2019)
Transpara (ScreenPoint Medical)	Interpretation of mammograms for screening.	Sasaki et al. (2020)
ProFound AI Software V2.1 (iCAD)	Interpretation of DBT.	Lyell et al. (2021)
JBD-01 K (JLK Inspection Inc.)	Diagnosis using mammograms.	Al-Tam and Narangale (2021)
b-box (b-rayZ GmbH)	Assess mammography image quality and breast density fusing mammograms.	Steinwendner (2020)
Genius AI Detection (Hologic, Inc.)	Identify potential abnormalities in breast tomosynthesis images.	Hamilton-Basich (2020)
Visage Breast Density (Visage Imaging)	Screening and diagnosis using full-field digital mammography.	Liu et al. (2022)
Imagio Breast Imaging System (Seno Medical Instruments, Inc.)	Improved classification of breast masses using ultrasound.	Overman (2022)
MammoScreen (Therapixel SA)	Screening and diagnosis using FFDM.	Dang et al. (2022)
Breast cancer image, computer-aided detection/diagnosis software	Diagnosis of breast cancer based on mammography.	Retson and Eghtedari (2020)
SmartTarget (SmartTarget Ltd.)	Image-guided procedures for intervention against and diagnosis.	Hamid et al. (2019)
QyScore software (Qynapse SAS)	Labelling, visualisation, and volumetric quantification of brain structures and lesions using MR images.	Cavedo et al. (2022)
LungQ (Thirona Corp)	Diagnosis and documentation of pulmonary tissue images.	Sadeghi et al. (2021)
Syngo.CT Lung CAD (Siemens Medical Solutions Inc. (parent company: Siemens AG))	Detects solid pulmonary nodules using multi-detector computed tomography examinations of the chest.	Poulter (2022)
Auto Lung Nodule Detection (Samsung Electronics Co. Ltd. (parent company: Samsung Group)	Detection and diagnosis of lung nodule using X-ray images.	Cha et al. (2019)

This table lists the AINDs used to analyse medical images for screening, diagnosis, prognosis, and monitoring purposes

European Medicines Agency (EMA) in the EU have attempted to regulate these devices and approval of these devices for clinical use, by developing new rules and guidelines (Food and Administration, 2019). The importance of regulating the use and approval of these devices is due to them being considered high-risk and innovative technology. The high-risk definition is based on the integral role they would play in diagnosis and treatment and any error in the device could easily result

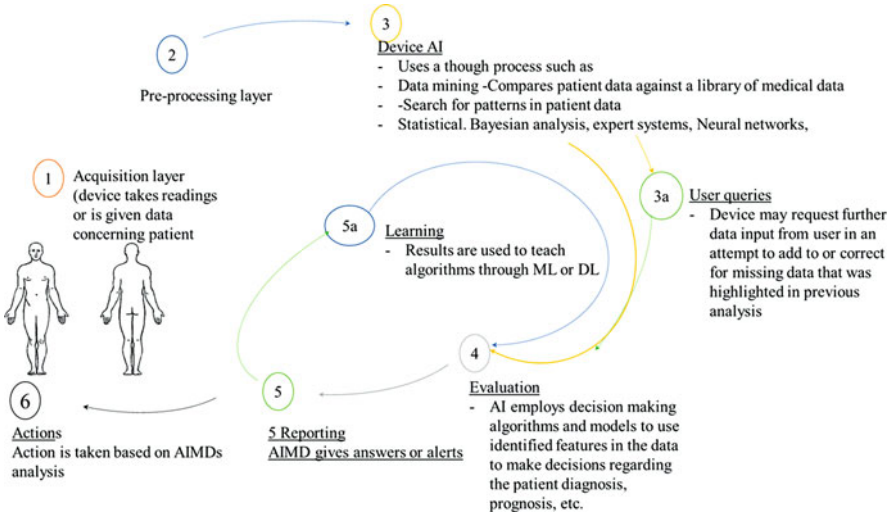


Fig. 10.2 An example of the workflow using a medical device. (1) In the acquisition step the medical device either takes readings or is given data concerning the patient (2) In the pre-processing step, the data is filtered, duplicates are removed, and complexity is reduced minimising loss of data. (3) The data is entered into the decision support structure of the device where it is compiled and organised, and if any problems are identified, this may lead to (3a) the device demanding more information to be entered by the user. (4) The device implements its built-in decision and analysis algorithms to answer the question which in this case is to offer a diagnosis, prognosis or management of a specific condition. The results of this analysis may be used to teach the algorithm (4a or generate a report (5) where the recommended action is conducted (6).

in fatal consequences for the patient. The fact that it is a new technology means that the consequences of using AI/ML to make medical decisions are unknown (Benjamens et al., 2020). This is complicated by how the algorithms update so frequently that any approval process would need to be continuously updated, which is probably not viable (Food and Administration, 2019).

Worldwide steps have been taken to create laws to regulate these devices. In the USA, the Twenty-First Century Cures Act was enacted in 2016. This acts as the approval process for these devices. A few months later the FDA approved the Digital Health Innovation Plan to improve the efficiency of regulation of digital technologies. This plan also includes the software pre-cert pilot program which allows the software to skip the medical device approval process (Lee & Kesselheim, 2018). The European Union (EU) introduced new directives and regulations between 2021 and 2022. These state that an IAMD is classified as a medical device and falls under the medical device classification system. The EU further revised the General Data Protection Regulation (GDPR) to allow the use of IAMDs for diagnosis and management of health conditions and the collection of medical data (Chance & Review, 2018).

10.7 Drawbacks and Limitations of AI Devices

Despite all the promises that AIMDs hold for advancing medicine and by their very nature precision medicine, there are many problems facing their implementation. As discussed above, the implementation of rules and regulations governing their use and application has yet to be properly developed (Food and Administration, 2019). Related to this are some of the ethical issues surrounding the use of these devices. For example, if an error occurs due to the failure of a device or incorrect conclusions reached by the AI, who is accountable? Blame could be placed on the operator, the manufacturer, or a regulatory body. This becomes even more complex once AI starts making autonomous decisions. This may currently limit their use as support tools (Pesapane et al., 2018). In addition to this, the cost of these devices is inherently high mostly due to their development costs. Unfortunately, this means that poorer under-resourced countries will have problems accessing them, maintaining them, and ensuring that staff are trained in using them.

Apart from ethical and financial considerations, the adoption of AI-based medical devices faces further problems such as problems of trust and problems involving real-world data resources. The trust issue comes down to the black box problem which is due to a lack of transparency. This is due to the actual function of the device not being understood, this is further complicated when learning algorithms alter the algorithms as the device is used and learns (Steinwendner, 2020). Another problem relates to the information given to a device to learn from. AI requires huge real-world data sets and this will take time to create and develop. If there is an inherent bias in the data, then the results given by the device will show a bias (Steinwendner, 2020). For instance, if a device is used only with patients of a specific ethnic group, then the results may be inaccurate when the device is used on other population groups.

Many of the problems facing the use of AIMDs can be solved through more specific and specialised regulations being in place. For instance, specialist software regulations need to be developed and implemented, as currently, most regulations focus only on hardware issues. Quality control regulations for AIMDs are also lacking. Security and privacy risks and concerns are also problems facing AIMDs. The communication function of AIMDs means that information can be stolen and cyberattacks can cause the device to malfunction. More prospective studies should be conducted to increase the amount of data that is available for inputs into AI systems.

10.8 Conclusion and Future Perspectives

It is well established that AI has a key role to play in precision oncology, but its introduction into medical devices is a recent development in cancer management. These devices can be considered by their very nature to be personalised medicine as they work with data gathered from individual patients. At the same time, they promise to advance a variety of fields in cancer management. These range from

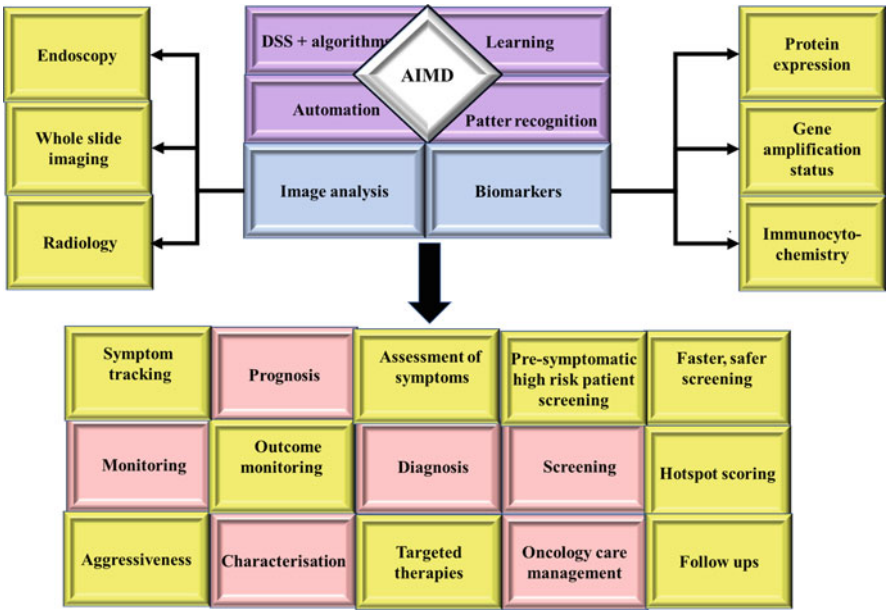


Fig. 10.3 Summary of the applications and underlying basis of AIMDs. AIMDs rely on decision support systems and algorithms to automate the process of pattern recognition and use machine and deep learning to improve these underlying systems with each data set analysed. These can be said to be the principles that guide the development of these devices. From their ability to learn to their basic role as pattern recognition machines no matter what form that data takes to their goal to be fully automated medical devices with minimal or no human interference. The pink blocks in the figure represent key roles that AIMDs play in precision oncology. The yellow boxes represent the key tasks the devices need to perform to accomplish these goals

screening (such as radiology, image analysis, and endoscopy) to diagnosis and disease monitoring (Fig. 10.3). They can accomplish these tasks more accurately and with less bias in a fraction of the time that human pathologists can manage.

The ability of these devices to use machine learning to improve their abilities makes them autonomous in that they do not require external upgrades and as they learn they may be able to operate independently from a human operator. However, they do face challenges such as the cost, regulations, ethical considerations, and a lack of trust. These problems are, however, not insurmountable. Each can be solved through the development of proper guidelines, which may become easier as the technology develops and the medical community gets more familiar with it. The development of the technology and the ability to build new devices based on technology developed for and based on the previous generation of devices will in all likelihood lead to a decrease in the cost of these devices.

References

Al-Tam, R. M., & Narangale, S. M. (2021). Breast cancer detection and diagnosis using machine learning: A survey. *Journal of Scientific Research, 65*.

Asada, K., Kobayashi, K., Joutard, S., Tubaki, M., Takahashi, S., Takasawa, K., Komatsu, M., Kaneko, S., Sese, J., & Hamamoto, R. (2020). Uncovering prognosis-related genes and pathways by multi-omics analysis in lung cancer. *Biomolecules, 10*, 524.

Ayling, R. M., Wong, A., & Cotter, F. (2021). Use of ColonFlag score for prioritisation of endoscopy in colorectal cancer. *BMJ Open Gastroenterology, 8*, e000639.

Ba, M. J. S., & Bakshi, B. (2021). Accelerating early identification of cancer in primary care using an artificial intelligence driven solution.

Baptist, J., Dixit, S., Busheri, L., Krishnan, L., Alhat, R., Verghese, B., Koppiker, C., & Development. (2017). Evaluating the performance of an advanced breast cancer diagnosis unit in India. *Indian Journal of Public Health Research, 8*.

Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: An online database. *NPJ Digital Medicine, 3*, 118.

Bhalla, S., & Laganà, A. (2022). Artificial intelligence for precision oncology. *Computational methods for precision oncology.* Springer.

Bhambhvani, H. P., Zamora, A., Shkolyar, E., Prado, K., Greenberg, D. R., Kasman, A. M., Liao, J., Shah, S., Srinivas, S., Skinner, E. C., & Shah, J. B. (2021). Development of robust artificial neural networks for prediction of 5-year survival in bladder cancer. *Urologic Oncology, 39*, 193. e7–193.e12.

Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., Allison, T., Arnaout, O., Abbosh, C., & Dunn, I. F. (2019). Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA: A Cancer Journal for Clinicians, 69*, 127–157.

Blanc, D., Racine, V., Khalil, A., Deloche, M., Broyelle, J.-A., Hammouamri, I., Sinitambirivoutin, E., Fiammante, M., Verdier, E., & Besson, T. (2020). Artificial intelligence solution to classify pulmonary nodules on CT. *Diagnostic and Interventional Imaging, 101*, 803–810.

Boese, A., Wex, C., Croner, R., Liehr, U. B., Wendler, J. J., Weigt, J., Walles, T., Vorwerk, U., Lohmann, C. H., & Friebe, M. (2022). Endoscopic imaging technology today. *J Diagnostics, 12*, 1262.

Borgers, L. (2021). The role of Artificial Intelligence (AI) in radiology: The current ctatus of FDA approved systems.

Cave, D. R., Gladney, D., Rinner, M., Foley, A., & Houghton, J. M. (2022). Initial experience with the use of a tether with the navicam video capsule system. *J Gastrointestinal Endoscopy, 95*, AB380–AB381.

Cavedo, E., Tran, P., Thoprakarn, U., Martini, J.-B., Movschin, A., Delmaire, C., Gariel, F., Heidelberg, D., Pyatigorskaya, N., & Ströer, S. (2022). Validation of an automatic tool for the rapid measurement of brain atrophy and white matter hyperintensity: QyScore®. *Journal of European Radiology, 32*, 2949–2961.

Cha, M. J., Chung, M. J., Lee, J. H., & Lee, K. S. (2019). Performance of deep learning model in detecting operable lung cancer with chest radiographs. *Journal of Thoracic Imaging, 34*, 86–91.

Chance, C. J. L., & Review, F. M. (2018). EU regulatory developments. *12*, 228–236.

Dang, L.-A., Chazard, E., Poncelet, E., Serb, T., Rusu, A., Pauwels, X., Parsy, C., Poclet, T., Cauliez, H., & Engelaere, C. (2022). Impact of artificial intelligence in breast cancer screening with mammography. *Breast Cancer*, 1–11.

De Hond, A. A. H., Leeuwenberg, A. M., Hooft, L., Kant, I. M. J., Nijman, S. W. J., Van Os, H. J. A., Aardoom, J. J., Debray, T. P. A., Schuit, E., Van Smeden, M., Reitsma, J. B., Steyerberg, E. W., Chavannes, N. H., & Moons, K. G. M. (2022). Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: A scoping review. *NPJ Digital Medicine, 5*, 2.

Fitting, D., Krenzer, A., Troya, J., Banck, M., Sudarevic, B., Brand, M., Böck, W., Zoller, W. G., Rösch, T., & Puppe, F. (2022). A video based benchmark data set (ENDOTEST) to evaluate computer-aided polyp detection systems. *Scandinavian Journal of Gastroenterology*, 1–7.

Food & Administration, D. (2019). Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD).

Gassner, U. M., & Juknat, U. (2019). Regulatory approaches to AI in medical practice. *EPLR, 3*, 176.

Goodpaster, E. (n.d.). Roche receives FDA approval for first companion diagnostic to identify dMMR solid tumour patients eligible for anti-PD-1 immunotherapy.

Gowda, V., Kwaramba, T., Hanemann, C., Garcia, J. A., & Barata, P. C. (2021). Artificial intelligence in cancer care: Legal and regulatory dimensions. *The Oncologist, 26*, 807–810.

Greffier, J., Hamard, A., Pereira, F., Barrau, C., Pasquier, H., Beregi, J. P., & Frandon, J. (2020). Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: A phantom study. *Journal of European Radiology, 30*, 3951–3959.

Gurupur, V., & Wan, T. T. H. (2020). Inherent bias in artificial intelligence-based decision support systems for healthcare. *Medicina (Kaunas), 56*.

Hamamoto, R., Suvarna, K., Yamada, M., Kobayashi, K., Shinkai, N., Miyake, M., Takahashi, M., Jinnai, S., Shimoyama, R., & Sakai, A. (2020a). Application of artificial intelligence technology in oncology: Towards the establishment of precision medicine. *Cancers, 12*, 3532.

Hamamoto, R., Suvarna, K., Yamada, M., Kobayashi, K., Shinkai, N., Miyake, M., Takahashi, M., Jinnai, S., Shimoyama, R., & Sakai, A. (2020b). Application of artificial intelligence technology in oncology: Towards the establishment of precision medicine. *Journal of Cancers, 12*, 3532.

Hamid, S., Donaldson, I. A., Hu, Y., Rodell, R., Villarini, B., Bonmati, E., Tranter, P., Punwani, S., Sidhu, H. S., & Willis, S. (2019). The SmartTarget biopsy trial: A prospective, within-person randomised, blinded trial comparing the accuracy of visual-registration and magnetic resonance imaging/ultrasound image-fusion targeted biopsies for prostate cancer risk stratification. *Journal of European Urology, 75*, 733–740.

Hamilton-Basich, M. (2020). Hologic receives FDA clearance for genius AI detection. *AXIS Imaging News*.

Harvey, H., Karpati, E., Khara, G., Korkinof, D., Ng, A., Austin, C., Rijken, T., & Kecskemethy, P. (2019). The role of deep learning in breast screening. *Current Breast Cancer Reports, 11*, 17–22.

Hassan, C., Wallace, M. B., Sharma, P., Maselli, R., Craviotto, V., Spadaccini, M., & Repici, A. (2020). New artificial intelligence system: First validation study versus experienced endoscopists for colorectal polyp detection. *Journal of Gut, 69*, 799–800.

Henkel, M., Horn, T., Leboutte, F., Trotsenko, P., Dugas, S. G., Sutter, S. U., Ficht, G., Engesser, C., Matthias, M., & Stalder, A. (2022). Initial experience with AI pathway companion: Evaluation of dashboard-enhanced clinical decision making in prostate cancer screening. *PLoS One, 17*, e0271183.

Herasevich, V., Kor, D. J., Subramanian, A., & Pickering, B. W. (2013). Connecting the dots: Rule-based decision support systems in the modern EMR era. *Journal of Clinical Monitoring and Computing, 27*, 443–448.

Hida, A. I., Omanovic, D., Pedersen, L., Oshiro, Y., Ogura, T., Nomura, T., Kurebayashi, J., Kanomata, N., & Moriya, T. (2020). Automated assessment of Ki-67 in breast cancer: The utility of digital image analysis using virtual triple staining and whole slide imaging. *Journal of Histopathology, 77*, 471–480.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science, 313*, 504–507.

Hizukuri, A., Nakayama, R., Nara, M., Suzuki, M., & Namba, K. (2021). Computer-aided diagnosis scheme for distinguishing between benign and malignant masses on breast DCE-MRI images using deep convolutional neural network with Bayesian optimization. *Journal of Digital Imaging, 34*, 116–123.

Hu, Q., & Giger, M. L. J. R. C. (2021). Clinical artificial intelligence applications. *Breast Imaging, 59*, 1027–1043.

Jaakkola, H., Henno, J., Mäkelä, J., & Thalheim, B. (2019). Artificial intelligence yesterday, today and tomorrow. 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). *IEEE*, 860–867.

Jiang, Y., Edwards, A. V., & Newstead, G. M. (2021). Artificial intelligence applied to breast MRI for improved diagnosis. *Radiology, 298*, 38–46.

Kårsnäs, A., Strand, R., Doré, J., Ebstrup, T., Lippert, M., Bjerrum, K. J. C. M. I. B., Imaging, B. E. Visualization. (2015). A histopathological tool for quantification of biomarkers with sub-cellular resolution. *3*, 25–46.

Kawtrakul, A., & Praneetpolgrang, P. (2014). A history of AI research and development in Thailand: Three periods, three directions. *AI Magazine, 35*, 83–92.

Kinar, Y., Kalkstein, N., Akiva, P., Levin, B., Half, E. E., Goldshtein, I., Chodick, G., & Shalev, V. (2016). Development and validation of a predictive model for detection of colorectal cancer in primary care by analysis of complete blood counts: A binational retrospective study. *Journal of the American Medical Informatics Association, 23*, 879–890.

Lee, T. T., & Kesselheim, A. S. J. A. O. I. M. (2018). US Food and Drug Administration precertification pilot program for digital health software: Weighing the benefits and risks. *168*, 730–732.

Li, K., Liu, K., Zhong, Y., Liang, M., Qin, P., Li, H., Zhang, R., Li, S., Liu, X., & Surgery. (2021). Assessing the predictive accuracy of lung cancer, metastases, and benign lesions using an artificial intelligence-driven computer aided diagnosis system. *Journal of Quantitative Imaging in Medicine, 11*, 3629.

Lindsay, R. K., Feigenbaum, E. A., Buchanan, B., & Lederberg, J. (1980). *Applications of artificial intelligence for chemical inference: The DENDRAL project*. McGraw-Hill.

Litjens, G., Sánchez, C. I., Timofeeva, N., Hermsen, M., Nagtegaal, I., Kovacs, I., Hulsbergen-Van de Kaa, C., Bult, P., Van Ginneken, B., & Van der Laak, J. (2016). Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Scientific Reports, 6*, 26286.

Liu, A., Yin, L., Ma, Y., Han, P., Wu, Y., Wu, Y., & Ye, Z. (2022). Quantitative breast density measurement based on three-dimensional images: A study on cone-beam breast computed tomography. *Acta Radiologica, 63*, 1023–1031.

Luchini, C., Pea, A., & Scarpa, A. (2022). Artificial intelligence in oncology: Current applications and future perspectives. *British Journal of Cancer, 126*, 4–9.

Luger, G. F. (2005). *Artificial intelligence: Structures and strategies for complex problem solving*. Pearson Education.

Lyell, D., Coiera, E., Chen, J., Shah, P., Magrabi, F., & Informatics, C. (2021). How machine learning is embedded to support clinician decision making: An analysis of FDA-approved medical devices. *BMJ Health, 28*.

Manickam, P., Mariappan, S. A., Murugesan, S. M., Hansda, S., Kaushik, A., Shinde, R., & Thipperudraswamy, S. P. (2022). Artificial Intelligence (AI) and Internet of Medical Things (IoMT) Assisted Biomedical Systems for Intelligent Healthcare. *Biosensors (Basel), 12*.

Maragoudakis, M., Lymberopoulos, D., Fakotakis, N., & Spiropoulos, K. (2008). A hierarchical, ontology-driven Bayesian concept for ubiquitous medical environments—A case study for pulmonary diseases. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008*, 3807–3810.

Min Seo, C., Jaehee, C., Nalee Kim, J. S. C., & Jinsung, K. (n.d.). Comparison of deep learning-based auto-segmentation algorithms for head and neck cancer.

Moor, J. J. A. M. (2006). The Dartmouth College artificial intelligence conference: The next fifty years. *AI Magazine, 27*, 87–87.

Murchison, J. T., Ritchie, G., Senyszak, D., Nijwening, J. H., Van Veenendaal, G., Wakkie, J., & Van Beek, E. J. (2022). Validation of a deep learning computer aided system for CT based lung nodule detection, classification, and growth rate estimation in a routine clinical population. *PloS One, 17*, e0266799.

Nilsson, N. J., & Nilsson, N. J. (1998). *Artificial intelligence: A new synthesis*. Morgan Kaufmann.

Okosun, J., Bödör, C., Batlevi, C., Nagy, N., Michot, J., Schneider, T., Alizadeh, H., Simon, Z., Vose, J., & Younes, A. (2019). EZH2 gain-of-function mutations are not associated with more favorable prognosis in relapsed/refractory follicular lymphoma (FL): A preliminary analysis on 590 patients. *Journal of Hematological Oncology, 37*, 192–193.

Overman, D. (2022). New research finds iCAD's ProFound AI risk for digital breast tomosynthesis is 2.4 x more accurate than traditional lifetime risk models. *AXIS Imaging News*.

Pai, R. K., Van Booven, D. J., Parmar, M., Lokeshwar, S. D., Shah, K., Ramasamy, R., & Arora, H. (2020). A review of current advancements and limitations of artificial intelligence in genitourinary cancers. *American Journal of Clinical and Experimental Urology, 8*, 152.

Pantanowitz, L., Quiroga-Garza, G. M., Bien, L., Heled, R., Laifenfeld, D., Linhart, C., Sandbank, J., Shach, A. A., Shalev, V., & Vecsler, M. (2020). An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: A blinded clinical validation and deployment study. *The Lancet Digital Health, 2*, e407–e416.

Papa, A., Mital, M., Pisano, P., Del Giudice, M. J. T. F., & Change, S. (2020). E-health and wellbeing monitoring using smart healthcare devices: An empirical investigation. *153*, 119226.

Pesapane, F., Codari, M., & Sardanelli, F. (2018). Artificial intelligence in medical imaging: Threat or opportunity? Radiologists again at the forefront of innovation in medicine. *European Radiology Experimental, 2*, 35.

Pesapane, F., Volonté, C., Codari, M., & Sardanelli, F. (2018). Artificial intelligence as a medical device in radiology: Ethical and regulatory issues in Europe and the United States. *Insights Into Imaging, 9*, 745–753.

Phillips, M., & Greenhalgh, J. (2020). Assessment of the change in accuracy of an artificial intelligence algorithm for the detection of skin cancer in camera images following diversification and training.

Poulter, R. (2022). Evaluation of the diagnostic performance of quantified dual energy cardiac computed tomography perfusion imaging in stable coronary artery disease.

Probe, V. (n.d.) Ventana HER2 Dual ISH DNA Probe Cocktail.

Qin, C., Yao, D., Shi, Y., & Song, Z. (2018). Computer-aided detection in chest radiography based on artificial intelligence: A survey. *Biomedical Engineering Online, 17*, 1–23.

Raciti, P., Sue, J., Ceballos, R., Godrich, R., Kunz, J. D., Kapur, S., Reuter, V., Grady, L., Kanan, C., & Klimstra, D. S. (2020). Novel artificial intelligence system increases the detection of prostate cancer in whole slide images of core needle biopsies. *Journal of Modern Pathology, 33*, 2058–2066.

Reed, M. J., Cooke, C., Mcmahon, N., Hands, K., Henderson, S., Knight, E., Littlewood, N., Latif, M., Tod, N., Mcgarvey, M., Hughes, N., Donald, M., Rowley, M., Innes, C., & Lockhart, S. (2020). Improvements in National Code Red transfusion practice in Scotland after adoption of recommendations from the Scottish National Code Red 2015 review. *Injury, 51*, 913–918.

Retson, T. A., & Eghtedari, M. (2020). Computer-aided detection/diagnosis in breast imaging: A focus on the evolving FDA regulations for using software as a medical device. *Current Radiology Reports, 8*, 1–7.

Reverter, J. L., Vázquez, F., & Puig-Domingo, M. (2019). Diagnostic performance evaluation of a computer-assisted imaging analysis system for ultrasound risk stratification of thyroid nodules. *American Journal of Roentgenology, 213*, 169–174.

Ryu, H. S., Jin, M.-S., Park, J. H., Lee, S., Cho, J., Oh, S., Kwak, T.-Y., Woo, J. I., Mun, Y., & Kim, S. W. (2019). Automated Gleason scoring and tumor quantification in prostate core needle biopsy images using deep neural networks and its comparison with pathologist-based assessment. *Cancers, 11*, 1860.

Sadeghi, A. H., Maat, A. P., Taverne, Y. J., Cornelissen, R., Dingemans, A.-M. C., Bogers, A. J., & Mahtab, E. A. (2021). Virtual reality and artificial intelligence for 3-dimensional planning of lung segmentectomies. *JTCVS Techniques, 7*, 309–321.

Sasaki, M., Tozaki, M., Rodríguez-Ruiz, A., Yotsumoto, D., Ichiki, Y., Terawaki, A., Oosako, S., Sagara, Y., & Sagara, Y. (2020). Artificial intelligence for breast cancer detection in

mammography: Experience of use of the ScreenPoint Medical Transpara system in 310 Japanese women. *Journal of Breast Cancer, 27*, 642–651.

Schmalz, O., Jacob, C., Ammann, J., Liss, B., Iivanainen, S., Kammermann, M., Koivunen, J., Klein, A., & Popescu, R. A. (2020). Digital monitoring and management of patients with advanced or metastatic non-small cell lung cancer treated with cancer immunotherapy and its impact on quality of clinical care: Interview and survey study among health care professionals and patients. *22*, e18655.

Scientific, T. J. S. E. (2017). ARL QUANT'X EDXRF.

Shortliffe, E. H. (1976). *Computer-based medical consultants: MYCIN*. Elsevier.

Singh, R., Kalra, M. K., Homayounieh, F., Nitiwarangkul, C., Mcdermott, S., Little, B. P., Lennes, I. T., Shepard, J.-A. O., Digumarthy, S. R., & Surgery. (2021). Artificial intelligence-based vessel suppression for detection of sub-solid nodules in lung cancer screening computed tomography. *Journal of Quantitative Imaging in Medicine, 11*, 1134.

Stålhammar, G., Fuentes Martinez, N., Lippert, M., Tobin, N. P., Mølholm, I., Kis, L., Rosin, G., Rantalainen, M., Pedersen, L., & Bergh, J. J. M. P. (2016). Digital image analysis outperforms manual biomarker assessment in breast cancer. *29*, 318–329.

Steinwendner, J. (2020). Klinische Entscheidungsunterstützungssysteme: von der Datenrepräsentation zur künstlichen Intelligenz. *Innovationen und Innovationsmanagement im Gesundheitswesen.* Springer.

Stephens, K. (2021). iCAD launches personalized risk assessment technology for 3D mammography. *AXIS Imaging News*.

Sweilam, N. H., Tharwat, A., & Moniem, N. A. (2010). Support vector machine for diagnosis cancer disease: A comparative study. *Egyptian Informatics Journal, 11*, 81–92.

Tan, V., Payne, J., Paquet, N., Iles, S., Rayson, D., Barnes, P., & Abdolell, M. (2020). The association between breast density and HER2-positive breast cancer: A population-based case-control study. *American Society of Clinical Oncology*.

Tartar, M., Le, L., Watanabe, A. T., & Enomoto, A. (2021). Artificial intelligence support for mammography: In-practice clinical experience. *Journal of the American College of Radiology, 18*, 1510–1513.

Thagaard, J. (2017). Detecting lymph node metastases in breast cancer using deep learning.

Turkbey, B., & Haider, M. A. (2022). Deep learning-based artificial intelligence applications in prostate MRI: Brief summary. *The British Journal of Radiology, 95*, 20210563.

Uijterwaal, M. H., Polman, N. J., Witte, B. I., Van Kemenade, F. J., Rijkaart, D., Berkhof, J., Balfoort-Van der Meij, G., Ridder, R., Snijders, P. J., & Meijer, C. J. (2015). Triaging HPV-positive women with normal cytology by p16/K i-67 dual-stained cytology testing: Baseline and longitudinal data. *International Journal of Cancer, 136*, 2361–2368.

Varma, M., Narahari, K., Mason, M., Oxley, J. D., & Berney, D. M. (2018). Contemporary prostate biopsy reporting: Insights from a survey of clinicians' use of pathology data. *Journal of Clinical Pathology, 71*, 874–878.

Vijayalakshmi, S., Gayathri, S., & Janarthanan, S. (2021). Blockchain security for artificial intelligence-based clinical decision support tool. *Internet of things, artificial intelligence and blockchain technology.* Springer.

Wang, X., Lin, R., Zeng, Y., Wang, Y., Wei, S., Lin, Z., Chen, S., Ye, Z., & Chen, L. (2022). High expression of FCRLB predicts poor prognosis in patients with colorectal cancer. *Frontiers in Genetics, 13*, 882307.

Wang, J. M., Ram, S., Labaki, W. W., Han, M. K., & Galbán, C. J. (2022). CT-based commercial software applications: Improving patient care through accurate COPD subtyping. *International Journal of Chronic Obstructive Pulmonary Disease, 17*, 919.

Wang, Y., Zhu, H., Madabushi, R., Liu, Q., Huang, S. M., Zineh, I., & Therapeutics. (2019). Model-informed drug development: Current US regulatory practice and future considerations. *Journal of Clinical Pharmacology, 105*, 899–911.

Weisberg, E., Chu, L., Park, S., Yuille, A., Kinzler, K., Vogelstein, B., & Fishman, E. (2020). Deep lessons learned: Radiology, oncology, pathology, and computer science experts unite around

artificial intelligence to strive for earlier pancreatic cancer diagnosis. *Diagnostic and Interventional Imaging, 101*, 111–115.

Woodhouse, R., Li, M., Hughes, J., Delfosse, D., Skoletsky, J., Ma, P., Meng, W., Dewal, N., Milbury, C., & Clark, T. (2020). Clinical and analytical validation of FoundationOne Liquid CDx, a novel 324-Gene cfDNA-based comprehensive genomic profiling assay for cancers of solid tumor origin. *PloS One, 15*, e0237802.

Woods, T. O. (2007). Standards for medical devices in MRI: Present and future. *Journal of Magnetic Resonance Imaging, 26*, 1186–1189.

Xu, X., Bao, L., Tan, Y., Zhu, L., Kong, F., Wang, W., & Biology. (2018). 1000-case reader study of radiologists' performance in interpretation of automated breast volume scanner images with a computer-aided detection system. *Journal of Ultrasound in Medicine, 44*, 1694–1702.

Xu, F., Pan, B., Zhu, X., Gulaka, P., Xiang, L., Gong, E., Zhang, T., Wang, J., Lin, L., & Ma, Y. (2020). Evaluation of deep learning based PET image enhancement method in diagnosis of Lymphoma. *Society Nuclear Medicine*.

Xue, Z., Novetsky, A. P., Einstein, M. H., Marcus, J. Z., Befano, B., Guo, P., Demarco, M., Wentzensen, N., Long, L. R., & Schiffman, M. (2020). A demonstration of automated visual evaluation of cervical images taken with a smartphone camera. *International Journal of Cancer, 147*, 2416–2423.

Yamada, M., Saito, Y., Imaoka, H., Saiko, M., Yamada, S., Kondo, H., Takamaru, H., Sakamoto, T., Sese, J., & Kuchiba, A. (2019). Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Scientific Reports, 9*, 1–9.

Yoshizawa, A. (2013). A trial of integrated telepathology (iTP) in Nagano prefecture. *Rinsho byori. The Japanese Journal of Clinical Pathology, 61*, 65–73.

Zhao, S., Wang, S., Pan, P., Xia, T., Chang, X., Yang, X., Guo, L., Meng, Q., Yang, F., Qian, W., Xu, Z., Wang, Y., Wang, Z., Gu, L., Wang, R., Jia, F., Yao, J., Li, Z., & Bai, Y. (2019). Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: A systematic review and meta-analysis. *Gastroenterology, 156*, 1661–1674.e11.

Chapter 11
Intelligent Drug Design and Use for Cancer Treatment: The Roles of AI and Precision Oncology in Targeting Patient-Specific Splicing Profiles

Rodney Hull, Mosebo A. Manabile, and Zodwa Dlamini

Abstract The development of new drugs is expensive, time-consuming, and often results in failure. These problems can partially be solved through the use of AI to identify drug targets, search for molecules capable of interacting with these targets, and then model the interactions of the drug and its target while modelling the physiochemical properties of this drug. Alternative splicing is commonly altered in cancer and as such has become a target for the designing of new drugs. While many drugs have been designed to target either the new isoforms that favour cancer development or proteins involved in the splicing pathway, AI can improve this by helping screen proteome and transcriptome databases to identify new splice variants. AI can also model the three-dimensional structure of new isoforms in order to screen for compounds that can bind exclusively to these isoforms.

Keywords Neural networks · Machine learning · Deep learning · Virtual screening · 3D modelling · Isoforms · Splice variants · Precision oncology

11.1 Introduction

The development of new drug treatment for cancer involves the investment of large periods of time and is associated with high costs and a low success rate (Waring et al., 2015). It is estimated that over 90% of new drugs fail between the initiation of phase 1 trials and the approval phase, mostly due to safety issues. This does not take into account the multitude of compounds that are screened before the few successful compounds enter phase 1 trials (Fleming, 2018). This means a lot of time and money is wasted. Figure 11.1 gives an indication of the costs of drug design and how much

R. Hull (✉) · M. A. Manabile · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: rodney.hull@up.ac.za; u20804289@tuks.ac.za; zodwa.dlamini@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_11

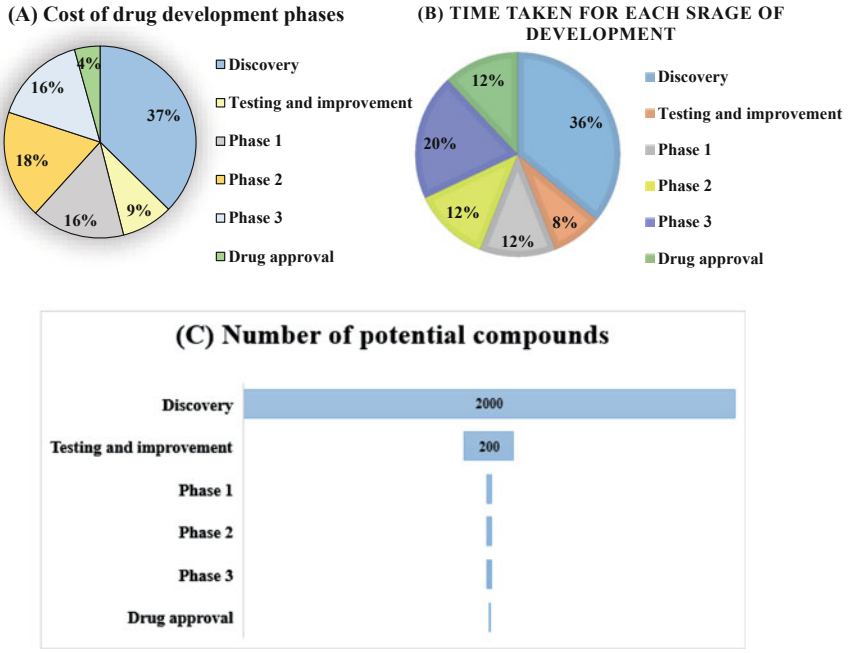


Fig. 11.1 The cost of drug design at various stages of the process and the chance of success. Drug discovery and design consists of multiple steps which are costly and time consuming. These steps include target selection and validation, therapeutic screening and lead compound optimization, pre-clinical and clinical trials, and manufacturing practices. The high failure rate of many drugs increases the cost of developing a new drug

cost is involved to complete each stage of the process. The costs shown are in both money (1a) and time (1b) and a graphical representation of the success rate is given in Figure 11.1c. The initial screening of compounds contributes a large percentage to both financial and time costs in the process. At this stage animal models need to be used and although this cost can be lowered using *in vitro* techniques, it still involves the screening of many compounds. The cost can be lowered even further using *in silico* approaches to only select the most promising compounds for further testing. The use of computer models can identify the most promising compounds to interact with specific targets and predict many of the physiochemical characteristics of these compounds, removing compounds that are less favourable for being developed as potential drug candidates (Zhang et al., 2017). These techniques which include virtual screening (VS) and molecular docking continue to have a low success rate because they are traditionally unable to accurately predict the *in vivo* actions of these compounds, which may significantly differ from those that are observed *in vitro* and cannot be foreseen by *in silico* analysis (Hughes et al., 2011). Artificial intelligence and machine learning give computer algorithms the ability to learn to recognize patterns in biological systems. Combined with the ability to learn to recognize and

interpret the properties of molecules based on their structure means that AI has the potential to interact with biological systems and predict side effects and toxicity.

The identification of specific targets for the design of new drugs is also an important role that AI can play in the treatment of cancer. These drug targets can be identified through the analysis of molecular profiles by AI. This would allow for the identification of molecular changes specific to a type of cancer, population group, or even a single patient. It is well known that the same treatments do not work in the same way across multiple population groups, with some groups responding poorly to a treatment that was used successfully in another population group. In addition to this, the more specific the molecular changes are to cancer, the better the chances of avoiding side effects or off-target effects. To assist in this, it is best to target a cancer-specific process. Alternative splicing results in multiple proteins and mRNA transcripts being coded for by a single gene. It is estimated that up to 90% of all human genes are alternatively spliced. Alterations in AS are common in cancer, with cancer-specific protein isoforms sometimes favouring the development and progression of the disease. This makes these protein-specific isoforms or the mechanisms that lead to the formation of these cancer-specific isoforms attractive targets for the design of new drugs. This chapter will discuss the use of AI in designing drugs based on cancer-specific splicing profiles to develop new treatments that can be personalized to specific individuals or population groups. It will discuss how AI can be used to identify specific targets, identify compounds that can interact with these targets, and demonstrate how AI can be used to assess the physiochemical properties of these compounds and predict any side effects they may have. Finally, it will discuss how AI can be used to predict how different treatments will behave based on biological data from an individual patient.

11.2 The Application of AI in Drug Design

High-throughput technologies such as next generation sequencing and mass spectrometry have given rise to large complex sets of biological data. This has been accompanied by improvements in computer hardware and the rise of machine learning and deep learning in AI. In addition to this, the data gathered through the use of these techniques have been entered into large data sharing platforms or databases. AI can make use of advanced computational approaches to help process and analyse these large sets of data. These include neural networks, decision trees, natural language processing (NLP), cloud computation, and graphics processing (Mirza et al., 2019; Dlamini et al., 2022). The ability of AI to aid in the rational design of drugs, identify drugs to be repurposed, optimize the design of clinical trials, and assist in the manufacturing and marketing of drugs means that AI can assist in the development of a pharmaceutical product from the bench to the bedside (Duch et al., 2007). AI's ability to effectively and efficiently identify molecules that could interact with a drug target relies on its ability to identify and predict chemical structure and characterize its pharmacophores and three-dimensional structure. This

relies on the use of algorithms and mathematical equations to predict a molecule's structure and biological activity. Despite the use of recently designed ML algorithms, it is important to note that these algorithms rely on classical chemistry (Sellwood et al., 2018). These algorithms that are used in virtual screening for drugs, including predicting their *in vivo* activity and toxicity, include algorithms such as Nearest-Neighbour classifiers, random forest decision trees (RF), extreme learning machines, support vector machines (SVMs), and deep neural networks (DNNs) (Álvarez-Machancoses & Fernández-Martínez, 2019; Dana et al., 2018) (Fig. 11.2).

The primary role played by AI in the analysis of big data sets is well known, which is due to its ability to organize and annotate the data which can then be used to identify features specific to cancer. In this case it is the identification of molecular patterns associated with alternative splicing. This can be the expression pattern of different protein isoforms, or the expression pattern of different splicing factors. It can also be the identification of different transcripts arising from alternative splicing and the level of these different transcripts. In addition to the ability of AI to identify specific molecular profiles for individual patients, AI can also assist in designing drugs able to target these different transcripts, protein isoforms, or even those regulatory proteins that control the splicing process. The latter drug target includes designing drugs that target differentially expressed splicing factors. AI's ability to assist in precise manufacturing of drugs means that AI can be used to make specific dosages of drugs or different mixtures of compounds in a single treatment (Blasiak et al., 2020) (Table 11.1).

11.3 The Role of AI in Drug Screening

Once AI has been used to identify molecular profiles unique to the disease within an individual, these unique molecular profiles can be targeted for treatment. AI can assist here by searching chemical databases and literature to identify molecules that will associate with the molecules that arise due to these molecular alterations. These search hits can then be used as lead compounds for further investigation. AI can then be used to determine pharmacological features of this compound, its toxicity and validate its bioactivity and practical usefulness as a drug (Mak & Pichika, 2019; Sellwood et al., 2018).

Drug design uses annotation systems such as the Simplified Molecular Input Line Entry System (SMILES). SMILES is a linguistic construct representing chemical structure not a computer data structure using atom and bond symbols and only a few grammar rules. This is a convenient notation for chemicals that AI can use (Weininger & Sciences, 1988). Algorithms such as DNN can be used to model structures or make feature predictions including data describing a molecule's potential energy measurement, electron density, and coordinates of atoms (Hessler & Baringhaus, 2018). Other algorithms involve DL using undirected graph recursive neural networks. Convolutional neural networks (CVNN) making use of

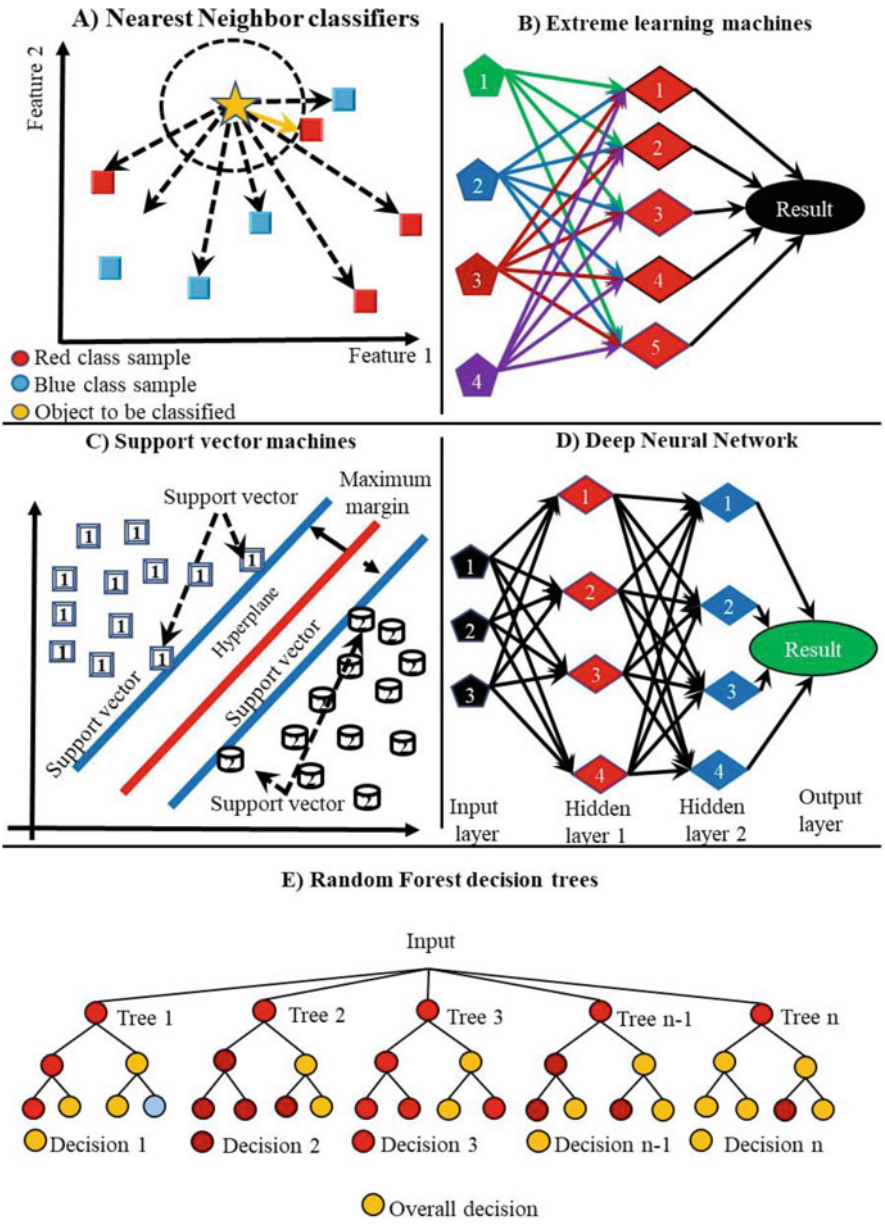


Fig. 11.2 Representations of techniques used by algorithms applied to virtual screening. (**a**) Nearest-neighbour classifiers are used to define entries based on known records by calculating the distance to tall training records based on their position as calculated by two features of the data points. The nearest neighbour is identified, and the labels of this neighbour are given to the unknown. (**b**) Extreme learning machines are neural networks with one or more hidden layers. (**c**) Support Vector Machine: Schematic representation of a linear SVM. Voxels (a unit of graphic information that difines a point in three-dimensional space as opposed to a pixel which defines a point in two dimensional space. The blue squares represent voxels belonging to cancer samples while the cylinders are voxels belonging to healthy samples. A hyperplane separates these which is

graph-based have been used to predict the solubility of molecules (Kumar et al., 2017) and ANN-based models using kernels and kernel ridge-based models (Rupp et al., 2010). Figure 11.3 is a depiction of the workflow that AI uses to design drugs.

This table lists the databases that contain omics data (largely fenomic) which can be used for drug design. It also lists databases that contain data concerning chemical compounds, for instance chemical structures and toxicity.

11.3.1 Prediction of Physicochemical Properties and Bioactivity Using AI

The physicochemical properties of a compound include properties such as how soluble it is and in what solvent it is most soluble, interatomic distance, partition coefficient, ionization, permeability, and intermolecular forces. It is important to establish these properties for a potential drug as these dictate the molecule's pharmacokinetics and its ability to associate with its target (Zang et al., 2017). Certain machine learning tools use large data sets that have been produced during previous compound optimization studies to predict physicochemical properties (Yang et al., 2019). AI can also be used to predict the toxicity of a compound. A compound that is toxic can be removed from the drug discovery pipeline thereby preventing further time and money being wasted on its development. This is also useful as AI can be used to alter the structure of a drug in a way as to decrease the toxicity of the drug but not its activity. Currently the most successful AI toxicity prediction tool is the ML-based DeepTox. This software predicts the toxicity of a compound based on 2500 toxicophoric features (Mayr et al., 2016). Another ML-based program called eToxPred was able to predict the toxicity with a 72% accuracy of small organic molecules (Mayr et al., 2016).

11.3.2 AI Predictions of the Mode of Action of Potential Drugs

By establishing the affinity of a molecule for its target dictates the efficacy of the potential drug. Drug affinity predictions can also be used to establish if there are any off-target molecules which the drug can bind to. Off-targets are any proteins or ligands the drug may bind to unintentionally. This can lead to undesirable side

Fig. 11.2 (continued) the maximum separation from the closest points of the two groups. Data points on the margins of separation are the support vectors. (d) Deep neural network uses multiple hidden layers with the data from all nodes in each layer being analysed by all nodes in the next. (e) The random forest decision tree increases the accuracy of the decision process by producing many trees and selecting the most common outcomes as the final decision

Table 11.1 List of databases that contain large amounts of data for drug design

Database	Purpose	Reference
PubChem	Public repository of chemical structure and associated biological properties	Wishart et al. (2018)
ChEMBL	Information on compounds' binding, function, and toxicity	Wishart et al. (2018)
DrugBank	Information on all approved drugs and their mechanisms, interactions, and apposite targets	Wishart et al. (2018)
NCBI Gene Expression Omnibus (GEO)	GWAS data	Edgar et al. (2002)
The Cancer Genome Atlas (TCGA)	GWAS data. Sequencing data related to cancer	Wang et al. (2016)
ArrayExpress	GWAS data	Parkinson et al. (2007)
GWAS central	GWAS data	Beck et al. (2014)
NHGRI-EBI GWAS Catalogue	GWAS data	Buniello et al. (2019)
Sequence read archive	Sequencing data	Leinonen et al. (2011)
The National Cancer Institute Genomic Data Commons (NCIGDC)	Sequencing data related to cancer	Jensen et al. (2017)
DriverML	Tool that can point out driver genes related to cancer	Han et al. (2019)

effects. Drug target binding affinity (DTBA) can be determined by AI-based methods. Generally, this is done in two ways. AI can determine any similarities or features that are shared between the drug and its target. The features method involves the use of the chemical moieties of the drug and the target to identify feature vectors. In the similarities method, the more similarities between the drug and the target, the higher the affinity they have for each other (Öztürk et al., 2018). There are web applications that have been designed to predict drug–target interactions (Lounkine et al., 2012). Tools such as Kronecker-regularized least squares (KronRLS) are used to compare the similarities of the drug and the target (Öztürk et al., 2018). Some tools such as SimBoo use both the feature and similarity methods. Other commonly used methods include SMILES, ligand maximum common substructure (LMCS), and extended connectivity fingerprinting (Öztürk et al., 2018). In cases where there is no accurate 3D model of the protein available then DL approaches such as DeepAffinity and Protein And Drug Molecule interaction prEdiction (PADME) can be used (Lounkine et al., 2012).

AI also allows for the automated mining of scientific literature for novel drug targets and lead compounds from the literature. These text mining-based tools use natural language processing (NLP) to turn unstructured text into structured data. This structured data can then be used to identify molecules of interest. NLP gives computers the ability to interpret human language, both speech and text, using AI

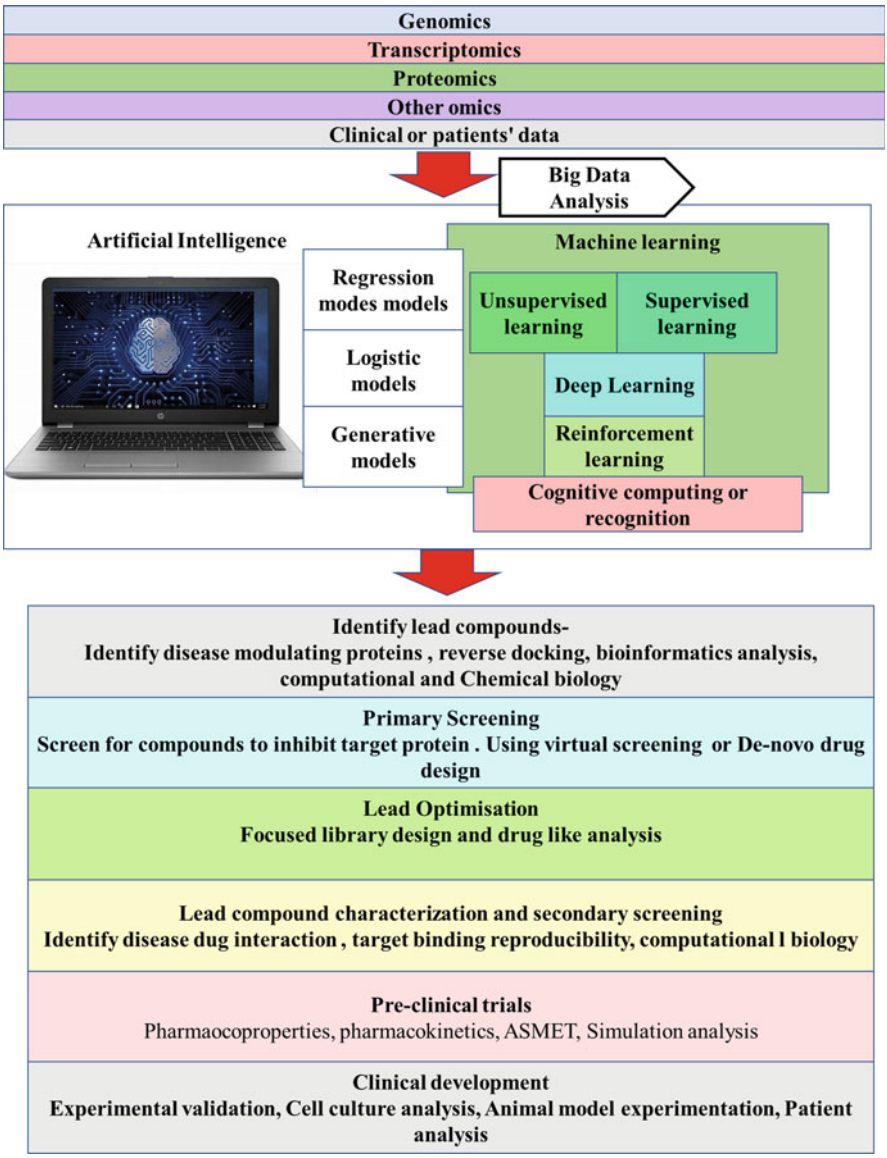


Fig. 11.3 Schematic of drug screening workflow. Large databases are analysed by AI to extract features, for example molecular profile changes, of interest using various algorithms and applying various forms of machine learning including deep learning. These features are used to search libraries and databases of chemical compounds to identify molecules that may interact with an identified molecule, most likely a protein. Once a molecule of interest is identified, more complex information gathering and analysis of the drug like properties of this compound are carried out. The molecule's binding affinity with the target and its pharmaco-properties can then be examined. Virtual drug testing can then be performed *in silico.* All these stages can be carried out with the aid of AI; however, the final validation steps must involve cell culture or animal models

algorithms (Ficenec et al., 2003). Molecular fingerprinting involves technique used to define the compound and is normally based on the structure of the compound. In the early stages of drug discovery a drug's chemical structure is normally used not only as a fingerprint but also to model drug–target interactions (Labute, 2000). This structure can then be used by computer-assisted drug design processes to search large databases of chemicals in order to identify similar molecules that may interact with the target. Alternately, the protein or ligand can be used to search the same databases in order to identify lead compounds. AI would allow for the development of other chemical descriptors besides structure. These include the use of nonlinear modelling algorithms (Zhu & Xia, 2016). The algorithms and AI tools that can be used to analyse drug activity, pharmaco-properties, and drug–target interactions are listed in Table 11.2.

The chemical space is a conceptual description of the space spanned by all possible molecules and compounds that fall under a set of given principles or boundaries. The virtual chemical space is the virtual or computational version of this chemical space. This virtual chemical space is enormous and can be interpreted as a map of the distributions of molecules based on their properties. Therefore, this space can be used to perform virtual screening (VS) by searching for bioactive compounds based on positional information about molecules within the space. *In silico* search methods use ligand-based approaches and can be much cheaper than other drug screening methods (Mak & Pichika, 2019).

11.4 Techniques and Tools for Computational Drug Discovery

The quantitative structure-activity relationship (QSAR) models assume that compounds with a similar structure will have similar activities. These models have been shown to identify many compounds which may serve as lead targets. It has also been shown to have the ability to make simple physicochemical predictions. However, when it comes to successful drug design and discovery these models only work using chemical structure and target activity, which is not enough to ensure the success of an identified compound (Russo & Zhu, 2022). This is due to the fact that these models lack the ability to model complex biological properties (Zhao et al., 2017). However, QSAR was first used in the 1960s and since then has evolved into AI-based QSAR approaches. These approaches use techniques such as linear discriminant analysis (LDA), support vector machines (SVMs), random forest (RF), and decision trees to increase the speed and accuracy of QSAR models (Zhang et al., 2017). These simple models also suffer when they are used as machine learning algorithm as they commonly only have access to small data training sets and may have to use unvalidated experimental data which may be full of errors (Zhao et al., 2017).

Table 11.2 List and description of drug activity, pharmocol-property, and binding databases

Database	Purpose	Reference
Absorption, distribution, metabolism, and excretion (ADME)	Determining the sites of metabolism of the drug	Gaulton et al. (2012)
DrugBank	Information on all approved drugs and their mechanisms, interactions, and apposite targets	Wishart et al. (2018)
DrugMatrix	Toxicogenomic information data of drugs	Gilson et al. (2016)
Binding Database (BindingDB)	Information regarding drug-target (protein/enzyme) binding	Armbrust et al. (2010)
Library of integrated network-based cellular signature (LINCS)	Change in gene expression signatures of human cell lines when treated with different chemical compounds	Keenan et al. (2018)
DriverML	Tool that can point out driver genes related to cancer	Han et al. (2019)
SimBoost	Considers both feature-based and similarity-based interactions to predict DTBA	Öztürk et al. (2018)
MANTRA	Groups compound based on similar gene expression profiles and therefore similar mechanism of action	Lounkine et al. (2012)
DeepNeuralNetQSAR	Detection of the molecular activity of compounds	Chan et al. (2019)
ORGANIC	A molecular generation tool	Brown (2015)
PotentialNet	Uses NNs to predict binding affinity of ligands	Pereira et al. (2016)
Hit Dexter	Predicts molecules that might respond to biochemical assays	Pereira et al. (2016)
DeltaVina	A scoring function for rescoring drug–ligand binding affinity	Pereira et al. (2016)
Neural graph fingerprint	Helps to predict properties of novel molecules	Pereira et al. (2016)
DeepTox	suitable candidate in drug discovery	Ciallella and Zhu (2019)
DeepChem	AI system that finds suitable candidates in drug discovery	Zhu (2020)

This table lists databases that contain data on the physiochemical nature of compounds such as molecular weight, melting temperature, and dissociation constant, as well as their binding affinities and biological activates.

AI can also use quantum mechanics in the form of predicting molecular orbitals and wave functions of organic molecules using a DL-driven tool named SchNOrb. This data can be used to predict the arrangement of chemical bonds and electronic properties of a molecule. This in turn can be used to identify reactive sites (Schütt et al., 2019). The behaviour of molecules or their molecular dynamics (MD) can be simulated to show how molecules interact at the atomic level (Gastegger et al., 2020). These simulations can then be used to study the interaction and binding of a drug to its target. Even though MD existed before AI and can be done using normal modelling techniques, this is time-consuming and labour-intensive. AI has the capacity to accelerate MD and lighten the workload of researchers (De Vivo et al., 2016). De novo drug design describes the process of developing novel compounds to serve as lead targets for drug development. Despite its usefulness and past success, traditional de novo drug design may involve complicated drug synthesis and an inability to accurately predict the bioactivity of the designed molecule (Yang et al., 2019). This process has been improved using AI as well as machine and deep learning. One of these tools, MolAICal, uses a DL genetic algorithm trained on data describing approved drugs from the US Food and Drug Administration (FDA) to design potential drugs. These drugs are then assessed for their molecular docking using DL algorithms trained on the ZINC database (Grzybowski et al., 2018). DNNs have been applied to the rules of organic chemistry and retrosynthesis. Using Monte Carlo tree search, this AI can predict the reaction process required to synthesize a molecule. They can also use the same methods to assist in drug discovery and design (Segler et al., 2018).

11.5 Protein Modelling and Docking

The most common type of target for most drugs are proteins. Their contribution to disease might be due to structural changes due to changes in their amino acid sequence or their expression level. In this chapter we are most interested in protein targets that arise due to splicing changes giving rise to pro-cancer isoforms. To target these proteins, it is important to be able to accurately predict the structure of these proteins. Protein folding can be predicted by replicating the four stages of protein folding virtually in a protein folding pipeline to produce a 3D model (Fiser et al., 2000). The starting material for 3D protein modelling is the amino acid sequence of the protein, its primary structure. Changes in the amino acid sequence can result in changes in how the protein folds into the secondary structure to form structures such as alpha helices and beta sheets. Neural networks can be trained to model how changes in the primary structure affect this initial folding into the secondary structure (Fiser et al., 2000). For example, changes that have occurred due to alternative splicing of mRNA will alter the amino acid sequence of the protein (Fig. 11.4). This

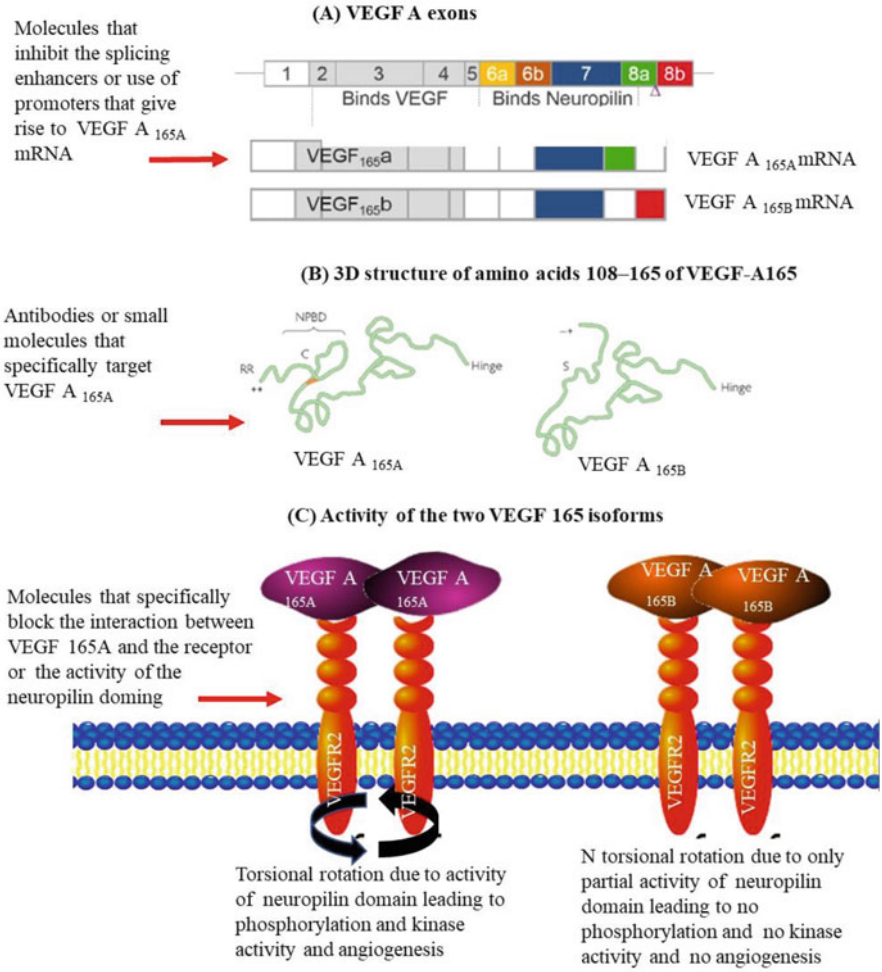


Fig. 11.4 Isoform modelling and interactions. Changes to three-dimensional structure of a protein caused by alternative splicing can be determined by examining the changes in mRNA sequence, which can then be used to determine the amino acid sequence and the structure of the protein. Using VEGF as an example the changes in the alternatively spliced mRNA of the pro-angiogenic 165 A and 165b show that the neuropilin domain will not form correctly in the anti-angiogenic 165b (A) as reflected in the structural change in the protein, as shown by the 3D models (B). Pathway analysis shows how this structural change can inhibit the angiogenic process (C). This also shows the various stages where drugs can be designed to target splicing or promote the expression of certain variants (red arrows)

is done by the calculation of the distances between pairs of residues. The tertiary structure of the protein is then modelled by arranging the secondary structural elements into a three-dimensional structure using specific algorithms, such as the Monte Carlo Metropolis algorithm (Fischer et al., 2016) and automated modelling

Table 11.3 List and description of multiple tools for structural prediction

Database	Purpose	Reference
Protein data bank (PDB)	Three-dimensional structures of proteins, DNA, RNA	Rose et al. (2017)
MolAIcal	Can design three-dimensional drugs in three-dimensional protein pockets	Bai et al. (2021)
AlphaFold	Predicts the 3D structure of proteins from their amino acid sequences	Powles and Hodson (2017)
Recurrent Geometric Network	Generates three-dimensional structure of proteins from amino acid sequence	AlQuraishi (2019)
SchNOrb	Predicts molecular orbitals and wave functions of organic molecules	Schütt et al. (2019)
ChemMapper	Predicting drug–target interactions	Lounkine et al. (2012)
Similarity ensemble approach (SEA)	Predicting drug–target interactions	Lounkine et al. (2012)

(Fiser et al., 2000). An AI-based tool named AlphaFold was created by Google's DeepMind and this tool was trained using 3D structure data from the Protein Data Bank (PDB). What is interesting about this AI is how it was trained, which was done in two steps. First a CNN was used to create a matrix of distances and torsion angles based on an amino acid sequence. These two matrices were then used to construct a three-dimensional protein model using a gradient optimization technique (Senior et al., 2020). Other AI applications that also use bond length and angles of rotation to predict 3D protein structure rely on DL-based Recurrent Geometric Network (RGN) (Segler et al., 2018). The AI-based tools that can be used to perform 3-dimensional modelling are listed in Table 11.3.

This table lists tools that can be used to predict or model the three-dimensional structure of proteins and compounds.

Once the 3D structure of the protein is known, it is possible to predict how the drug will interact with the protein. The drug can also be designed to interact with the protein target site in a specific chemical environment. In terms of alternate isoforms of the protein which may be missing entire domains or contain extra amino acids, the structure prediction may show a completely altered protein and this may lead to completely different structures and drug designs (Wan & Zeng, 2016). Random forest models have also been used to predict drug–protein interactions using integrated pharmacological and chemical data. These RF models could also predict target–disease and target–target associations that could be used for future drug design (Yu et al., 2012). Repurposed drugs qualify for Phase II clinical trials (Mak & Pichika, 2019), making its development faster. It is also cheaper, costing only half of what the development of a new drug costs (Persidis, 2011). Understanding the interaction between drugs and proteins is also important when a drug is being repurposed. It also prevents polypharmacotherapy or off-site interactions (Wan & Zeng, 2016). The ability of repurposing a drug can be investigated using the 'guilt by association' approach. Here similarities between different targets, diseases, drugs,

chemical structures, and gene expression profiles are used to search for new uses for an existing drug (Koromina et al., 2019; Park, 2019) (Fig. 11.4).

11.6 Drugs Targeting Alternative Splicing

Alternative splicing increases the diversity and number of proteins produced by a limited number of genes by splicing pre-mRNA allowing for one gene to produce multiple protein isoforms. The involvement of alternative splicing (AS) in many aspects of cellular homeostasis, differentiation as well as tissue growth and development, means that any dysregulation of this process could result in the development of disease (Baralle & Giudice, 2017). Aberrations in the splicing machinery and profiles of spliced mRNA and protein isoforms are a regular occurrence in many cancers. This is especially true if those produced isoforms favour the development or progression of cancer or influence pathways in the hallmarks of cancer such as the generation of malignant protein isoforms (Wang & Aifantis, 2020). The range of molecules that may be affected in cancer that can lead to changes in the splicing process and isoform expression profile include RNA–protein complexes and associated regulatory proteins (Oltean & Bates, 2014). This has led to the adoption of alternative splicing regulators and the pro-oncogenic isoforms AS produces, as targets for the development of drugs.

Multiple different classes of drugs have been developed that target alternative splicing. These include small molecule inhibitors that act as spliceosome inhibitors. Some of these may also target the auxiliary proteins of the spliceosome. Some of these are phytochemicals or their synthetic analogues and are structurally related compounds (Martínez-Montiel et al., 2016). The targets of these molecules include the splicing factor SF3b and its subunits or spliceosomal associated proteins (Martínez-Montiel et al., 2016; Teng et al., 2017). Other targets of these small molecules such as pladienolide B include the actual pro-cancer isoforms. Pladienolide B increased the levels of pro-apoptotic p73 while decreasing the levels of anti-apoptotic p73 (Zhang et al., 2019). Serine arginine protein kinases (SRPKs) and Cdc-like kinases (CLKs) are also targets for drugs that aim to influence AS. Aberrantly expressed SR proteins have been noted to greatly affect malignancy and cancer development (Kędzierska & Piekiełko-Witkowska, 2017). The CLK inhibitor TG-003 suppressed SR protein phosphorylation (ElHady et al., 2017). SRPIN340 and SPHINX are two SRPK inhibitors that are well studied (Fukuhara et al., 2006; Gammons et al., 2013). SRPIN340 has high inhibitory activity against SRPK1 and SRPK2. It has been demonstrated that it is able to reduce melanoma tumour growth *in vivo*. It affects the splicing of VEGF and angiogenesis, reducing the expression of pro-angiogenic isoforms of VEGF (Gammons et al., 2014). SRPIN340 also reduces cell migration, invasion, and colony number formation (Moreira et al., 2018). Despite these effects SRPIN340 failed to be developed as a drug due to its poor absorption *in vivo* and poor pharmacokinetic properties (Gammons et al., 2014). SPHINX was a more potent inhibitor of SRPK1 compared

to SRIN340 and was shown to significantly downregulate pro-angiogenic $VEGF_{165}$ expression (Mavrou et al., 2015).

Protein arginine methyltransferases (PMRTs) are splicing regulators that are upregulated in many cancers (Maron et al., 2022). PRMTs methylate arginine to post translationally modify proteins. This modification affects the binding of RNA binding proteins (RBPs) which act to regulate splicing and maintain accurate splicing (Dowhan et al., 2012). As such PRMTs are attractive drug targets for the treatment of cancer (Hwang et al., 2021). Small molecules that inhibit PRMTs have been developed, with many in the pre-clinical trial stage (Guccione & Richard, 2019). AMI-1 targets PRMT 5 and is able to inhibit solid tumour progression, while GSK3368715 and MS023 are PRMT1 inhibitors that are known to affect the assembly of the spliceosome by inhibiting the methylation of RBPs (Guccione & Richard, 2019). PF-0639999 is a type II PRMT inhibitor that has been shown to be effective in treating lung cancer (Yang et al., 2022).

The identification of splicing events, leading to the discovery of new drug targets can be performed optimally using AI/ML models. The web tool AVISPA (Barash et al., 2013) predicts if a specific exon is alternatively spliced and what regulatory elements are associated with it by using a Bayesian NN classifier and a DNN model performed even better. This model used RNA-Seq data to predict AS across tissues (Leung et al., 2014). Another AI model was developed based upon these two previous models that can use unlabelled data and semi-supervised learning algorithms to predict AS events from DNA sequences (Stanescu et al., 2016; Xiong et al., 2015). AI algorithms can be used to help detect splicing isoforms from transcriptome data. Four different software employing AI, CypRules, MetaSite, MetaPred, SMARTCyp, and WhichCyp have been used to identify cytochrome P450 isoforms. These specific isoforms were identified when the authors were looking for increase in the metabolism of various drugs and their presence is an indicator of drug resistance. These programs all made use of SVM-based algorithms and were able to accurately identify multiple CYO450 isoforms with high accuracy (Pu et al., 2019).

11.7 Other Applications of AI in Drug Design

Apart from drug design, screening, and virtual testing, AI can be used to assist in manufacturing, marketing, and the design of clinical trials. For instance, once a drug has been designed and has shown promising results, its dosage and desired delivery characteristics need to be determined. In addition to this mass manufacture and synthesis of the drug need to be considered. Traditionally this involved trial-and-error testing, a process that can now be replaced by AI (Kalra et al., 2002). Computational tools can also be used to solve formulation problems (Mehta et al., 2019). Computational fluid dynamics (CFD) and discrete element modelling (DEM) can be used to optimize manufacturing by predicting the movement, dispersal, and packaging of drugs using automated production processes (Rantanen & Khinast,

2015). Clinical trials are needed to determine the safety and efficacy of a drug in humans. A process that requires 6–7 years and is costly (Hay et al., 2014). These trials often fail or are complicated due to inappropriate patient selection and infrastructure problems (Harrer et al., 2019). AI can be used to improve patient selection for recruitment in Phase II and III clinical trials using patient-specific genome or exposome profile analysis. This can give an indication of the presence and level of drug targets in patients and therefore, the effectiveness of the drug (Harrer et al., 2019; Mak & Pichika, 2019).

11.8 Limitations to AI-Based Drug Design

AI-based drug design promises to speed up the creation of new drugs and can aid the pharmaceutical industry in manufacturing and testing, resulting in lower costs and cheaper drugs. However, there are challenges that AI in drug design faces that must be overcome. AI depends on substantial amounts of data to learn from and interpret. This data must be reliable and be of a high quality. AI also requires skilled AI-based platforms operators and concerns arising due to potential job losses as AI replaces humans. There is also apprehension due to the black box problem, where users do not know how an AI has arrived at the result of the analysis due to algorithms changing and evolving due to DL and ML (Lamberti et al., 2019). However, human intervention is required as AI needs to be trained by humans. Data security is also a concern as AI-driven personalized medicine requires omics data of an individual and this constitutes personal information. The ethics surrounding the use and sharing of this information need to be established.

11.9 Conclusion

High costs and low efficiency are the greatest challenges that new drug design faces and this chapter has demonstrated how AI can be used to help solve these challenges and lead to shorter development times for drugs. This has been facilitated due to advances in computing and the ability to obtain big data sets from high-throughput omics technologies. AI also allows for the integration of data from different sources and levels of cell biology. In the case of drugs targeting alternative splicing, it allows for the integration of proteomic and transcriptomic data to identify splicing events and search drug databases for molecules that can target different protein isoforms. It also allows data regarding molecules that regulate the splicing process, such as expression and transcription profiles of RNA binding proteins and splicing factors to be integrated into the analyses. Despite the discussed challenges many new drugs targeting splicing and many new drug targets that arise due to splicing or are involved in the splicing process have been identified using AI (Fig. 11.5).

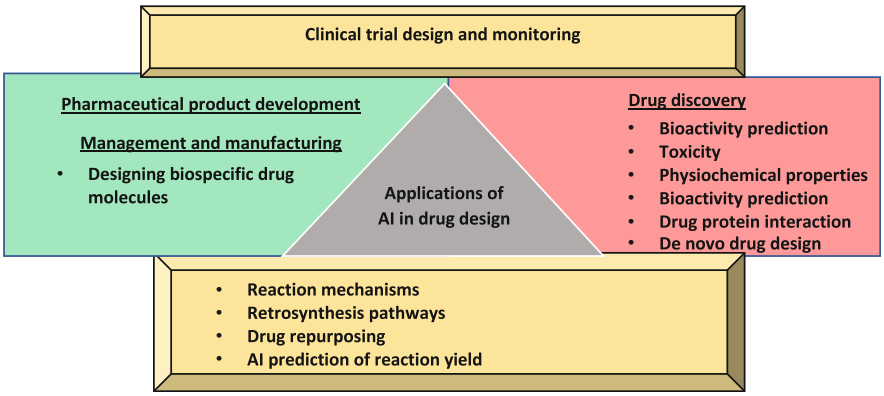


Fig. 11.5 The role of AI in drug discovery. The schematic represents the myriad of ways that AI can be used for the development of new drugs

References

Alquraishi, M. (2019). End-to-end differentiable learning of protein structure. *Cell Systems, 8*, 292–301.e3.

Álvarez-Machancoses, Ó., & Fernández-Martínez, J. L. (2019). Using artificial intelligence methods to speed up drug discovery. *Expert Opinon ionDrug Discovery, 14*, 769–777.

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A. & Stoica, I. J. C. O. T. A. 2010. A view of cloud computing. *53*, 50–58.

Bai, Q., Tan, S., Xu, T., Liu, H., Huang, J., & Yao, X. (2021). MolAICal: A soft tool for 3D drug design of protein targets by artificial intelligence and classical algorithm. *Briefings in Bioinformatics, 22*.

Baralle, F. E., & Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. *Nature Reviews. Molecular Cell Biology, 18*, 437–451.

Barash, Y., Vaquero-Garcia, J., González-Vallinas, J., Xiong, H. Y., Gao, W., Lee, L. J., & Frey, B. J. (2013). AVISPA: A web tool for the prediction and analysis of alternative splicing. *Genome Biology, 14*, R114.

Beck, T., Hastings, R. K., Gollapudi, S., Free, R. C., & Brookes, A. J. (2014). GWAS Central: A comprehensive resource for the comparison and interrogation of genome-wide association studies. *European Journal of Human Genetics, 22*, 949–952.

Blasiak, A., Khong, J., & Kee, T. (2020). CURATE.AI: Optimizing personalized medicine with artificial intelligence. *SLAS Technology, 25*, 95–105.

Brown, N. (2015). *In silico medicinal chemistry: Computational methods to support drug design*. Royal Society of Chemistry.

Buniello, A., Macarthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., Mcmahon, A., Morales, J., Mountjoy, E., Sollis, E., Suveges, D., Vrousgou, O., Whetzel, P. L., Amode, R., Guillen, J. A., Riat, H. S., Trevanion, S. J., Hall, P., Junkins, H., Flicek, P., Burdett, T., Hindorff, L. A., Cunningham, F., & Parkinson, H. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research, 47*, D1005–d1012.

Chan, H. C. S., Shan, H., Dahoun, T., Vogel, H., & Yuan, S. (2019). Advancing drug discovery via artificial intelligence. *Trends in Pharmacological Sciences, 40*, 592–604.

Ciallella, H. L., & Zhu, H. (2019). Advancing computational toxicology in the big data era by artificial intelligence: Data-driven and mechanism-driven modeling for chemical toxicity. *Chemical Research in Toxicology, 32*, 536–547.

Dana, D., Gadhiya, S. V., St Surin, L. G., Li, D., Naaz, F., Ali, Q., Paka, L., Yamin, M. A., Narayan, M., Goldberg, I. D., & Narayan, P. (2018). Deep learning in drug discovery and medicine; Scratching the surface. *Molecules, 23*.

De Vivo, M., Masetti, M., Bottegoni, G., & Cavalli, A. (2016). Role of molecular dynamics and related methods in drug discovery. *Journal of Medicinal Chemistry, 59*, 4035–4061.

Dlamini, Z., Skepu, A., Kim, N., Mkhabele, M., Khanyile, R., Molefi, T., Mbatha, S., Setlai, B., Mulaudzi, T., Mabongo, M., & Hull, R. O. (2022). AI and precision oncology in clinical cancer genomics: From prevention to targeted cancer therapies-an outcomes based patient care. *Informatics in Medicine Unlocked, 100965*.

Dowhan, D. H., Harrison, M. J., Eriksson, N. A., Bailey, P., Pearen, M. A., Fuller, P. J., Funder, J. W., Simpson, E. R., Leedman, P. J., Tilley, W. D., Brown, M. A., Clarke, C. L., & Muscat, G. E. (2012). Protein arginine methyltransferase 6-dependent gene expression and splicing: Association with breast cancer outcomes. *Endocrine-Related Cancer, 19*, 509–526.

Duch, W., Swaminathan, K., & Meller, J. (2007). Artificial intelligence approaches for rational drug design and discovery. *Current Pharmaceutical Design, 13*, 1497–1508.

Edgar, R., Domrachev, M., & Lash, A. E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research, 30*, 207–210.

Elhady, A. K., Abdel-Halim, M., Abadi, A. H., & Engel, M. (2017). Development of selective Clk1 and -4 inhibitors for cellular depletion of cancer-relevant proteins. *Journal of Medicinal Chemistry, 60*, 5377–5391.

Ficenec, D., Osborne, M., Pradines, J., Richards, D., Felciano, R., Cho, R. J., Chen, R. O., Liefeld, T., Owen, J., Ruttenberg, A., Reich, C., Horvath, J., & Clark, T. (2003). Computational knowledge integration in biopharmaceutical research. *Briefings in Bioinformatics, 4*, 260–278.

Fischer, A. W., Heinze, S., Putnam, D. K., Li, B., Pino, J. C., Xia, Y., Lopez, C. F., & Meiler, J. (2016). CASP11—An evaluation of a modular BCL::Fold-based protein structure prediction pipeline. *PLoS One, 11*, e0152517.

Fiser, A., Do, R. K., & Sali, A. (2000). Modeling of loops in protein structures. *Protein Science, 9*, 1753–1773.

Fleming, N. (2018). How artificial intelligence is changing drug discovery. *Nature, 557*, S55–s57.

Fukuhara, T., Hosoya, T., Shimizu, S., Sumi, K., Oshiro, T., Yoshinaka, Y., Suzuki, M., Yamamoto, N., Herzenberg, L. A., Herzenberg, L. A., & Hagiwara, M. (2006). Utilization of host SR protein kinases and RNA-splicing machinery during viral replication. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 11329–11333.

Gammons, M. V., Fedorov, O., Ivison, D., Du, C., Clark, T., Hopkins, C., Hagiwara, M., Dick, A. D., Cox, R., Harper, S. J., Hancox, J. C., Knapp, S., & Bates, D. O. (2013). Topical antiangiogenic SRPK1 inhibitors reduce choroidal neovascularization in rodent models of exudative AMD. *Investigative Ophthalmology & Visual Science, 54*, 6052–6062.

Gammons, M. V., Lucas, R., Dean, R., Coupland, S. E., Oltean, S., & Bates, D. O. (2014). Targeting SRPK1 to control VEGF-mediated tumour angiogenesis in metastatic melanoma. *British Journal of Cancer, 111*, 477–485.

Gastegger, M., Mcsloy, A., Luya, M., Schütt, K. T., & Maurer, R. J. (2020). A deep neural network for molecular wave functions in quasi-atomic minimal basis representation. *The Journal of Chemical Physics, 153*, 044123.

Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., Mcglinchey, S., Michalovich, D., Al-Lazikani, B., & Overington, J. P. (2012). ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Research, 40*, D1100–D1107.

Gilson, M. K., Liu, T., Baitaluk, M., Nicola, G., Hwang, L., & Chong, J. (2016). BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Research, 44*, D1045–D1053.

Grzybowski, B. A., Szymkuć, S., Gajewska, E. P., Molga, K., Dittwald, P., Wołos, A., & Klucznik, T. J. C. 2018. Chematica: A story of computer code that started to think like a chemist. *4*, 390–398.

Guccione, E., & Richard, S. (2019). The regulation, functions and clinical relevance of arginine methylation. *Nature Reviews. Molecular Cell Biology, 20*, 642–657.

Han, Y., Yang, J., Qian, X., Cheng, W. C., Liu, S. H., Hua, X., Zhou, L., Yang, Y., Wu, Q., Liu, P., & Lu, Y. (2019). DriverML: A machine learning algorithm for identifying driver genes in cancer sequencing studies. *Nucleic Acids Research, 47*, e45.

Harrer, S., Shah, P., Antony, B., & Hu, J. (2019). Artificial intelligence for clinical trial design. *Trends in Pharmacological Sciences, 40*, 577–591.

Hay, M., Thomas, D. W., Craighead, J. L., Economides, C., & Rosenthal, J. (2014). Clinical development success rates for investigational drugs. *Nature Biotechnology, 32*, 40–51.

Hessler, G., & Baringhaus, K. H. (2018). Artificial intelligence in drug design. *Molecules, 23*.

Hughes, J. P., Rees, S., Kalindjian, S. B., & Philpott, K. L. (2011). Principles of early drug discovery. *British Journal of Pharmacology, 162*, 1239–1249.

Hwang, J. W., Cho, Y., Bae, G. U., Kim, S. N., & Kim, Y. K. (2021). Protein arginine methyltransferases: Promising targets for cancer therapy. *Experimental & Molecular Medicine, 53*, 788–808.

Jensen, M. A., Ferretti, V., Grossman, R. L., & Staudt, L. M. (2017). The NCI Genomic Data Commons as an engine for precision medicine. *Blood, 130*, 453–459.

Kalra, G., Peng, Y., Guo, M., & Augsburger, L. L. (2002). A hybrid intelligent system for formulation of BCS Class II drugs in hard gelatin capsules. *Proceedings of the 9th international conference on neural information processing. ICONIP'02, 2002. IEEE*, 1987–1991.

Kędzierska, H., & Piekiełko-Witkowska, A. (2017). Splicing factors of SR and hnRNP families as regulators of apoptosis in cancer. *Cancer Letters, 396*, 53–65.

Keenan, A. B., Jenkins, S. L., Jagodnik, K. M., Koplev, S., He, E., Torre, D., Wang, Z., Dohlman, A. B., Silverstein, M. C., Lachmann, A., Kuleshov, M. V., Ma'ayan, A., Stathias, V., Terryn, R., Cooper, D., Forlin, M., Koleti, A., Vidovic, D., Chung, C., Schürer, S. C., Vasiliauskas, J., Pilarczyk, M., Shamsaei, B., Fazel, M., Ren, Y., Niu, W., Clark, N. A., White, S., Mahi, N., Zhang, L., Kouril, M., Reichard, J. F., Sivaganesan, S., Medvedovic, M., Meller, J., Koch, R. J., Birtwistle, M. R., Iyengar, R., Sobie, E. A., Azeloglu, E. U., Kaye, J., Osterloh, J., Haston, K., Kalra, J., Finkbiener, S., Li, J., Milani, P., Adam, M., Escalante-Chong, R., Sachs, K., Lenail, A., Ramamoorthy, D., Fraenkel, E., Daigle, G., Hussain, U., Coye, A., Rothstein, J., Sareen, D., Ornelas, L., Banuelos, M., Mandefro, B., Ho, R., Svendsen, C. N., Lim, R. G., Stocksdale, J., Casale, M. S., Thompson, T. G., Wu, J., Thompson, L. M., Dardov, V., Venkatraman, V., Matlock, A., Van Eyk, J. E., Jaffe, J. D., Papanastasiou, M., Subramanian, A., Golub, T. R., Erickson, S. D., Fallahi-Sichani, M., Hafner, M., Gray, N. S., Lin, J. R., Mills, C. E., Muhlich, J. L., Niepel, M., Shamu, C. E., Williams, E. H., Wrobel, D., Sorger, P. K., Heiser, L. M., Gray, J. W., Korkola, J. E., Mills, G. B., Labarge, M., Feiler, H. S., Dane, M. A., Bucher, E., Nederlof, M., Sudar, D., Gross, S., et al. (2018). The library of integrated network-based cellular signatures NIH program: System-level cataloging of human cells response to perturbations. *Cell Systems, 6*, 13–24.

Koromina, M., Pandi, M. T., & Patrinos, G. P. (2019). Rethinking drug repositioning and development with artificial intelligence, machine learning, and omics. *Omics, 23*, 539–548.

Kumar, R., Sharma, A., Siddiqui, M. H., & Tiwari, R. K. (2017). Prediction of human intestinal absorption of compounds using artificial intelligence techniques. *Current Drug Discovery Technologies, 14*, 244–254.

Labute, P. (2000). A widely applicable set of descriptors. *Journal of Molecular Graphics & Modelling, 18*, 464–477.

Lamberti, M. J., Wilkinson, M., Donzanti, B. A., Wohlhieter, G. E., Parikh, S., Wilkins, R. G., & Getz, K. (2019). A study on the application and use of artificial intelligence to support drug development. *Clinical Therapeutics, 41*, 1414–1426.

Leinonen, R., Sugawara, H., & Shumway, M. (2011). The sequence read archive. *Nucleic Acids Research, 39*, D19–D21.

Leung, M. K., Xiong, H. Y., Lee, L. J., & Frey, B. J. (2014). Deep learning of the tissue-regulated splicing code. *Bioinformatics, 30*, i121–i129.

Lounkine, E., Keiser, M. J., Whitebread, S., Mikhailov, D., Hamon, J., Jenkins, J. L., Lavan, P., Weber, E., Doak, A. K., Côté, S., Shoichet, B. K., & Urban, L. (2012). Large-scale prediction and testing of drug activity on side-effect targets. *Nature, 486*, 361–367.

Mak, K. K., & Pichika, M. R. (2019). Artificial intelligence in drug development: Present status and future prospects. *Drug Discovery Today, 24*, 773–780.

Maron, M. I., Casill, A. D., Gupta, V., Roth, J. S., Sidoli, S., Query, C. C., Gamble, M. J., & Shechter, D. (2022). Type I and II PRMTs inversely regulate post-transcriptional intron detention through Sm and CHTOP methylation. *eLife, 11*.

Martínez-Montiel, N., Rosas-Murrieta, N. H., Martínez-Montiel, M., Gaspariano-Cholula, M. P., & Martínez-Contreras, R. D. (2016). Microbial and natural metabolites that inhibit splicing: A powerful alternative for cancer treatment. *BioMed Research International, 2016*, 3681094.

Mavrou, A., Brakspear, K., Hamdollah-Zadeh, M., Damodaran, G., Babaei-Jadidi, R., Oxley, J., Gillatt, D. A., Ladomery, M. R., Harper, S. J., Bates, D. O., & Oltean, S. (2015). Serine-arginine protein kinase 1 (SRPK1) inhibition as a potential novel targeted therapeutic strategy in prostate cancer. *Oncogene, 34*, 4311–4319.

Mayr, A., Klambauer, G., Unterthiner, T., & Hochreiter, S. (2016). DeepTox: Toxicity prediction using deep learning. *Frontiers in Environmental Science, 3*, 80.

Mehta, C. H., Narayan, R., & Nayak, U. Y. (2019). Computational modeling for formulation design. *Drug Discovery Today, 24*, 781–788.

Mirza, B., Wang, W., Wang, J., Choi, H., Chung, N. C., & Ping, P. (2019). Machine learning and integrative analysis of biomedical big data. *Genes (Basel), 10*.

Moreira, G. A., Lima, G. D. A., Siqueira, R. P., Barros, M. V. A., Adjanohoun, A. L. M., Santos, V. C., Barbosa, É. A. A., Loterio, R. K., Paiva, J. C., Gonçalves, V. H. S., Viol, L. C. S., Marques-Da-Silva, E. A., Júnior, A. S., Almeida, M. R., Fietto, J. L. R., Machado-Neves, M., Ferreira, R. S., Teixeira, R. R., & Bressan, G. C. (2018). Antimetastatic effect of the pharmacological inhibition of serine/arginine-rich protein kinases (SRPK) in murine melanoma. *Toxicology and Applied Pharmacology, 356*, 214–223.

Oltean, S., & Bates, D. O. (2014). Hallmarks of alternative splicing in cancer. *Oncogene, 33*, 5311–5318.

Öztürk, H., Özgür, A., & Ozkirimli, E. (2018). DeepDTA: Deep drug-target binding affinity prediction. *Bioinformatics, 34*, i821–i829.

Park, K. (2019). A review of computational drug repurposing. *Translational and Clinical Pharmacology, 27*, 59–63.

Parkinson, H., Kapushesky, M., Shojatalab, M., Abeygunawardena, N., Coulson, R., Farne, A., Holloway, E., Kolesnykov, N., Lilja, P., Lukk, M., Mani, R., Rayner, T., Sharma, A., William, E., Sarkans, U., & Brazma, A. (2007). ArrayExpress—A public database of microarray experiments and gene expression profiles. *Nucleic Acids Research, 35*, D747–D750.

Pereira, J. C., Caffarena, E. R., & Dos Santos, C. N. (2016). Boosting docking-based virtual screening with deep learning. *Journal of Chemical Information and Modeling, 56*, 2495–2506.

Persidis, A. J. D. D. W. (2011). The benefits of drug repositioning. *12*, 9–12.

Powles, J., & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health Technology (Berlin), 7*, 351–367.

Pu, L., Naderi, M., Liu, T., Wu, H. C., Mukhopadhyay, S., & Brylinski, M. (2019). eToxPred: A machine learning-based approach to estimate the toxicity of drug candidates. *BMC Pharmacology and Toxicology, 20*, 2.

Rantanen, J., & Khinast, J. (2015). The future of pharmaceutical manufacturing sciences. *Journal of Pharmaceutical Sciences, 104*, 3612–3638.

Rose, P. W., Prlić, A., Altunkaya, A., Bi, C., Bradley, A. R., Christie, C. H., Costanzo, L. D., Duarte, J. M., Dutta, S., Feng, Z., Green, R. K., Goodsell, D. S., Hudson, B., Kalro, T., Lowe,

R., Peisach, E., Randle, C., Rose, A. S., Shao, C., Tao, Y. P., Valasatava, Y., Voigt, M., Westbrook, J. D., Woo, J., Yang, H., Young, J. Y., Zardecki, C., Berman, H. M., & Burley, S. K. (2017). The RCSB protein data bank: Integrative view of protein, gene and 3D structural information. *Nucleic Acids Research, 45*, D271–d281.

Rupp, M., Körner, R., & Tetko, I. V. (2010). Estimation of acid dissociation constants using graph kernels. *Molecular Informatics, 29*, 731–740.

Russo, D. P., & Zhu, H. (2022). High-throughput screening assay profiling for large chemical databases. *Methods in Molecular Biology, 2474*, 125–132.

Schütt, K. T., Gastegger, M., Tkatchenko, A., Müller, K. R., & Maurer, R. J. (2019). Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions. *Nature Communications, 10*, 5024.

Segler, M. H. S., Preuss, M., & Waller, M. P. (2018). Planning chemical syntheses with deep neural networks and symbolic AI. *Nature, 555*, 604–610.

Sellwood, M. A., Ahmed, M., Segler, M. H., & Brown, N. (2018). Artificial intelligence in drug discovery. *Future Medicinal Chemistry, 10*, 2025–2028.

Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Žídek, A., Nelson, A. W. R., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D. T., Silver, D., Kavukcuoglu, K., & Hassabis, D. (2020). Improved protein structure prediction using potentials from deep learning. *Nature, 577*, 706–710.

Stanescu, A., Tangirala, K., Caragea, D., & Bioinformatics. (2016). Predicting alternatively spliced exons using semi-supervised learning. *International Journal of Data Mining, 14*, 1–21.

Teng, T., Tsai, J. H., Puyang, X., Seiler, M., Peng, S., Prajapati, S., Aird, D., Buonamici, S., Caleb, B., Chan, B., Corson, L., Feala, J., Fekkes, P., Gerard, B., Karr, C., Korpal, M., Liu, X., Lowe, J. T., Mizui, Y., Palacino, J., Park, E., Smith, P. G., Subramanian, V., Wu, Z. J., Zou, J., Yu, L., Chicas, A., Warmuth, M., Larsen, N., & Zhu, P. (2017). Splicing modulators act at the branch point adenosine binding pocket defined by the PHF5A-SF3b complex. *Nature Communications, 8*, 15522.

Wan, F., & Zeng, J. M. (2016). Deep learning with feature embedding for compound-protein interaction prediction. bioRxiv, 086033.

Wang, E., & Aifantis, I. (2020). RNA splicing and cancer. *Trends Cancer, 6*, 631–644.

Wang, Z., Jensen, M. A., & Zenklusen, J. C. (2016). A practical guide to The Cancer Genome Atlas (TCGA). *Methods in Molecular Biology, 1418*, 111–141.

Waring, M. J., Arrowsmith, J., Leach, A. R., Leeson, P. D., Mandrell, S., Owen, R. M., Pairaudeau, G., Pennie, W. D., Pickett, S. D., Wang, J., Wallace, O., & Weir, A. (2015). An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nature Reviews. Drug Discovery, 14*, 475–486.

Weininger, D. J. J. O. C. I., & Sciences, C. (1988). SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *28*, 31–36.

Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., Pon, A., Knox, C., & Wilson, M. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research, 46*, D1074–D1082.

Xiong, H. Y., Alipanahi, B., Lee, L. J., Bretschneider, H., Merico, D., Yuen, R. K., Hua, Y., Gueroussov, S., Najafabadi, H. S., Hughes, T. R., Morris, Q., Barash, Y., Krainer, A. R., Jojic, N., Scherer, S. W., Blencowe, B. J., & Frey, B. J. (2015). RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science, 347*, 1254806.

Yang, X., Wang, Y., Byrne, R., Schneider, G., & Yang, S. (2019). Concepts of artificial intelligence for computer-assisted drug discovery. *Chemical Reviews, 119*, 10520–10594.

Yang, Z., Xiao, T., Li, Z., Zhang, J., & Chen, S. (2022). Novel chemicals derived from Tadalafil exhibit PRMT5 inhibition and promising activities against breast cancer. *International Journal of Molecular Sciences, 23*.

Yu, H., Chen, J., Xu, X., Li, Y., Zhao, H., Fang, Y., Li, X., Zhou, W., Wang, W., & Wang, Y. (2012). A systematic prediction of multiple drug-target interactions from chemical, genomic, and pharmacological data. *PLoS One, 7*, e37608.

Zang, Q., Mansouri, K., Williams, A. J., Judson, R. S., Allen, D. G., Casey, W. M., & Kleinstreuer, N. C. (2017). In silico prediction of physicochemical properties of environmental chemicals using molecular fingerprints and machine learning. *Journal of Chemical Information and Modeling, 57*, 36–49.

Zhang, Q., Di, C., Yan, J., Wang, F., Qu, T., Wang, Y., Chen, Y., Zhang, X., Liu, Y., Yang, H., & Zhang, H. (2019). Inhibition of SF3b1 by pladienolide B evokes cycle arrest, apoptosis induction and p73 splicing in human cervical carcinoma cells. *Artificial Cells Nanomedicine and Biotechnology, 47*, 1273–1280.

Zhang, L., Tan, J., Han, D., & Zhu, H. (2017). From machine learning to deep learning: Progress in machine intelligence for rational drug discovery. *Drug Discovery Today, 22*, 1680–1685.

Zhao, L., Wang, W., Sedykh, A., & Zhu, H. (2017). Experimental errors in QSAR modeling sets: What we can do and what we cannot do. *ACS Omega, 2*, 2805–2812.

Zhu, H. (2020). Big data and artificial intelligence modeling for drug discovery. *Annual Review of Pharmacology and Toxicology, 60*, 573–589.

Zhu, H., & Xia, M. (2016). *High-throughput screening assays in toxicology*. Springer.

Chapter 12
Applying Artificial Intelligence Prediction Tools for Advancing Precision Oncology in Immunotherapy: Future Perspectives in Personalized Care

Botle Precious Damane, Zilungile Lynette Mkhize-Kwitshana, Mahlatse Cordelia Kgokolo, Thifhelimbilu Luvhengo, and Zodwa Dlamini

Abstract Artificial intelligence (AI) has revolutionized the medical field more evidently in managing and treating cancer as it continues to be a global burden. Artificial intelligence has been imperative in the screening and detection of cancer for decades. As technology evolves, AI has gone as far as predicting the risk of cancer development or recurrence. More profoundly, machine learning (ML) can advance individualized cancer therapy (precision medicine) through molecularly

B. P. Damane (✉)
Department of Surgery, Steve Biko Academic Hospital, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
e-mail: precious.setlai@up.ac.za

Z. L. Mkhize-Kwitshana
Department of Medical Microbiology, School of Laboratory Medicine & Medical Sciences, Medical School Campus, College of Health Sciences, University of KwaZulu-Natal-Natal, Durban, South Africa

SAMRC Research Capacity Development Division, South African Medical Research Council, Tygerberg, Cape Town, South Africa
e-mail: mkhizekwitshanaz@ukzn.ac.za

M. C. Kgokolo
Department of Dermatology, Bridge A, Level 9, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: mahlatse.kgokolo@up.ac.za

T. Luvhengo
Department of Surgery, Charlotte Maxeke Johannesburg Academic Hospital, University of the Witwatersrand, Johannesburg, South Africa
e-mail: Thifhelimbilu.Luvhengo@wits.ac.za

Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: zodwa.dlamini@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_12

targeted therapies or cancer immunotherapy. Machine learning algorithms are utilized to develop models that can monitor treatment response and resistance to cancer immunotherapy. However, the ability of ML to deal with large data sets accurately and precisely requires careful construction and articulation of prediction models. This chapter will discuss some of the models which are used to identify diagnostic and predictive biomarkers for stratification of cancer patients who will respond positively to immunotherapy and/or stratification based on the likelihood of developing resistance to treatment. Furthermore, the chapter will address how AI plays a role in advancing precision medicine in cancer immunotherapy.

Keywords Artificial intelligence · Precision medicine · Cancer immunotherapy · Prediction model · Machine learning · Immune checkpoints

12.1 Introduction

Cancer immunotherapy refers to the manipulation of the patient's immune system to fight cancer (Waldman et al., 2020). Precision medicine refers to the identification of a specific molecular biomarker of a disease and using the same marker to target it. Precision medicine is selective and thus keeps healthy cells and the surrounding tissue unharmed as it avoids the traditional one-size-fits-all approach when using chemotherapy and radiation (Lewis & Yap, 2020). In the broader sense of the concept, precision medicine refers to customized treatment practices for specific groups of individuals or specific tumors or disease profiles. Artificial intelligence can monitor the use of cancer immunotherapy, predict patient tolerance, and optimize treatment response. Prediction models which are generated through AI provide information about the patient's health status and assist with diagnosis, patient outcome, and identification of prognostic factors. This information is based on the input data such as the patient's symptoms, environment, tissue or anatomical imaging, and molecular biomarkers (de Hond et al., 2022). Subtypes of AI include machine learning (ML), deep learning (DL), robotics, computer vision, natural language processing, and a lot more (Liang et al., 2020; Sarker, 2021), however not all AI tools are applicable instantaneously (Shao et al., 2021).

The development of neural networks comes with the disadvantage of having a multitude of ineffectual data that need to be filtered to create a workable lucrative system. Several ML classifier algorithms including artificial neural networks (ANNs), support vector machines (SVMs), and decision trees (DTs) can accurately predict the existence of cancer and/or prognosis (Shaikh & Rao, 2022). Deep learning is capable of processing input of medical images to solve complex data involving classification, segmentation, and image texture. Deep learning comes with neural networks that can process high-quality output images fast and efficiently. Deep learning programs include deep neural networks (DNNs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs) (Shao et al., 2021). The deep learning program is useful for monitoring of the response to treatment inclusive of immunotherapeutic agents in patients diagnosed with cancer. Trebeschi et al.

designed the prognostic AI-monitor (PAM) that can accurately predict the overall survival of patients during treatment with immune checkpoint inhibitors (ICs) (Trebeschi et al., 2021). This chapter will cover the existing AI tools that are used to predict therapeutic responses to cancer immunotherapy. The feasibility of the AI tools will be addressed as well as the measures taken to improve the efficacy of cancer immunotherapy using AI tools.

12.1.1 Cancer Immunotherapy

Exposure to environmental carcinogens (Koual et al., 2020) or genetic predispositions (Wang, 2016) that contribute to cancer development can be fought off at the cellular level by the immune system until such time that cancer cells manage to escape immunoediting and grow beyond control. Cancer cells use specific escape mechanisms which include manipulation of the immune system to recognize cancer cells as part of the normal body cells. Similar mechanisms explain how immunotherapy is used to manipulate and heighten the host immune system to enable it to identify and destroy cancer cells (Suresh & O'Donnell, 2021). Adverse events (AEs) seen with cancer immunotherapy are due to heightened immune responses which are intended to elicit anticancer immune responses. The type and severity of AEs depend on the type of immunotherapy used and can be mild or even life-threatening (Barber, 2019).

12.1.1.1 The Efficacy of Cancer Immunotherapy

Key to the success of cancer immunotherapy is its ability to target cancer cells while preserving normal cells. The effectiveness of immunotherapy has improved significantly over the years with remission reported in some of the patients (Ventola, 2017). The success of immunotherapy has led to the Food and Drug Administration (FDA) approving more immunotherapeutic drugs for cancer treatment particularly immune checkpoint inhibitors, especially for hard-to-treat cancers (Twomey & Zhang, 2021). Table 12.1 categorizes and highlights some of the FDA-approved therapies and their mechanisms of action.

Immunotherapy has its pros and cons. The variations in treatment response between different cancers and individuals with the same cancer can be due to lack of specificity, the ability of anticancer immune cells to recognize and destroy cancer cells, and drug resistance. Moreover, tumor mechanisms such as the dense tumor microenvironment which may prevent drug delivery to the core of the tumor, the reliance on the individual's immunocompetence, and the involvement of microbiota all contribute to immunotherapeutic response (Sambi et al., 2019). Similar to all other treatment strategies, challenges with immunotherapy include side effects such as overstimulation of the immune response (cytokine release syndrome). This aspect, however, falls outside the scope of this chapter and will not be dealt with in detail.

Table 12.1 Categorization of cancer immunotherapies

FDA-approved drug	Mode of Action	Reference
Immunotherapy: Vaccines		
Sipuleucel-T (Provenge)	Dendritic cells (DCs) activate T cells to target prostatic acid phosphatase.	Anassi and Ndefo (2011), Ventola (2017), DeMaria and Bilusic (2019)
TICE (BCG vaccine)	BCG activates T cells immune response and promotes phagocytosis of cancer cells by macrophages. BCG cell wall surface marker and glycosaminoglycan on the surface of normal urothelial cells bear negative charges that make it challenging for BCG to enter normal cells. Instead, BCG gains access preference to positively charged cancer cells.	Han et al. (2020)
Immunotherapy: Immunomodulators		
Cytokines		
IFN-α	Enhances MHC class I expression, mediates maturation of DCs, and activates effector cells.	Lee and Margolin (2011)
IL-2	IL-2 signaling forms an integral part of regulatory T cells' (Tregs) suppressive function. Tregs block the anti-tumor functional capacity of cytotoxic T cells.	
Immune checkpoint inhibitors:		
Multiple (listed in references)	Generally, CTLA-4 and PD-1 are normal co-inhibitors that downregulate an immune response to maintain immune homeostasis and can thus block anti-tumor T cells immune responses. CTLA-4 and PD-1 inhibitors act by halting the activities of these molecules thereby enabling anti-tumor immune responses to continue. Otherwise, there is a list of drugs developed to inhibit these immune checkpoints with each having a specific side and mode of action.	Vaddepally et al. (2020), Twomey and Zhang (2021)
Agonists:		
Mostly under clinical trials	Activates T cells and induces the expression of costimulatory molecules resulting in inflammatory responses.	Kaczanowska et al. (2013), Choi et al. (2020)
Immunotherapy: Adoptive cell therapies		
Chimeric antigen receptor (CAR) T-cell therapies	CAR-Ts are genetically modified autologous T cells that specifically target and kill cancer cells. First, autologous circulating T lymphocytes will be isolated by leukapheresis. Then lymphocytes will be genetically engineered with chimeric tumor-specific antigens and these tumor-infiltrating lymphocytes (TIL) are expanded and transferred back into the same patient as a form of therapy.	Rohaan et al. (2019), Fischer and Bhattarai (2021)

(continued)

Table 12.1 (continued)

FDA-approved drug	Mode of Action	Reference
Immunotherapy: Oncolytic virus therapy		
T-VEC	Similar to BCG, T-VEC promotes lysis of cancer cells and activation of immune T cells responses in the tumor microenvironment. Furthermore, activation of T cells via DCs is facilitated by GM-CSF inserted within the virus.	DeMaria and Bilusic (2019), Ferrucci et al. (2021)
Immunotherapy: Monoclonal antibodies		
Multiple (listed in references)	Monoclonal antibodies exhibit a variety of anticancer mechanisms such as enhanced identification of cancer cells, improvement of cytotoxicity, blocking cancer cell growth, promotion of apoptosis, specific delivery of radio- or chemotherapy and blocking angiogenesis	Zahavi and Weiner (2020), Coulson et al. (2014)

12.1.2 AI and Biomarker Prediction Tools

12.1.2.1 Identification of Genomic Immune Signatures

The response of cancer to systemic treatment is influenced by its stage and molecular profile. Hepatocellular carcinoma is a heterogeneous disease with a divergent response to local and systemic therapy. Prediction of therapeutic response by identification of immune profiles that can be utilized as clinically actionable biomarkers in hepatocellular carcinoma patients was determined using DL prediction tools. Upregulated immune signatures were identified using The Cancer Genome Atlas (TCGA). Immune signatures were selected based on their known ability to improve immunotherapeutic response. Training of CNN was performed based on these immune signatures and digital histopathology (Fig. 12.1).

Patch-based model is used to correct the background noise caused by a random signal that interferes with data acquisition. This noise could occur during transmission or due to the quality of input data (Alkinani & El-Sakka, 2017). For the patch-based model (Fig. 12.2), 500 patches were chosen from the slides. Patches were predicted using pre-trained ShuffleNet and categorized into high or medium/low clusters. Hierarchical clustering was performed using the Ward2 algorithm and Euclidean distance. In classic multiple-instance learning, bags of unlabeled instances were considered positive if a bag contained one or more positive instances (cluster high) and negative if all instances were negative. The model was able to predict digital histological images outside of the region of interest (ROI) which was accomplished by combining annotations from the most predictive patches. Clustering-constrained attention multiple-instance learning (CLAM) is another method that was used in the study to identify patches that were of clinical significance. The identified patches were then used to make predictions of the whole slide

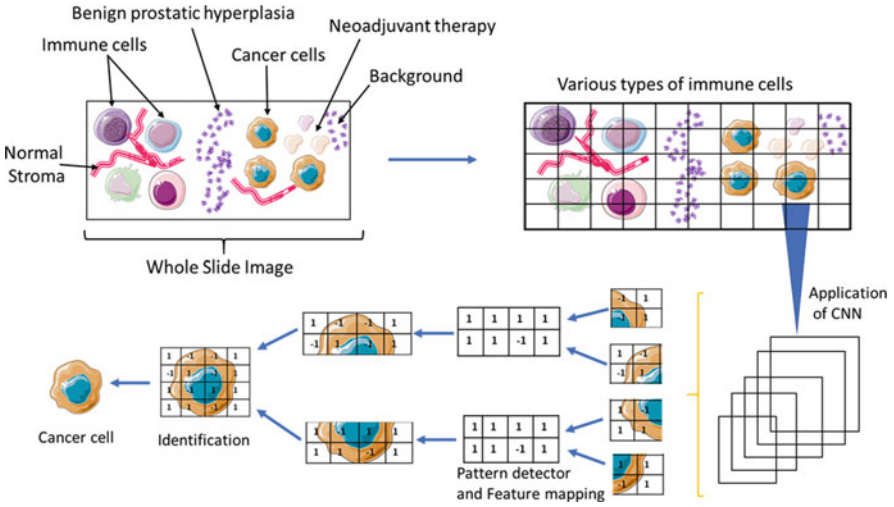


Fig. 12.1 Simplified illustration of identification of immune cells on a whole slide image by digital pathology using deep learning prediction tools. When training a classifier to identify an image, a neural network will create a simplified way to recognize the image by creating numerical patterns. These patterns are calculated and filtered into a feature map which will allow the placement of specific features of a cell until accurate identification is achieved

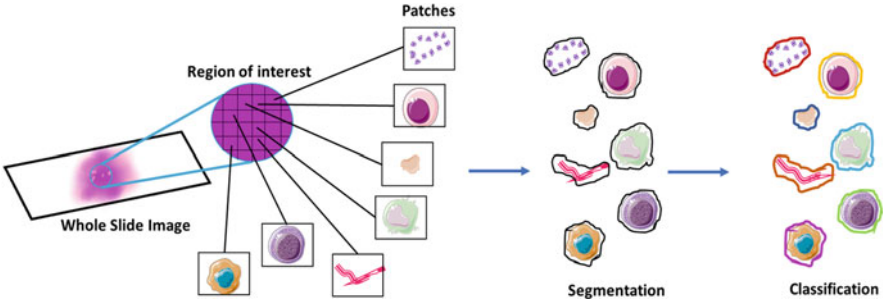


Fig. 12.2 Classification of immunomorphological features on a whole slide image. An ROI will first be selected. This region encompasses characteristics that represent the entire image. Boundaries are created around each cell. Segmentation is performed to separate each cell according to its boundaries. A classifier will separate each cell according to morphological features and characteristics. Here the computer is trained to differentiate types of immune cells. These include the ability to isolate immune parameters from morphological features such as the stroma, cells dying from neoadjuvant therapy, cancerous and normal cells

images (WSI). The AI methods allowed the authors to identify the highest expression of immune gene signatures that were likely to correlate with a subset of hepatocellular carcinoma (HCC) cells that would most probably be sensitive to cancer immunotherapy. The CLAM model was able to accurately predict tumor

immune clusters in 38/42 cases that would benefit from targeted immunotherapy (Zeng et al., 2022).

12.1.2.2 Long Noncoding RNAs as Prognostic Markers

Another set of biomarker tools are long noncoding RNAs (lncRNAs) which have been used as predictive markers of overall survival in several cancers. Novel immune-related lncRNA signatures are becoming more popular as prognostic or overall survival markers in bladder (Ren et al., 2022), breast (Ma et al., 2020), nasopharyngeal (Liang et al., 2022), lung (Zhou et al., 2022), and colon (Lin et al., 2020) cancers. Although several authors have confirmed the role of immune-related lncRNA signature (IRLS) as predictive biomarkers, IRLS have limited predictive value in colorectal cancers. High-risk patients have to undergo chemotherapy while low-risk patients show improved responses to immunotherapy such as bevacizumab. The selection of patients suitable for immune checkpoint therapy is based on several parameters including expression of PD-L1, tumor mutation burden, and identification of defective mismatch repair markers such as microsatellite instability (MSI) or deficient DNA mismatch repair (dMMR) which are not always reliable or accurate.

There is a plethora of identified gene expression markers that are not clinically actionable due to multiple factors including the need for consistent and reliable ML tools. Data sets from multiple studies with clinical and molecular information can be collected and ML applied to detect specific IRLS that can be used as reliable predictive markers for colorectal cancer patients who will benefit from myeloid-derived suppressor cells (MDSCs) immunotherapy. From 235 IRLS identified, 43 were prognostic. These were subjected to ML-based integration to get a consensus. The TCGA-CRC data set (101 prediction models) was subjected to Leave-One-Out Cross-Validation (LOOCV) framework. The use of lasso cox followed by stepwise cox identified 16 IncRNAs. This set was used to determine patient risk scores and overall survival (Liu et al., 2022).

12.1.2.3 MicroRNAs as Prognostic Markers

The use of anti-PD-1 as a predictive marker for cancer immunotherapeutic outcome is limited to 30–50% of cases which could be due to the heterogeneous cell density of the tumor microenvironment. Machine learning can be employed to handle the heterogeneous data collected from the tumor microenvironment and develop decision-making AI tools to support precision oncology and personalized care. The microRNA (miRNA) signature classifier (MSC) has been validated as an independent prognostic indicator in lung cancers and an ML tool was developed to predict the efficacy and therapeutic response to cancer immunotherapy in non-small cell lung cancer (NSCLC). Data on patient demographics, medical history, tumor stage, PD-L1 expression, molecular diagnosis, radiology, patient management, and therapeutic response was combined to develop the predictive algorithm. The MSC

algorithm consisted of low, intermediate, high, and hemolyzed categories. The hemolyzed category was thought to be due to the presence of microRNA in the blood cells (Prelaj et al., 2022).

Recognition of blood miRNA is validated by the study done by Rajakumar et al. who identified five miRNA signatures for prediction of the overall survival of patients with advanced NSCLC on anti-PD-L1 therapy from whole blood miRNA profiling (Rajakumar et al., 2022). The authors further assessed if the revised miRNA risk (miRisk) score could predict patient response to chemoimmunotherapy. Patients selected for the study had high expression levels of PD-1 ligand (≥50%) pre-immunotherapeutic treatment. The miRisk score was stratified according to the median risk score within the low-risk $\leq -0.0725 <$ high-risk training cohort. The low miRisk cohort had significantly increased overall survival which did not correlate with that of combinatorial therapy suggesting that immunotherapeutic markers cannot be used as predictive markers for combinatorial therapy, specifically with chemotherapy (Rajakumar et al., 2022).

The most relevant feature for ML training was selected based on the available literature and clinical experience paired with a linear correlation higher than 0.8. Only one of the features was chosen. The classification was performed using a feedforward neural network (FFNN), logistic regression (LR), K-nearest neighbors (K-NN), support vector machines (SVM), and random forest. The best-suited feature was selected using the Akaike information criterion (AIC). The model that was selected based on performance for predicting therapeutic response was the LR. Features included functional status as measured by the Eastern Cooperative Oncology Group (ECOG) performance status, immunotherapy, line of therapy, the neutrophil-to-lymphocyte ratio (NLR), the MSC test, and PD-L1. The model could accurately predict treatment response and patient survival. The limitation of this study was the exclusion of the radiogenomics and sample size (Prelaj et al., 2022).

12.1.2.4 Radiomics as Therapeutic Response Monitoring Tools

A multi-cohort study using a radiomics-based biomarker of tumor-infiltrating CD8 cells in patients on anti-PD-1 monotherapy was performed to assess predictive immunotherapeutic response in multiple solid cancers. A radiomic signature profile was identified with high contrast computerized tomography (CT) scan images and RNA-sequencing genomic data (CD8B gene). The radiomic signature was validated on the TGCA database. The CD8 cell expression signature was constructed using the elastic-net regularized regression method. A high radiomic score was correlated with an increased objective response rate (ORR) at 3 months compared to patients with progressive disease or stable disease. These results were also associated with improved patient overall survival (Sun et al., 2018).

Yang et al. utilized a multi-omics DL approach to predict response to cancer immunotherapy (anti-PD-1/PD-L1) in NSCL cancer. Data collection included radiomics, laboratory, and baseline clinical data. A unified DL tool was used to integrate this data and multimodal serial data from CT scans. The CT scan data was

grouped according to RECIST. A group of responders was based on the reports that showed complete or partial treatment response and non-responders were patients who had progressive disease. Radiomics was used to analyze radiographic features. A DL model with Simple Temporal Attention (SimTA) modules was developed to process the asynchronous clinical time series between radiomic and blood test features. A 60-day ($SimTA_{60d}$) and 90-day ($SimTA_{90d}$) response models were developed. The RNN DL models using baseline PD-L1 expression, blood profile, and radiomics were used to validate the use of SimTA modules. The data obtained was integrated with clinical data. All these data were incorporated into a multiple layer perceptron (MLP) structure for the prediction of responders and non-responders. The neural network was trained for classification purposes and optimized with an Adam optimizer. The module was able to correctly group responders and non-responders; however, the RNN model was not as good in terms of distinguishing responders from non-responders. The study also indicated that low-risk patients had significantly improved progression-free survival (PFS) and overall survival than high-risk patients (Yang et al., 2021).

Trebeschi et al. [8] developed a prognostic AI-monitor (PAM) model to monitor immunotherapeutic response in metastatic urothelial cancer patients. Whole body information was collected from the chest and abdominal CT scans which also provided an overview of disease complications and patients' immune response to immunotherapy. The AI tool, PAM consisted of three modules. The localizer module used CNN trained to focus on the chest and the abdomen in two separate images and the last two modules were named trackers. These modules used CNN with instances trained for chest and abdominal imaging, respectively, for analysis of morphological changes. Trackers were designed to match anatomical landmarks and shapes of two 3D radiological images and quantify anatomical differences between them (Trebeschi et al., 2021).

12.1.2.5 Other Approaches

Paracrine-regulated cross-talk between various cell types, including immune and tumor cells, and cell–matrix interactions may be involved in promoting tumor progression (Place et al., 2011; Quail & Joyce, 2013; Burkholder et al., 2014; Ungefroren et al., 2011). The cell–matrix interactions can be manipulated to enhance treatment and improve patient clinical outcomes (Bracci et al., 2014; Beatty & Gladney, 2015). Wang and colleagues developed an image classifier that could quantify TIL and analyze morphological features of the image to assess the interaction between cancer and immune cells. The aforementioned characteristics were analyzed in comparison with patient clinical outcomes and CNN was used for nuclei segmentation. A watershed and feature-based approach was used to detect TIL and cancer nuclei and thereafter, CNN was used for the characterization of the normal stroma and epithelial tissue. High infiltration of lymphocytes and cancer cells was associated with poor prognosis. The groups were separated based on lymphocytic and non-lymphocytic clusters. The density and spatial closeness between

lymphocytes and non-lymphocytic cells were used as quantitative image features. These parameters could be used in predicting which patients will respond well to cancer immunotherapy (Wang et al., 2022).

12.2 Integration of AI Tools in the Enhancement of Cancer Immunotherapies

The response to immune checkpoint inhibitors is predicted based on the detection of specific IC (e.g., PD-L1) associated with the disease but patients do not always benefit from treatment with ICs. There is variability in PD-L1 expression amongst patients with the same cancer, patients with different cancer types, and the frequency of tumor-infiltrating immune cells (Davis & Patel, 2019). To overcome the variability of PD-L1 expression, adjuvants are used to improve the efficacy of cancer vaccines by heightening specific anticancer immune responses. For example, Cervarix vaccine is utilized as a preventative measure against human papillomavirus (HPV) 16 and 18 types. Approximately 70% of cervical cancers are attributable to HPV 16 and 18 (Ahmed et al., 2017; Dubensky & Reed, 2010). Cervarix vaccine is formulated with a toll-like receptor (TLR)-4 targeted adjuvant known as monophosphoryl lipid A (Dubensky & Reed, 2010). The TLRs modulate immune responses in multiple ways which include reduction of cancer favoring immunosuppressive effect of regulatory T cells (Han et al., 2019), MDSCs, tumor-associated macrophages (TAMs) (Liu et al., 2020), and stimulation of neutrophils (Hayashi et al., 2003) which may promote metastatic spread of cancer (Xiong et al., 2021). TLRs can also increase cancer immune responses by activation of DCs, hence the use of the CpG vaccine adjuvant which activates DCs through TLR9 (Chen et al., 2022).

Chaudhury et al. combined immune response profiles with data integration and ML to investigate how adjuvants modulate vaccine-induced immune responses. Three liposomal formulations, namely Alum (ALFA) with or without lQS21 (ALFQ) were used and a multivariate prediction model was developed based on their effect on the immune response. The self-assembling protein nanoparticles (SAPN) vaccine (FMP014) was administered to rhesus monkeys. The levels of CD4+ and C8+ T cells and cytokine profiles were measured at different time points and the effect of vaccination was accessed at each time point. Immune responses were based on the difference with adjuvant and not antigen dose-related immune responses and so most of the measurements were vaccine-induced. A difference in the immune response of ALFA and ALFQ animal groups was observed. A random forest model was used to categorize immune profiles between these two groups. An ML tool was trained to analyze the adjuvant used by accessing its immune profile. The model predicted which animals received a vaccine with ALFA or ALFQ as an adjuvant with 92% accuracy (Chaudhury et al., 2018).

The development of cancer vaccines depends on the ability of immune cells to recognize tumor antigens and facilitate their destruction by antigen-presenting cells such as DCs. Adoptive cell therapies are based on the augmentation of anticancer immune responses by harvesting tumor antigen-exposed tumor-infiltrating lymphocytes (TILs). Precision and specific identification of tumor antigens by human leukocyte antigen (HLA) molecules are central to the efficacy of these immunotherapies. The HLA epitope prediction algorithms that have been utilized thus far are based on the binding of the peptide–HLA complex. However, these models have not been successful in predicting HLA presentation. Bulik-Sullivan developed an HLA class I epitope prediction algorithm that could accurately predict HLA presentation crucial to the development of personalized immunotherapies by using data sets collected from mass spectrophotometry. A neural network model, the epitope discovery in cancer genomes (EDGE) was trained from the data set of tissue samples with paired class I HLA peptide sequences, HLA types, and transcriptome sequencing. The HLA peptide presentation was directly proportional to the mRNA expression of their source gene. Peptides were considered positively labeled if they were identified using mass spectrometry and negative if identified via the reference proteome not detectable by mass spectrometry. From here data sets were categorized into validation and testing sets. The model accurately identified HLA alleles based on gene RNA expression and gene-specific presentation propensity (Bulik-Sullivan et al., 2019).

12.2.1 AI Tools for the Prediction of Novel Immune-Related Adverse Events

It is of outmost importance that healthcare providers ensure that the risk of developing side effects is minimal, particularly in patients who are less likely to benefit from cancer immunotherapy. Some patients falling under the category of responders might still develop side effects. Kichloo et al. reviewed systemic adverse effects caused by cancer immunotherapy and grouped them according to specific organ toxicity which ranged from cardiotoxicity mainly attributed to ICs and chimeric antigen receptor (CAR) T-cell therapy to dermatological, pancreatic, endocrine, and gastrointestinal problems and possibly more. It is predicted that with more discoveries of cancer immunotherapeutic agents, understanding of their mode of action, and the information provided by patients on therapy, the benefits of cancer immunotherapy will eventually outweigh the risks (Kichloo et al., 2021).

Martins et al. noted that ICs have not only reshaped the traditional therapeutic golden standard-of-care approaches but have also brought new hope to patients. This statement is supported by real stories of patients who are in remission and remain in remission years after they have stopped cancer immunotherapy. However, a new spectrum of adverse effects has been reported with characteristics that have not been observed with chemotherapy. With the discovery of new immunotherapeutic agents,

new and rare adverse effects might surface (Martins et al., 2019). Clinicians should therefore be prepared and find ways to learn about immune-related adverse events (irAEs) in advance so they can act timeously and accordingly.

Having to detect the possibility of patients developing irAEs could save some of the patients from the additional psychological strain caused by a cancer diagnosis and irAEs which would more likely have a positive impact on the patient's quality of life and overall survival. Machine learning can be used to determine the possible development of irAEs from data collected from previously reported clinical outcomes post-cancer immunotherapeutic immune response. To predict whether a patient has irAEs (presence) or the odds of a patient developing irAEs (onset), ML binary classification model was used. A questionnaire with 18 monitored symptoms was collected from cancer patients receiving ICs using a digital platform and then data sets were collected from a treating clinician with records of symptoms from when treatment was initiated and terminated. These included the type of toxicity and the location thereof.

Several ML models were applied including logistic regression, elastic-net regression, support vector machines, light gradient boosting machine (LightGBM), and random forest. However, these were excluded except for python library extreme gradient boosting (XGBoost) which gave the best fit for the objectives of the study. The input data for the first model contained the electronic patient-reported outcomes (ePRO) data and was trained to detect the presence of irAEs. The second model was trained to detect the onset of irAEs. The developed model was able to accurately predict the presence of irAEs but was not able to predict the onset of irAEs (Iivanainen et al., 2021).

An increase in cancer progression (hyperprogressive disease) has been observed in patients on ICs. Hyperprogressive disease is associated with many factors including age, genetic mutations, and metastatic disease. Identification and stratification of patients at risk of developing hyperprogressive disease will be an additional important tool in clinical decision-making (Vaidya et al., 2020). An increase in quantitative vessel tortuosity (QVT) serves as an indication of a positive response to treatment with ICs. Baseline contrast CT scans were used to study tumor vasculature and correlated it with overall survival and predictive response to ICs in lung cancer patients. RECIST was used to categorize patients into responders and non-responders. Using an unsupervised clustering algorithm, two survival groups were identified. This allowed for successful prediction of responders and improvement of overall survival (Alilou et al., 2019).

Patients with hyperprogressive disease will have distinct radiomics features that can be used to develop predictive AI tools useful for the stratification of advanced NSCL patients on ICs therapy who are at risk of developing hyperprogressive disease. After RECIST categorization, the effect of immunotherapy on tumor size per unit time (months) was used to determine the hyperprogressive disease. Unsupervised clustering (heatmaps and K-mean) was performed on radiomics features and a supervised algorithm was applied to identify hyper-responders. The random forest, linear discriminant analysis, diagonal linear discriminant analysis, quadratic discriminant analysis, and support vector machine were tested as training

models. Ultimately the best fit model was a random forest classifier which was also used for the validation set. Prediction of hyper-responders also revealed that this group of patients had lower overall survival than the other groups (Vaidya et al., 2020).

12.2.2 Implementation of AI Tools for Monitoring Patient Compliance to Cancer Immunotherapy

The benefits of cancer immunotherapy include the ability of the body to regain effective immunoediting processes that recognize and destroy cancer cells. The patient immune system regains memory of cancer cells and develops mechanisms that keep these cells at bay until such time that the patient is in remission. The immune system is armed in such a way that recurrence is limited/prevented. However, these benefits can be hampered by patient's poor adherence to treatment or hospital visits. A concern that is managed while the patient is still hospitalized. Non-compliance by patients is attributed to several factors including not taking the medication at prescribed times, stopping medications once the patient feels better or due to certain side effects such as nausea, diarrhea and fatigue, taking incorrect dosage, forgetfulness, inconvenience, and family support. Non-compliance has a significant impact on therapeutic response and clinical outcomes (Jin et al., 2008; Nizet et al., 2022).

The AI tools have been developed to assist patients with adherence. These can come in a form of phone applications (apps) with a personalized patient profile that sends reminders with schedules for daily medications. These AI technologies have information that can assist patients with frequently asked questions related to the medication, how to control possible adverse effects, and when to visit their oncologist. Web-based apps allow patients to constantly report adverse effects to their healthcare providers assisting with patient monitoring. These apps have built-in alerts set to alarm healthcare providers should the patient need immediate attention. Graetz et al. evaluated the use of a remote patient care monitoring platform to improve management of adverse effects and adherence to breast cancer therapy. Alerts were sent by email to the hospital team who would then pay attention to reports of new symptoms, increased severity, or if 2 or more dosages were missed and respond to patients within 24 hours of receiving alerts. The apps with reminders were utilized better (73.5% of participants) than apps without. Email alerts received from this group of patients were also higher. Adherence to breast cancer therapy was significantly higher in patients who had the app and reminder (100%) than those who only had the app (72.7%) (Graetz et al., 2018). However, the study did not compare compliance with patients who had no app.

Chatter robots (Chatbots) are also referred to as digital assistants or intelligent agents. Chatbots is a software application developed from ML that can be used to have a conversation with a patient via text (Xu et al., 2021). Conveniently so,

chatbots are integrated into WhatsApp messaging that is more familiar and widely used in African countries as well (Belfin et al., 2019). Chatbots would allow a patient to take a photo of a skin lesion for cancer diagnosis. Data sets of images of patients confirmed with skin cancer are used to train CNN models to make predictions of whether the lesion is malignant or benign. If the lesion is malignant, the model will be able to accurately classify the type of skin cancer the patient has. The patient will be able to communicate with the software and ask questions related to the diagnosis and get urgent medical advice and intervention (An, 2021). This platform can further be used to monitor patient response to skin cancer immunotherapy. Adverse effects can be managed effectively and images of improved or worse responses can be uploaded to help the patient cope with the disease outside clinical visits.

These AI applications could assist with improved patient overall survival and reduce the burden of disease which would in return reduce hospital visits and consequently transport costs. This will also provide healthcare professionals with the opportunity to manage the burden of service delivery and reduce costs related to patient management, especially in the public sector. Remote monitoring can help patients keep track of their medications and reduce the risk of cancer recurrence (Xu et al., 2021).

12.3 Challenges of AI in Cancer Immunotherapy

Implementation of AI use, particularly in LMICs can pose significant challenges. Most African countries have regions where there is limited access to technology such as communication towers needed to transmit cell phone signals or connect to the internet. Thus, web-based apps used to monitor patient treatment response and adherence can be challenging in these setting. Although such interventions have been reported to improve patient adherence, reduce racial disparities in cancer treatment strategies, and improve cancer survival outcomes, these advantages are yet to be leveraged in LMICs (Graetz et al., 2018). It should be noted that most cancer patients require psychological and emotional support and these applications do not provide human interaction and emotional support. Consequences related to psychological disturbances should the diagnosis be positive and the risks associated with self-misdiagnosis can be detrimental. The apps are also not equipped with detailed treatment plans (Xu et al., 2021).

A multidisciplinary team (MDT) consisting of oncologists, radiologists, surgeons, and pathologists was employed to assess the effectiveness of an assistant decision-making system trained by Memorial Sloan Kettering [Watson for Oncology (WFO)]. This AI technology is used to recommend treatment regimens for cancer patients. Discrepancies in the concordance rate amongst participating countries and the types of cancer were observed. Out of the 1738 cases that were studied, 959 (55.18%) cases were consistent with the MDT treatment schemes. In some cases, 166 (9.55%) fell under the "not recommend" scheme while 110 cases (6.33%) were not available for recommendation on WFO. This data serves as an indicator

that there is still more work to be done to improve the reliability of AI tools used for treatment recommendations with either chemo or immunotherapeutic interventions. The use of technologies such as WFO in LMICs is not always practical as it does not accommodate locally available treatment modalities though it is meant to generate treatment recommendations based on the data captured (Jie et al., 2021).

More applications use imaging features in combination with non-imaging data and real-world data to train AI models for precision oncology. However, the quality of data sets especially in studies where data banks such as the TCGA database are not utilized can produce unreliable results. The concept of putting junk in and getting junk out applies in every aspect of medical research. Training of models using large data sets requires a multi-center approach. The challenge here is that annotation is labor intensive and requires patience. This might result in discrepancies in segmentations that might affect feature mapping. Thus, when collecting data from different study sites, these issues should be taken into consideration during the validation processes of the AI model. Ethical considerations might also be an issue as research done on images require waiver of patient informed consent. The issue around the protection of the personal information act (POPIA) could also be a concern as most of the imaging data comes from archives and getting permission from patients themselves can limit the feasibility of research studies (Cheung & Rubin, 2021).

12.4 Future Perspectives

The application of AI technologies has improved the efficiency and accuracy of diagnosis and prediction of treatment responses in cancer. This includes the identification of biomarkers that have allowed for non-invasive diagnostic and prediction methods. The precise identification of HLA alleles and presentation based on mRNA expression is one of the breakthroughs in the development of future cancer immunotherapies. Future efforts to improve therapeutic response to ICs and limit adverse effects should include analysis of other immune pathways involving suppressive cells such as Tregs. Amoozgar et al. found that targeting glucocorticoid-induced TNFR-related receptor (GITR) in Tregs can ameliorate anti-PD-1 therapeutic response by several mechanisms including converting Tregs into cytotoxic T cells and improving MHC recognition of tumor antigens. The authors noted significant alleviations of therapeutic resistance and eradication of tumors in murine models (Amoozgar et al., 2021). AI application in combination with RNA-Seq data sets could improve the efficacy and accuracy of these studies. Validation of data using AI tools could also facilitate their translation into clinical settings.

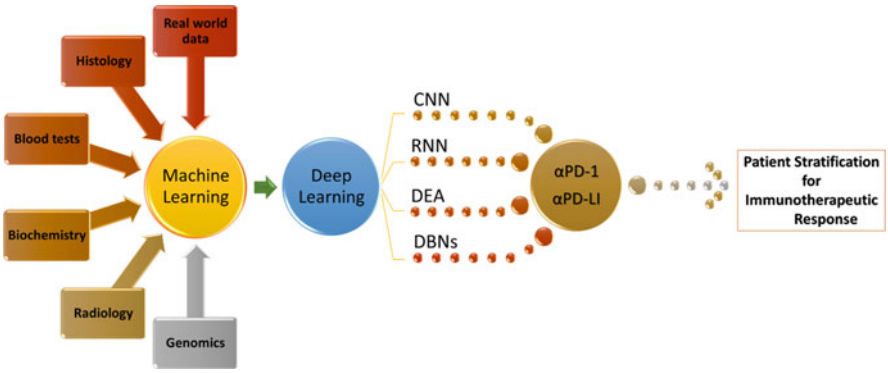


Fig. 12.3 Multiple DL algorithms can be tested in determining patients that will benefit from cancer immunotherapy. The most commonly used DL algorithms are listed on the diagram. AI accomplishes what the human mind cannot comprehend in a short space of time with precision and accuracy. The use of AI allows for the collection of data from multifaceted medical diagnostic platforms. Large data sets can be analyzed to group patients into responders and non-responders, saving time, and costs related to patient care/management and improving overall survival. DAE—Denoising autoencoder, DBNs—deep belief networks, LSTM—long short-term memory

12.5 Conclusion

The introduction of AI in medicine has allowed for big data sets to be processed by integrating imaging data and data collected from several medical disciplines to facilitate decision-making and improve precision oncology and clinical outcomes (Fig. 12.3). AI has allowed for the stratification of patients into responders and non-responders so that only patients who will benefit from cancer immunotherapy are treated accordingly. This also assists with limiting irAEs and promotes patient overall survival. Other treatment modalities can then be considered in a group of patients categorized as non-responders. As noted in the text, the use of AI has shown that even combinatorial therapies with chemotherapy and immunotherapy might still not be beneficial in patients grouped as non-responders.

Acknowledgments Parts of the figure were drawn by using pictures from Servier Medical Art licensed under a Creative Commons Attribution 3.0 Unsupported License.

References

Ahmed, H. G., Bensumaidea, S. H., Alshammari, F. D., Alenazi, F. S. H., Ba, A. L., Alturkstani, M. Z., & Aladani, I. A. (2017). Prevalence of human papillomavirus subtypes 16 and 18 among Yemeni patients with cervical cancer. *Asian Pacific Journal of Cancer Prevention, 18*, 1543–1548.

Alilou, M., Patil, P., Fu, P., Bera, K., Velcheti, V., Madabhushi, A., & Vaidya, P. (2019). P1.04-25 CT based vessel tortuosity features are prognostic of overall survival and predictive of immunotherapy response in NSCLC patients. *Journal of Thoracic Oncology, 14*, S449.

Alkinani, M. H., & El-Sakka, M. R. (2017). Patch-based models and algorithms for image denoising: A comparative review between patch-based images denoising methods for additive noise reduction. *EURASIP Journal on Image and Video Processing, 2017*, 58.

Amoozgar, Z., Kloepper, J., Ren, J., Tay, R. E., Kazer, S. W., Kiner, E., Krishnan, S., Posada, J. M., Ghosh, M., Mamessier, E., Wong, C., Ferraro, G. B., Batista, A., Wang, N., Badeaux, M., Roberge, S., Xu, L., Huang, P., Shalek, A. K., Fukumura, D., Kim, H.-J., & Jain, R. K. (2021). Targeting Treg cells with GITR activation alleviates resistance to immunotherapy in murine glioblastomas. *Nature Communications, 12*, 2582.

An, Q. (2021). Design of a skin cancer diagnosing web application based on convolutional neural network model and chatterbot application programming interface. *Journal of Physics: Conference Series, 2078*, 012039.

Anassi, E., & Ndefo, U. A. (2011). Sipuleucel-T (provenge) injection: the first immunotherapy agent (vaccine) for hormone-refractory prostate cancer. *P & T: A Peer-Reviewed Journal for Formulary Management, 36*, 197–202.

Barber, F. D. (2019). Adverse events of oncologic immunotherapy and their management. *Asia-Pacific Journal of Oncology Nursing, 6*, 212–226.

Beatty, G. L., & Gladney, W. L. (2015). Immune escape mechanisms as a guide for cancer immunotherapy. *Clinical Cancer Research, 21*, 687–692.

Belfin, R. V., Shobana, A. J., Manilal, M., Mathew, A. A., & Babu, B. (2019). A graph based chatbot for cancer patients. *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, 15–16 March 2019. pp. 717–721.

Bracci, L., Schiavoni, G., Sistigu, A., & Belardelli, F. (2014). Immune-based mechanisms of cytotoxic chemotherapy: Implications for the design of novel and rationale-based combined treatments against cancer. *Cell Death and Differentiation, 21*, 15–25.

Bulik-Sullivan, B., Busby, J., Palmer, C. D., Davis, M. J., Murphy, T., Clark, A., Busby, M., Duke, F., Yang, A., Young, L., Ojo, N. C., Caldwell, K., Abhyankar, J., Boucher, T., Hart, M. G., Makarov, V., De Montpreville, V. T., Mercier, O., Chan, T. A., Scagliotti, G., Bironzo, P., Novello, S., Karachaliou, N., Rosell, R., Anderson, I., Gabrail, N., Hrom, J., Limvarapuss, C., Choquette, K., Spira, A., Rousseau, R., Voong, C., Rizvi, N. A., Fadel, E., Frattini, M., Jooss, K., Skoberne, M., Francis, J., & Yelensky, R. (2019). Deep learning using tumor HLA peptide mass spectrometry datasets improves neoantigen identification. *Nature Biotechnology, 37*, 55–63.

Burkholder, B., Huang, R. Y., Burgess, R., Luo, S., Jones, V. S., Zhang, W., Lv, Z. Q., Gao, C. Y., Wang, B. L., Zhang, Y. M., & Huang, R. P. (2014). Tumor-induced perturbations of cytokines and immune cell networks. *Biochimica et Biophysica Acta, 1845*, 182–201.

Chaudhury, S., Duncan, E. H., Atre, T., Storme, C. K., Beck, K., Kaba, S. A., Lanar, D. E., & Bergmann-Leitner, E. S. (2018). Identification of immune signatures of novel adjuvant formulations using machine learning. *Scientific Reports, 8*, 17508.

Chen, X., Zhang, Y., & Fu, Y. (2022). The critical role of Toll-like receptor-mediated signaling in cancer immunotherapy. *Medicine in Drug Discovery, 14*, 100122.

Cheung, H. M. C., & Rubin, D. (2021). Challenges and opportunities for artificial intelligence in oncological imaging. *Clinical Radiology, 76*, 728–736.

Choi, Y., Shi, Y., Haymaker, C. L., Naing, A., Ciliberto, G., & Hajjar, J. (2020). T-cell agonists in cancer immunotherapy. *Journal for Immunotherapy of Cancer, 8*, e000966.

Coulson, A., Levy, A., & Gossell-Williams, M. (2014). Monoclonal antibodies in cancer therapy: Mechanisms, successes and limitations. *The West Indian Medical Journal, 63*, 650–654.

Davis, A. A., & Patel, V. G. (2019). The role of PD-L1 expression as a predictive biomarker: An analysis of all US Food and Drug Administration (FDA) approvals of immune checkpoint inhibitors. *Journal for Immunotherapy of Cancer, 7*, 278.

De Hond, A. A. H., Leeuwenberg, A. M., Hooft, L., Kant, I. M. J., Nijman, S. W. J., Van Os, H. J. A., Aardoom, J. J., Debray, T. P. A., Schuit, E., Van Smeden, M., Reitsma, J. B., Steyerberg, E. W., Chavannes, N. H., & Moons, K. G. M. (2022). Guidelines and quality

criteria for artificial intelligence-based prediction models in healthcare: A scoping review. *npj Digital Medicine, 5*, 2.

Demaria, P. J., & Bilusic, M. (2019). Cancer vaccines. *Hematology/Oncology Clinics of North America, 33*, 199–214.

Dubensky, T. W., Jr., & Reed, S. G. (2010). Adjuvants for cancer vaccines. *Seminars in Immunology, 22*, 155–161.

Ferrucci, P. F., Pala, L., Conforti, F., & Cocorocchio, E. (2021). Talimogene Laherparepvec (T-VEC): An intralesional cancer immunotherapy for advanced melanoma. *Cancers, 13*, 1383.

Fischer, J. W., & Bhattarai, N. (2021). CAR-T cell therapy: Mechanism, management, and mitigation of inflammatory toxicities. *Frontiers in Immunology, 12*.

Graetz, I., Mckillop, C. N., Stepanski, E., Vidal, G. A., Anderson, J. N., & Schwartzberg, L. S. (2018). Use of a web-based app to improve breast cancer symptom management and adherence for aromatase inhibitors: A randomized controlled feasibility trial. *Journal of Cancer Survivorship, 12*, 431–440.

Han, J., Gu, X., Li, Y., & Wu, Q. (2020). Mechanisms of BCG in the treatment of bladder cancer-current understanding and the prospect. *Biomedicine & Pharmacotherapy, 129*, 110393.

Han, S., Toker, A., Liu, Z. Q., & Ohashi, P. S. (2019). Turning the tide against regulatory T cells. *Frontiers in Oncology, 9*, 279.

Hayashi, F., Means, T. K., & Luster, A. D. (2003). Toll-like receptors stimulate human neutrophil function. *Blood, 102*, 2660–2669.

Iivanainen, S., Ekstrom, J., Virtanen, H., Kataja, V. V., & Koivunen, J. P. (2021). Electronic patient-reported outcomes and machine learning in predicting immune-related adverse events of immune checkpoint inhibitor therapies. *BMC Medical Informatics and Decision Making, 21*, 205.

Jie, Z., Zhiying, Z., & Li, L. (2021). A meta-analysis of Watson for Oncology in clinical application. *Scientific Reports, 11*, 5792.

Jin, J., Sklar, G. E., Min Sen Oh, V., & Chuen Li, S. (2008). Factors affecting therapeutic compliance: A review from the patient's perspective. *Therapeutics and Clinical Risk Management, 4*, 269–286.

Kaczanowska, S., Joseph, A. M., & Davila, E. (2013). TLR agonists: Our best frenemy in cancer immunotherapy. *Journal of Leukocyte Biology, 93*, 847–863.

Kichloo, A., Albosta, M., Dahiya, D., Guidi, J. C., Aljadah, M., Singh, J., Shaka, H., Wani, F., Kumar, A., & Lekkala, M. (2021). Systemic adverse effects and toxicities associated with immunotherapy: A review. *World Jof ournalClinical Oncology, 12*, 150–163.

Koual, M., Tomkiewicz, C., Cano-Sancho, G., Antignac, J.-P., Bats, A.-S., & Coumoul, X. (2020). Environmental chemicals, breast cancer progression and drug resistance. *Environmental Health, 19*, 117.

Lee, S., & Margolin, K. (2011). Cytokines in cancer immunotherapy. *Cancers, 3*, 3856–3893.

Lewis, P., & Yap, T. A. (2020). Optimizing precision oncology and immunotherapy strategies: Moving into the next stage of cancer medicine. *Journal of Immunotherapy and Precision Oncology, 2*, 1–3.

Liang, G., Fan, W., Luo, H., & Zhu, X. (2020). The emerging roles of artificial intelligence in cancer drug development and precision therapy. *Biomedicine & Pharmacotherapy, 128*, 110255.

Liang, Y.-L., Zhang, Y., Tan, X.-R., Qiao, H., Liu, S.-R., Tang, L.-L., Mao, Y.-P., Chen, L., Li, W.-F., Zhou, G.-Q., Zhao, Y., Li, J.-Y., Li, Q., Huang, S.-Y., Gong, S., Zheng, Z.-Q., Li, Z.-X., Sun, Y., Jiang, W., Ma, J., Li, Y.-Q., & Liu, N. (2022). A lncRNA signature associated with tumor immune heterogeneity predicts distant metastasis in locoregionally advanced nasopharyngeal carcinoma. *Nature Communications, 13*, 2996.

Lin, Y., Pan, X., Chen, Z., Lin, S., & Chen, S. (2020). Identification of an immune-related nine-lncRNA signature predictive of overall survival in colon cancer. *Frontiers in Genetics, 11*, 318.

Liu, Z., Liu, L., Weng, S., Guo, C., Dang, Q., Xu, H., Wang, L., Lu, T., Zhang, Y., Sun, Z., & Han, X. (2022). Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. *Nature Communications, 13*, 816.

Liu, Z., Xie, Y., Xiong, Y., Liu, S., Qiu, C., Zhu, Z., Mao, H., Yu, M., & Wang, X. (2020). TLR 7/8 agonist reverses oxaliplatin resistance in colorectal cancer via directing the myeloid-derived suppressor cells to tumoricidal M1-macrophages. *Cancer Letters, 469*, 173–185.

Ma, W., Zhao, F., Yu, X., Guan, S., Suo, H., Tao, Z., Qiu, Y., Wu, Y., Cao, Y., & Jin, F. (2020). Immune-related lncRNAs as predictors of survival in breast cancer: A prognostic signature. *Journal of Translational Medicine, 18*, 442.

Martins, F., Sofiya, L., Sykiotis, G. P., Lamine, F., Maillard, M., Fraga, M., Shabafrouz, K., Ribi, C., Cairoli, A., Guex-Crosier, Y., Kuntzer, T., Michielin, O., Peters, S., Coukos, G., Spertini, F., Thompson, J. A., & Obeid, M. (2019). Adverse effects of immune-checkpoint inhibitors: epidemiology, management and surveillance. *Nature Reviews Clinical Oncology, 16*, 563–580.

Nizet, P., Touchefeu, Y., Pecout, S., Cauchin, E., Beaudouin, E., Mayol, S., Fronteau, C., & Huon, J.-F. (2022). Exploring the factors influencing adherence to oral anticancer drugs in patients with digestive cancer: A qualitative study. *Supportive Care in Cancer, 30*, 2591–2604.

Place, A. E., Jin Huh, S., & Polyak, K. (2011). The microenvironment in breast cancer progression: Biology and implications for treatment. *Breast Cancer Research, 13*, 227.

Prelaj, A., Boeri, M., Robuschi, A., Ferrara, R., Proto, C., Lo Russo, G., Galli, G., De Toma, A., Brambilla, M., Occhipinti, M., Manglaviti, S., Beninato, T., Bottiglieri, A., Massa, G., Zattarin, E., Gallucci, R., Galli, E. G., Ganzinelli, M., Sozzi, G., De Braud, F. G. M., Garassino, M. C., Restelli, M., Pedrocchi, A. L. G., & Trovo, F. (2022). Machine learning using real-world and translational data to improve treatment selection for NSCLC patients treated with immunotherapy. *Cancers (Basel), 14*.

Quail, D. F., & Joyce, J. A. (2013). Microenvironmental regulation of tumor progression and metastasis. *Nature Medicine, 19*, 1423–1437.

Rajakumar, T., Horos, R., Jehn, J., Schenz, J., Muley, T., Pelea, O., Hofmann, S., Kittner, P., Kahraman, M., Heuvelman, M., Sikosek, T., Feufel, J., Skottke, J., Nötzel, D., Hinkfoth, F., Tikk, K., Daniel-Moreno, A., Ceiler, J., Mercaldo, N., Uhle, F., Uhle, S., Weigand, M. A., Elshiaty, M., Lusky, F., Schindler, H., Ferry, Q., Sauka-Spengler, T., Wu, Q., Rabe, K. F., Reck, M., Thomas, M., Christopoulos, P., & Steinkraus, B. R. (2022). A blood-based miRNA signature with prognostic value for overall survival in advanced stage non-small cell lung cancer treated with immunotherapy. *NPJ Precision Oncology, 6*, 19.

Rajakumar, T., Horos, R., Kittner, P., Kahraman, M., Sikosek, T., Hinkfoth, F., Tikk, K., Mercaldo, N. D., Stenzinger, A., Rabe, K. F., Reck, M., Thomas, M., Christopoulos, P., & Steinkraus, B. R. (2022). Brief report: A blood-based microRNA complementary diagnostic predicts immunotherapy efficacy in advanced-stage NSCLC with high programmed death-ligand 1 expression. *JTO Clinical and Research Reports, 3*, 100369.

Ren, W., Zuo, S., Yang, L., Tu, R., Wang, P., & Zhang, X. (2022). Identification of a novel immune-related long noncoding RNA signature to predict the prognosis of bladder cancer. *Scientific Reports, 12*, 3444.

Rohaan, M. W., Wilgenhof, S., & Haanen, J. B. A. G. (2019). Adoptive cellular therapies: The current landscape. *Virchows Archiv: An International Journal of Pathology, 474*, 449–461.

Sambi, M., Bagheri, L., & Szewczuk, M. R. (2019). Current challenges in cancer immunotherapy: Multimodal approaches to improve efficacy and patient response rates. *Journal of Oncology, 2019*, 4508794–4508794.

Sarker, I. H. (2021). Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science, 2*, 420.

Shaikh, F. J., & Rao, D. S. (2022). Prediction of cancer disease using machine learning approach. *Materials Today: Proceedings, 50*, 40–47.

Shao, D., Dai, Y., Li, N., Cao, X., Zhao, W., Cheng, L., Rong, Z., Huang, L., Wang, Y., & Zhao, J. (2021). Artificial intelligence in clinical research of cancers. *Briefings in Bioinformatics, 23*.

Sun, R., Limkin, E. J., Vakalopoulou, M., Dercle, L., Champiat, S., Han, S. R., Verlingue, L., Brandao, D., Lancia, A., Ammari, S., Hollebecque, A., Scoazec, J. Y., Marabelle, A., Massard, C., Soria, J. C., Robert, C., Paragios, N., Deutsch, E., & Ferté, C. (2018). 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. *The Lancet Oncology, 19*, 1180–1191.

Suresh, S., & O'Donnell, K. A. (2021). Translational control of immune evasion in cancer. *Trends Cancer, 7*, 580–582.

Trebeschi, S., Bodalal, Z., Van Dijk, N., Boellaard, T. N., Apfaltrer, P., Tareco Bucho, T. M., Nguyen-Kim, T. D. L., van der Heijden, M. S., Aerts, H. J. W. L., & Beets-Tan, R. G. H. (2021). Development of a prognostic AI-monitor for metastatic urothelial cancer patients receiving immunotherapy. *Frontiers in Oncology, 11*.

Twomey, J. D., & Zhang, B. (2021). Cancer immunotherapy update: FDA-approved checkpoint inhibitors and companion diagnostics. *The AAPS Journal, 23*, 39.

Ungefroren, H., Sebens, S., Seidl, D., Lehnert, H., & Hass, R. (2011). Interaction of tumor cells with the microenvironment. *Cell Communication and Signaling: CCS, 9*, 18.

Vaddepally, R. K., Kharel, P., Pandey, R., Garje, R., & Chandra, A. B. (2020). Review of indications of FDA-approved immune checkpoint inhibitors per NCCN guidelines with the level of evidence. *Cancers, 12*, 738.

Vaidya, P., Bera, K., Patil, P. D., Gupta, A., Jain, P., Alilou, M., Khorrami, M., Velcheti, V., & Madabhushi, A. (2020). Novel, non-invasive imaging approach to identify patients with advanced non-small cell lung cancer at risk of hyperprogressive disease with immune checkpoint blockade. *Journal for Immunotherapy of Cancer, 8*, e001343.

Ventola, C. L. (2017). Cancer immunotherapy, Part 2: Efficacy, safety, and other clinical considerations. *P & T: A Peer-Reviewed Journal for Formulary Management, 42*, 452–463.

Waldman, A. D., Fritz, J. M., & Lenardo, M. J. (2020). A guide to cancer immunotherapy: From T cell basic science to clinical practice. *Nature Reviews Immunology, 20*, 651–668.

Wang, Q. (2016). Cancer predisposition genes: Molecular mechanisms and clinical impact on personalized cancer care: examples of Lynch and HBOC syndromes. *Acta Pharmacologica Sinica, 37*, 143–149.

Wang, X., Barrera, C., Bera, K., Viswanathan, V. S., Azarianpour-Esfahani, S., Koyuncu, C., Velu, P., Feldman, M. D., Yang, M., Fu, P., Schalper, K. A., Mahdi, H., Lu, C., Velcheti, V., & Madabhushi, A. (2022). Spatial interplay patterns of cancer nuclei and tumor-infiltrating lymphocytes (TILs) predict clinical benefit for immune checkpoint inhibitors. *Science Advances, 8*, eabn3966.

Xiong, S., Dong, L., & Cheng, L. (2021). Neutrophils in cancer carcinogenesis and metastasis. *Journal of Hematology & Oncology, 14*, 173.

Xu, L., Sanders, L., Li, K., & Chow, J. C. L. (2021). Chatbot for health care and oncology applications using artificial intelligence and machine learning: Systematic review. *JMIR Cancer, 7*, e27850.

Yang, Y., Yang, J., Shen, L., Chen, J., Xia, L., Ni, B., Ge, L., Wang, Y., & Lu, S. (2021). A multi-omics-based serial deep learning approach to predict clinical outcomes of single-agent anti-PD-1/PD-L1 immunotherapy in advanced stage non-small-cell lung cancer. *American Journal of Translational Research, 13*, 743–756.

Zahavi, D., & Weiner, L. (2020). Monoclonal antibodies in cancer therapy. *Antibodies (Basel, Switzerland), 9*, 34.

Zeng, Q., Klein, C., Caruso, S., Maille, P., Laleh, N. G., Sommacale, D., Laurent, A., Amaddeo, G., Gentien, D., Rapinat, A., Regnault, H., Charpy, C., Nguyen, C. T., Tournigand, C., Brustia, R., Pawlotsky, J. M., Kather, J. N., Maiuri, M. C., Loménie, N., & Calderaro, J. (2022). Artificial intelligence predicts immune and inflammatory gene signatures directly from hepatocellular carcinoma histology. *Journal of Hepatology, 77*, 116–127.

Zhou, D., Wang, J., & Liu, X. (2022). Development of six immune-related lncRNA signature prognostic model for smoking-positive lung adenocarcinoma. *Journal of Clinical Laboratory Analysis, 36*, e24467.

Chapter 13
Employing AI-Powered Decision Support Systems in Recommending the Most Effective Therapeutic Approaches for Individual Cancer Patients: Maximising Therapeutic Efficacy

Thulo Molefi, Rahaba Marima, Demetra Demetriou, Afra Basera, and Zodwa Dlamini

Abstract Artificial Intelligence (AI) has made unique advances in anti-cancer drug discovery, development and treatment. Although AI efforts in cancer therapy are not meant to replace human capabilities, its benefits cannot be ignored either, as humans' capacity to generate most appropriate personalised treatment may be limited. Every patient responds uniquely to treatment and AI-tools offer opportunities to study unique mechanisms of response and even recommend the most effective personalised treatment combination. The success of precision oncology not only depends on patient-drug optimum combination but also relies on innovative therapeutic approaches that may include new drug discovery, development and drug repurposing. AI holds great potential in maximising therapeutic efficacy for each patient and improving patient outcome. Furthermore, it is emerging that adaptive therapy empowered by AI may transform the landscape of precision oncology in favour of improved patient outcome. This chapter will focus on AI-enhanced decision support systems in optimising therapeutic approaches, advancing precision oncology and improving patient outcome.

T. Molefi (✉)
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Department of Medical Oncology, Steve Biko Academic Hospital and University of Pretoria, Pretoria, South Africa
e-mail: thulo.molefi@up.ac.za

R. Marima · D. Demetriou · A. Basera · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: rahaba.marima@up.ac.za; dd.demetriou@up.ac.za; u22800370@tuks.co.za; zodwa.dlamini@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_13

Keywords Artificial Intelligence (AI) · Precision oncology · Chemotherapy · Drug development · Drug combination · Therapeutic efficacy

13.1 Introduction

Artificial intelligence (AI) is a branch of computer science that deals with human behaviour simulation in computers. AI relies on computers to follow human established algorithms. These algorithms may also be learned by computers in supporting human decisions or executing specific human tasks (Hosny et al., 2018). Machine learning (ML), a sub-field of AI, represents processes by which computers can improve their own performance by continuous incorporation of new data into an existing model (Rajkomar et al., 2019). Deep learning (DL) is an ML sub-field in which mathematical algorithms are arrayed using multi-layered computational units that resemble human cognition (Farina et al., 2022). DL can find hidden information in images that may not be recognisable by the human eye, cancer research and management is one of the most AI benefiting fields (Tartar et al., 2014; Van der Waal, 2018; Li et al., 2019). DL has strong learning and reasoning abilities, thus simulating human capacity (Houssami et al., 2019; Sherbet et al., 2018).

The growing load of cancer on healthcare systems and the necessity to minimise its adverse effect on cancer patients' lives necessitate the generation of care approaches supported by predictive, personalised, preventive and participatory (P4) systems (Hood & Auffray, 2013). Thus, employing AI-powered decision support systems in recommending the most effective therapeutic approaches for individual cancer patients in maximising therapeutic efficacy may be beneficial. In this era, the widely employed treatment methods are steered by clinical practice guidelines crafted by oncology experts. For instance, the American Society of Clinical Oncology (ASCO) supplies oncology sites with guidelines that report their practices via its Quality Oncology Practice Initiative. ASCO reports back with a clinical site evaluation hinged on high-value measures that are patient-centred procedures, including pain regulation (Abernethy et al., 2010; Rocque et al., 2019). Notably, as large amounts of cancer patients' clinical and molecular data become available, AI and ML are being investigated for their potential to aid in the multi-faceted cancer treatment decision-making (El Naqa et al., 2018; Walsh et al., 2019). Diagnostics, prognostics, prediction of therapy outcomes and therapy prescriptions are all possible applications of AI in oncology. For example, deep neural networks (DNNs) and convolutional neural networks (CNNs) are utilised to categorise skin cancer lesions (Esteva et al., 2017), lung cancer patterns (Gertych et al., 2019; Wei et al., 2019), forecast HLA-peptide ability to bind for immunotherapy (Martins et al., 2019) and define target quantity for radiotherapy (Boon et al., 2018). The ML applications also include breast cancer therapy advice to deter metastasis (Jiang et al., 2019) and using Bayesian networks to aid in the therapy plan (Sesen et al., 2013; Cypko et al., 2017). Bayesian networks and logistic regression have also been used to predict cancer recurrence (Witteveen et al., 2018). Other uses of ML in cancer studies include predicting the short-term risk of death in patients beginning

chemotherapy (Elfiky et al., 2018), indicating which patients will benefit from adjuvant therapy (Steele et al., 2014). Additionally, AI could be essential to precision oncology by the detection and discovery of novel biomarkers that can be targeted for population and personalised based therapy (Derbal, 2022; West et al., 2020). This chapter will discuss AI-enhanced decision support systems, improving therapeutic efficacies for personalised cancer care and improving patient outcome.

13.2 AI-Tools in Optimising Drug Combinations and Enhancing Effective Cancer Therapeutics: From Drug Development to Personalised Care

Artificial intelligence (AI) tools have the potential to transform various facets of cancer care, management and therapy. AI-empowered cancer therapy approaches may include anti-cancer drug discovery, development and repurposing. These AI-augmented efforts may also facilitate how these drugs are clinically validated and administered (Ho, 2020). To date, drug discovery and development processes are time consuming and costly. Variability between patients' treatment outcomes is also a challenge. Thus, AI holds promising potential to bridge such existing gaps in optimising most effective therapeutic approaches, particularly for individual cancer patients, Figs. 13.1 and 13.2.

AI-tools have also been reported to play a significant role in drug repurposing by effectively predicting drug behaviour using genomic and chemical data (Menden et al., 2013). Furthermore, AI-reinforcement learning was successfully reported to design a new drug compound in 3 weeks, compared to standard timelines of about 1 year. Reinforcement learning uses the punishment and reward approach to train AI-algorithms with a goal to achieve an intended drug structure. In addition to AI-augmented anti-cancer drug discovery, drug pharmacokinetic properties can also be deduced. For example, Zhavoronkov et al.'s (2019) study demonstrated that the generative tensorial reinforcement learning (GENTRL) was trained by chemical structures dataset that target the tyrosine kinase discoidin domain receptor 1 (DDR1) (Zhavoronkov et al., 2019). This receptor has been reported to promote progression of various cancer cells. AI-based tools predicted receptor binding and lead compound, with an aim of minimising off-targets or other tyrosine kinase isoforms, thus improving drug-target binding. Table 13.1 summarises AI-based tools in drug discovery.

AI-technology has the potential to transform the traditional drug-combination based approach using the already available clinical trials data and resolving extensive drug and dose parameters. This novel AI-approach will distinctly broaden the pool of drugs in consideration and optimise combinations, even rare combinations that may be effective towards maximising therapeutic efficacy. For instance, the quadratic phenotypic optimisation platform (QPOP) has the potential to be used as a clinical decision support platform to identify appropriate patient-specific drug

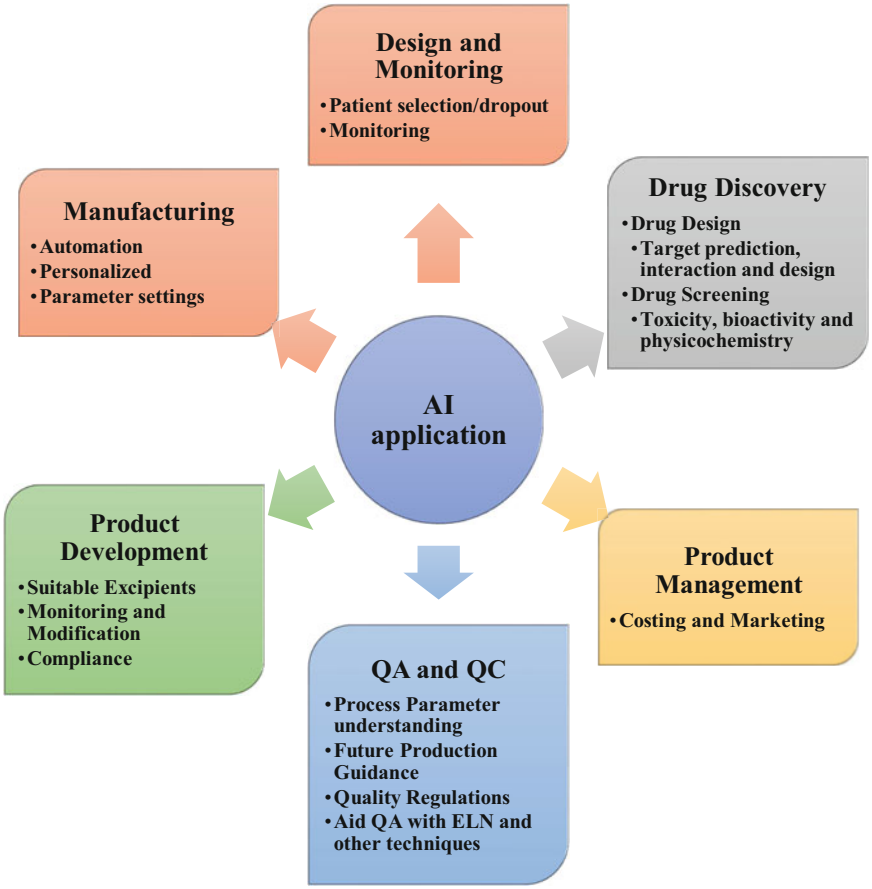


Fig. 13.1 AI-application in drug development. AI allows the optimisation of clinical trial drug design and minoring by improving target prediction and interaction, thus reducing toxicity, optimising bioactivity and physicochemical properties while upholding quality assurance and control

Fig. 13.2 Cancer therapy developmental stages. The first step in therapy is to identify the drug target. Lead compound must be generated and optimised to ensure adequate binding to the target. The process will enter pre-clinical development and tested before entering the clinical trial phases. This process will be optimised towards personalised cancer therapy

combinations. QPOP does not depend on previous assumptions such as cancer molecular mechanisms, associated drug targets and drug combinations. This AI-platform uses parabola represented quadratic relationships, thus correlating

Table 13.1 AI-tools used in drug discovery

Tool	Information	Website URL and Reference
DeepTox	This software predicts the toxicity of drugs.	www.bioinf.jku.at/research/DeepTox (Ciallella & Zhu, 2019)
ORGANIC	This is a molecular generation tool that assists in the creations of molecules with the required properties.	https://github.com/aspuru-guzik-group/ORGANIC (Brown, 2015)
DeepChem	This is an MLP model that uses a python-based AI system. This tool can be used to find an adequate drug discovery candidate.	https://github.com/deepchem/deepchem (Zhu, 2020)
DeltaVina	This is a scoring function that allows rescoring of drug–ligand binding affinity.	https://github.com/chengwang88/deltavina
DeepNeural-NetQSAR	This is a Python-based system driven by computational tools. It can assist in the detection of compound molecular activity.	https://github.com/Merck/DeepNeuralNet-QSAR (Chan et al., 2019)
Chemputer	Chemputer assists in the reporting procedure of chemical synthesis and standardisation.	https://zenodo.org/record/1481731
PotentialNet	Predicts binding affinity of ligands using NNs.	https://pubs.acs.org/doi/full/10.1021/acscentsci.8b00507 (Pereira et al., 2016)
AlphaFold	3D protein structures prediction.	https://deepmind.com/blog/alphafold
Hit Dexter	Uses ML techniques to predict molecules that respond to biochemical assays.	http://hitdexter2.zbh.uni-hamburg.de
Neural graph fingerprint	Predicts properties of novel molecules.	https://github.com/HIPS/neural-fingerprint

inputs with optimal output sets. In such cases, inputs would be drugs and doses, while outputs will be pre-clinical reduction in tumour with minimised drug toxicity. Various studies have evaluated the QPOP platform in improving therapeutic efficacy. For example, Rashid et al. (2018) used the QPOP AI-platform to evaluate 14 chemotherapy drugs in synergy with other drugs to treat multiple myeloma mouse models (Rashid et al., 2018). Unexpectedly, QPOP identified rare drug combinations such as mitomycin C and decitabine. Reportedly, monotherapy drug efficacy of each of these two drugs was not achieved. This combination noticeably improved MM mouse model outcomes when compared to standard clinical drug combinations (Rashid et al., 2018). This platform can be used to explore novel targeted therapies combinations in chemotherapy and immunotherapy.

AI has also been reported to predict toxicity related to radiation and chemotherapy (Isaksson et al., 2020; Oyaga-Iriarte et al., 2019; Cuplov & André, 2020). ML algorithms can be trained to develop models that can predict response to new anticancer drugs or drug combinations using high-throughput screening data (Liang et al., 2020; Simon et al., 2020; Meng et al., 2020; Goecks et al., 2020). Scientists are also reported to advance ML-drug discovery in generating and creating molecules' reverse synthesis pathways. Reverse synthesis of molecules, also known as

retrosynthesis, is a chemical synthesis method involving the deconstruction of a target molecule into its readily available, simple starting materials to assess the best synthetic route (Liang et al., 2020). Large amounts of data are usually created from the generation of new drugs, thus offering new opportunities to chemical-molecular data processing and results generation that will aid in drug development (Nascimento et al., 2019; Sharma & Rani, 2020; Watson et al., 2019). Additionally, ML can assist in data processing that has been collected over decades in a short space of time (Vamathevan et al., 2019). ML also can facilitate the support of making informed decisions that would otherwise require experimentation costs (Koromina et al., 2019; Klambauer et al., 2019; Ballester, 2019).

The approval rate of oncology drug candidates has unfortunately been reported to be below 5% (Wong et al., 2019). It has also been reported that identifying actionable biomarkers as potential treatment response predictive indicators improves patient outcomes (Lin et al., 2019). AI-empowered biomarker discovery is also a promising area towards improved treatment outcomes in advancing precision oncology (Lin et al., 2019; Harrer et al., 2019). Through harnessing of electronic health records (EHRs), patients' genomic data and biomarker information, AI may also enable remote clinicians' teams to collectively work together in providing inputs to diagnosis and prognosis. Furthermore, a neural network AI-platform, CURATE.AI was used to study patient-specific combination therapy by modulating multidrug dosing. This platform used a patient's data solely by employing a second-order algebraic algorithm. Optimal dosing associated with combination therapy safety and tumour reduction was observed at various treatment time points (Pantuck et al., 2018; Zarrinpar et al., 2016). However, this AI-platform to maximise therapeutic efficacy through innovative combination therapy is being tested on smaller patient cohorts and will therefore require larger patient pools for validation. Additionally, AI-optimised compounds that combine with other standard therapies at a sub-optimal level are not likely to significantly improve patient outcome.

Lind et al.'s (2019) research group developed an AI-random forest model that integrates screening data and ML. This AI-model can predict anti-cancer drug activity based on the mutation state of genomes of cancer cells. Furthermore, Wang et al. (2018) group developed an AI-model, called elastic regression predicting drug sensitivity. It has also been reported that ML models were able to predict various cancers' drug sensitivity. These included ovarian (tamoxifen treated), gastric (treated with 5-FU) and endometrial (paclitaxel treated) cancers (Hossain et al., 2019; Paik et al., 2019; McDonald, 2018; Li et al., 2019; Taninaga et al., 2019; Liu et al., 2019; Stanzione et al., 2021; Günakan et al., 2019). AI also may play an important role in addressing cancer drug resistance by learning and analysing large cancer drug-resistant data (Beck et al., 2020; Goldenberg et al., 2019; Leventakos et al., 2019). The implementation of AI in anti-cancer therapeutics will require involving the AI-tools from drug discovery, development, testing, validation and administration. AI-tools also hold a promise to increasing the pool of combination therapies for personalised care by tailoring bespoke treatments that combine multiple therapeutic strategies. Table 13.2 outlines the benefits and challenges of AI-enhanced cancer therapy.

Table 13.2 Benefits and challenges of the aspects of AI-enhanced cancer therapy

Aspect	Benefit	Challenge
Target Discovery	Decrease of toxicity and non-target effects. Increased drug exposure.	Identification of the optimal targets. Validation of AI-designed drugs.
Development	Drug and dosage optimisation. Patient and trial matching.	Improved trial results. Stratification of patient data.
Administration	Approved dosages. Decrease in resistance.	Validation. Similar benefits in various cancers.

13.3 AI-Empowered Clinical Decision Support: Applications in Chemotherapy, Radiotherapy, Immunotherapy

AI-chemotherapy applications focus more on patient-drug response. The key achievements include chemotherapy drug management use, chemotherapy drug tolerance prediction and chemotherapy program optimisation (Chen et al., 2018; Levine et al., 2019; Smaïl-Tabbone & Rance, 2019; Zhu et al., 2012). In a recent study by Pantuck et al. (2018), optimal dose combination of zen-3694 and enzalutamide was determined using CURATE.AI (Pantuck et al., 2018). Additionally, breast cancer patients with homologous repair (HR) defects were detected by a DL system which predicted 74% HR defects accuracy. This prediction enabled breast cancer patients to benefit from PARP inhibitors (Gulhan et al., 2019). In addition, another study demonstrated the link between taxol and gemcitabine chemotherapy drugs and breast cancer patients' genes using an ML algorithm. This AI-platform was able to predict patients' drug tolerance and differentiate between the two chemotherapy drugs' effects (Dorman et al., 2016). Furthermore, DL platform was used by a different study in risk stratification and induction chemotherapy guidance for nasopharyngeal carcinoma (NSC) (Tang et al., 2019). This AI-method was illustrated to be significantly better when compared to Epstein-Barr Virus (EBV)-DNA-based model (Peng et al., 2019).

In radiotherapy, the application of AI is more precise. It is reported that AI can automatically plan radiotherapy treatment and aiding radiologists identify key target areas (Fiorino et al., 2020; Lou et al., 2019; Meyer et al., 2018). Lin et al. (2019), for example, used a 3D CNN (3D CNN) in NSC delineation and achieved 79%, which is reported to be equivalent to that of a radiologist. Furthermore, radiotherapy treatment response in bladder cancer was also evaluated using a DL approach. This study combined DL with radiomics to construct a predictive model in determining treatment response. Additionally, Babier et al. (2018) developed an AI software that demonstrated reduction in radiation therapy planning time period, indicating similar radiation therapy treatment plan but reduced time in planning compared to a radiologist. AI-intervention in cancer overtreatment has also been demonstrated. This was shown by ML algorithm that could analyse digital images from women's

cervix. This AI-platform was illustrated to accurately recognise precancerous lesions to be treated over non-cancerous lesions, thus reducing patients' overtreatment (Hu et al., 2019). Furthermore, Bahl et al. (2018) developed an ML platform that could evaluate high-risk breast cancer lesions, aiding physicians in recommending appropriate treatment and avoid unnecessary surgery.

Compared to chemotherapy and radiotherapy, AI-immunotherapy applications primarily focus on treatment effect evaluation and aiding physicians adjust the treatment plan (Jabbari & Rezaei, 2019; Trebeschi et al., 2019; Abbasi, 2019; Tan et al., 2020). Sun et al. (2018) developed a DL platform that could precisely predict programmed cell death protein 1 (PD-1) inhibitors' therapeutic effects by evaluating immunotherapy effects in patients who are PD-1 sensitive with advanced solid tumours (Sun et al., 2018). Another research study developed an ML method which could improve identifying cancer neoantigen and cancer immunotherapy efficiency. This AI-method was based on mass spectrometry database of human leukocyte antigen (HLA). DL technologies have been reported to augment or support physician's treatment decision systems, and not to replace them. AI can help combine most appropriate treatment plan through learning from cancer patients' clinical big data (Meyer et al., 2018; Liu et al., 2018; Bogani et al., 2018; Golden, 2017; Walsh et al., 2019; Blackledge et al., 2019). For example, Prinzt (2017) developed an AI-platform, known as a Clinical Decision Support System (CDSS). This is a DL technology that can extract, evaluate and generate suitable cancer treatment options using large clinical data amounts from patients' medical records. However, it is evident that AI-advances are recent and still at infancy, illustrating a long road ahead to adopting and implementing AI-powered clinical decision support in cancer care, let alone personalised cancer care.

13.4 AI-Enabled Adaptive Cancer Therapy

AI has also been reported to play a role in adaptive therapy. In standard cancer therapy, the maximum tolerated doses remove drug-sensitive cancer cells. However, this might not apply to drug-resistant cells. Adaptive therapy is being explored to combat this challenge, using dose-reduction AI-algorithms with an ultimate goal of preventing the outnumbering of drug-sensitive cells by drug-resistant cells (Chmielecki et al., 2011; Jedeszko et al., 2015). A recent study demonstrated that AI-adaptive therapy improved mouse breast cancer model treated with paclitaxel, compared to standard established therapies with high doses (Enriquez-Navas et al., 2016). However, adaptive therapy real-world clinical oncology applications remain to be elucidated. A recent study by Zhang et al. (2017) translated adaptive therapy into a pilot clinical oncology study. This research group evaluated adaptive therapy on a cohort of prostate cancer (PCa) patients receiving abiraterone hormone therapy (Zhang et al., 2017). In this study, PCa patients received 47% of standard abiraterone dose. At reporting time, 1/11 adaptive therapy trial participants experienced tumour progression. This reporting aided in median estimation of time to

progression, using the prostate-specific antigen (PSA) biomarker and radiographic imaging of more than 27 months old. This was better than 11 months of PSA and 16.5 months of radiographic imaging for PCa patients on uninterrupted abiraterone therapy (Zhang et al., 2017). In advancing personalised care through adaptive therapy, population-based dose adjustments may yield superior results when customising this therapy to each patient's response to therapy. Mathematical models are reported to be suitable tools in facilitating adaptive therapy in cancer management.

Adaptive therapy involves consecutive treatment selections based on surveyed tumour burden to achieve an intended outcome, which could be cure or ailment regulation (Gatenby & Brown, 2020). Reinforcement learning (Alpaydın, 2004; Sutton & Barto, 2018) is a potentially beneficial AI-method for therapeutic agent shortlisting and sequencing in the setting of adaptive cancer treatment. Given a consistently apparent ailment condition, the objective is to choose a treatment series that will influence the malignancy to a lag state, like remission, while optimising an incentive function detailed by the clinical outcome and toxic effects. In the language of reinforcement learning, the treatment choice-maker is referred to as the agent acting on the disease, or the environment, by administering a series of treatments, each of which is chosen depending on the state of the illness while ensuring a maximum incremental reward function for the decision-makers conduct. The incentive may be detailed because of information-fuelled forecasts of the toxic effects of treatment and findings to steer therapy choice towards the objective of enhancing survival rates with negligible treatment toxicity as a clinical cut-off. Nonetheless, clinical trials on adaptive therapeutic interventions in malignancy care may be required to crystallise the bounds and processes of an efficient and secure AI-empowered adaptive cancer treatment, Fig. 13.3. Adaptive therapy clinical applications face numerous obstacles, including minimising treatment toxicity and acquiring a statistically valid burden of disease projections throughout care delivery, from detection and initial therapy to cancer management.

13.5 Challenges and Limitations

Challenges to AI-application in oncology include the lack of transparency of AI-algorithms associated with inadequacy of quality annotated data availability at population levels. Furthermore, tumour heterogeneity and evolutionary dynamics make tumour therapy responses unique, to every individual and populations. Thus AI-platforms trained on closed systems (such as a population) may not yield reliable treatment recommendations when placed to perform as open systems. Thus AI-empowered decision support in recommending most effective treatment approaches should be evaluated. Challenges surrounding AI-implementation in real-world cancer management systems also include lack of adequate data, access to the already available data, socio-and-legal issues and high costs. How exactly, like

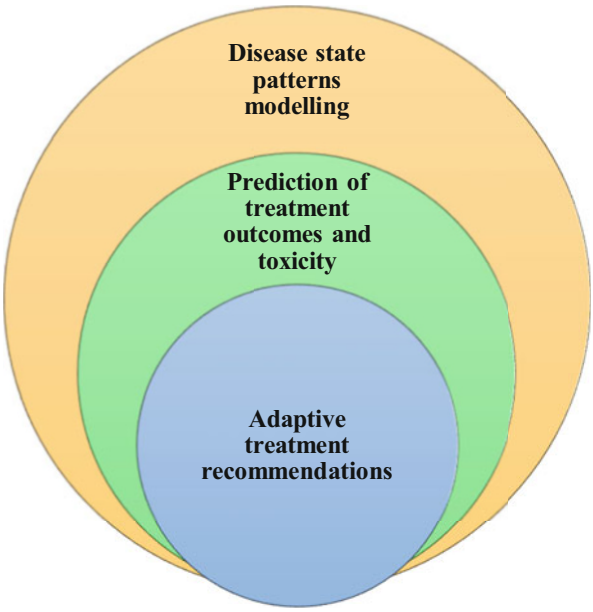


Fig. 13.3 Adaptive cancer treatment. Treatment can be adapted based on the results of the disease sate patterns modelling. The modelling allows the prediction of treatment and toxicity. Based on the results, recommendations on the treatments are made

humans do, does AI-platform reach conclusions—the black-box phenomenon is also a challenge, in cancer therapy and decision-making.

Unprecedented anti-cancer drug toxicity is another challenge in anti-cancer drug discovery. In traditional drug development processes, approved and trial compounds are frequently co-delivered preclinically and clinically. This is so done to manage multiple drug targets while improving efficacy of treatment. The successive augmentation studies of clinical dose are done to attain drug synergy. This drug synergy is expected to have increased efficacy than those of individual drugs. In most cases unfortunately, off-target drug effects can impede with the approval of the drug (Lin et al., 2019). On the other hand, successfully designed compounds can demonstrate efficacy when delivered at sub-optimal doses, while attempting to elude drug toxicity. Ideally, optimised drug-combination concurrently best links compounds and their doses, with minimised toxicity while aiming at the right targets. Experimentally, testing all these parameters for optimal combination therapy is almost unfeasible. Even though AI-tools can reduce this burden by significantly reducing the number of experiments, ethical, socio-economic and AI-associated high costs especially in low-middle income countries (LMICs) may be constraining. Additionally, the existing therapeutic interventions such as chemotherapy still far supersede the proposed innovative therapy approaches such as adaptive therapy. Furthermore, inclusive medical genomics' studies may also be a significant limitation towards successful global implementation of AI-decision support making in oncology personalised care. Nonetheless, novel AI-approaches hold the potential to improving therapeutic efficacy in personalised cancer care.

13.6 Conclusions

AI-enhanced decision support systems in recommending the most effective therapeutic approaches for individual cancer patients hold the potential to maximising therapeutic efficacy. Challenges currently facing cancer therapeutic approaches present as complex. These include a need to customise bespoke regimens that will benefit each cancer patients. Furthermore, wet-lab and clinical oncology workflows are insufficient to meeting these specific needs, and thus there is an existing gap that AI-based tools hold great potential to bridging. AI-platforms such as CURATIVE. AI and QPOP already show great promise towards personalised care through unique drug combination and improved patient outcomes. Like all new technologies, these platforms still need to be validated in larger patient cohorts and monitored for drifting in various populations. These AI-platforms are not meant to replace human efforts but to augment human capabilities. Furthermore, adaptive therapy is key in the advancement of precision oncology and AI-technology is key to the success of this therapy. In addition to unique, optimal and customised drug combination for individual patients, AI plays a key role in drug discovery, design, development and administration and may help alleviate costs associated with new drug development, common in oncology therapeutics. It is a long road ahead to effectively employ AI in oncology therapy decision support systems. However, emerging studies demonstrate the potential of AI in solidifying the road towards precision oncology by recommending and enhancing most optimum therapeutic approaches in personalised care, Fig. 13.4.

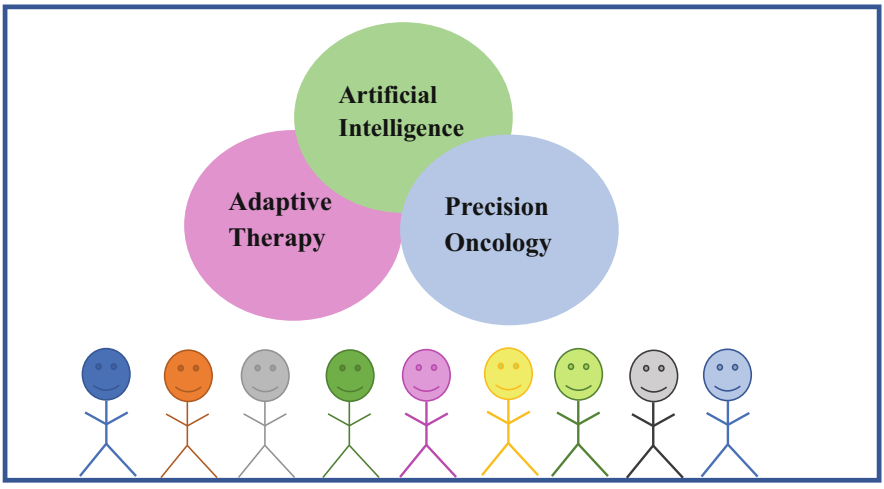


Fig. 13.4 AI-enhanced therapeutic efficacy. AI holds the potential to facilitate adaptive therapy in advancing precision oncology by optimising unique drug combinations, drug dose and safety administration thus recommending the most effective therapeutic approaches

References

Abbasi, J. (2019). "Electronic nose" predicts immunotherapy response. *JAMA, 322*, 1756.

Abernethy, A. P., Etheredge, L. M., Ganz, P. A., Wallace, P., German, R. R., Neti, C., Bach, P. B., & Murphy, S. B. (2010). Rapid-learning system for cancer care. *Journal of Clinical Oncology, 28*, 4268–4274.

Alpaydin, E. (2004). Introduction to machine learning.

Babier, A., Boutilier, J. J., Mcniven, A. L., & Chan, T. C. Y. (2018). Knowledge-based automated planning for oropharyngeal cancer. *Medical Physics, 45*, 2875–2883.

Bahl, M., Barzilay, R., Yedidia, A. B., Locascio, N. J., Yu, L., & Lehman, C. D. (2018). High-risk breast lesions: A machine learning model to predict pathologic upgrade and reduce unnecessary surgical excision. *Radiology, 286*, 810–818.

Ballester, P. J. (2019). Machine learning for molecular modelling in drug design. *Biomolecules, 9*.

Beck, J. T., Rammage, M., Jackson, G. P., Preininger, A. M., Dankwa-Mullan, I., Roebuck, M. C., Torres, A., Holtzen, H., Coverdill, S. E., Williamson, M. P., Chau, Q., Rhee, K., & Vinegra, M. (2020). Artificial intelligence tool for optimizing eligibility screening for clinical trials in a large community cancer center. *JCO Clinical Cancer Informatics***,** 50–59.

Blackledge, M. D., Winfield, J. M., Miah, A., Strauss, D., Thway, K., Morgan, V. A., Collins, D. J., Koh, D. M., Leach, M. O., & Messiou, C. (2019). Supervised machine-learning enables segmentation and evaluation of heterogeneous post-treatment changes in multi-parametric MRI of soft-tissue sarcoma. *Frontiers in Oncology, 9*, 941.

Bogani, G., Rossetti, D., Ditto, A., Martinelli, F., Chiappa, V., Mosca, L., Leone Roberti Maggiore, U., Ferla, S., Lorusso, D., & Raspagliesi, F. (2018). Artificial intelligence weights the importance of factors predicting complete cytoreduction at secondary cytoreductive surgery for recurrent ovarian cancer. *Journal of Gynecologic Oncology, 29*, e66.

Boon, I. S., Au Yong, T. P. T., & Boon, C. S. (2018). Assessing the role of Artificial Intelligence (AI) in clinical oncology: utility of machine learning in radiotherapy target volume delineation. *Medicine, 5*.

Brown, N. (2015). *In silico medicinal chemistry: Computational methods to support drug design*. Royal Society of Chemistry.

Chan, H. C. S., Shan, H., Dahoun, T., Vogel, H., & Yuan, S. (2019). Advancing drug discovery via artificial intelligence. *Trends in Pharmacological Sciences, 40*, 592–604.

Chen, G., Tsoi, A., Xu, H., & Zheng, W. J. (2018). Predict effective drug combination by deep belief network and ontology fingerprints. *Journal of Biomedical Informatics, 85*, 149–154.

Chmielecki, J., Foo, J., Oxnard, G. R., Hutchinson, K., Ohashi, K., Somwar, R., Wang, L., Amato, K. R., Arcila, M., & Sos, M. L. (2011). Optimization of dosing for EGFR-mutant non-small cell lung cancer with evolutionary cancer modeling. *Science Translational Medicine, 3*, 90ra59.

Ciallella, H. L., & Zhu, H. (2019). Advancing computational toxicology in the big data era by artificial intelligence: Data-driven and mechanism-driven modeling for chemical toxicity. *Chemical Research in Toxicology, 32*, 536–547.

Cuplov, V., & André, N. (2020). Machine learning approach to forecast chemotherapy-induced haematological toxicities in patients with rhabdomyosarcoma. *Cancers, 12*, 1944.

Cypko, M. A., Stoehr, M., Kozniewski, M., Druzdzel, M. J., Dietz, A., Berliner, L., & Lemke, H. U. (2017). Validation workflow for a clinical Bayesian network model in multidisciplinary decision making in head and neck oncology treatment. *International Journal of Computer Assisted Radiology and Surgery, 12*, 1959–1970.

Derbal, Y. (2022). Can artificial intelligence improve cancer treatments? *Health Informatics Journal, 28*, 14604582221102314.

Dorman, S. N., Baranova, K., Knoll, J. H. M., Urquhart, B. L., Mariani, G., Carcangiu, M. L., & Rogan, P. K. (2016). Genomic signatures for paclitaxel and gemcitabine resistance in breast cancer derived by machine learning. *Molecular Oncology, 10*, 85–100.

El Naqa, I., Kosorok, M. R., Jin, J., Mierzwa, M., & Ten Haken, R. K. (2018). Prospects and challenges for clinical decision support in the era of big data. *JCO Clinical Cancer Informatics*, 1–12.

Elfiky, A. A., Pany, M. J., Parikh, R. B., & Obermeyer, Z. (2018). Development and application of a machine learning approach to assess short-term mortality risk among patients with cancer starting chemotherapy. *JAMA Network Open, 1*, e180926–e180926.

Enriquez-Navas, P. M., Kam, Y., Das, T., Hassan, S., Silva, A., Foroutan, P., Ruiz, E., Martinez, G., Minton, S., & Gillies, R. J. (2016). Exploiting evolutionary principles to prolong tumor control in preclinical models of breast cancer. *Science Translational Medicine, 8*, 327ra24.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*, 115–118.

Farina, E., Nabhen, J. J., Dacoregio, M. I., Batalini, F., & Moraes, F. Y. (2022). An overview of artificial intelligence in oncology. *Future Science OA, 8*, Fso787.

Fiorino, C., Guckenberger, M., Schwarz, M., Van der Heide, U. A., & Heijmen, B. (2020). Technology-driven research for radiotherapy innovation. *Molecular Oncology, 14*, 1500–1513.

Gatenby, R. A., & Brown, J. S. (2020). Integrating evolutionary dynamics into cancer therapy. *Nature Reviews Clinical Oncology, 17*, 675–686.

Gertych, A., Swiderska-Chadaj, Z., Ma, Z., Ing, N., Markiewicz, T., Cierniak, S., Salemi, H., Guzman, S., Walts, A. E., & Knudsen, B. S. (2019). Convolutional neural networks can accurately distinguish four histologic growth patterns of lung adenocarcinoma in digital slides. *Scientific Reports, 9*, 1483.

Goecks, J., Jalili, V., Heiser, L. M., & Gray, J. W. (2020). How machine learning will transform biomedicine. *Cell, 181*, 92–101.

Golden, J. A. (2017). Deep learning algorithms for detection of lymph node metastases from breast cancer: Helping artificial intelligence be seen. *JAMA, 318*, 2184–2186.

Goldenberg, S. L., Nir, G., & Salcudean, S. E. (2019). A new era: Artificial intelligence and machine learning in prostate cancer. *Nature Reviews Urology, 16*, 391–403.

Gulhan, D. C., Lee, J. J.-K., Melloni, G. E. M., Cortés-Ciriano, I., & Park, P. J. (2019). Detecting the mutational signature of homologous recombination deficiency in clinical samples. *Nature Genetics, 51*, 912–919.

Günakan, E., Atan, S., Haberal, A. N., Küçükyildiz, İ. A., Gökçe, E., & Ayhan, A. (2019). A novel prediction method for lymph node involvement in endometrial cancer: Machine learning. *International Journal of Gynecologic Cancer, 29*, 320.

Harrer, S., Shah, P., Antony, B., & Hu, J. (2019). Artificial intelligence for clinical trial design. *Trends in Pharmacological Sciences, 40*, 577–591.

Ho, D. (2020). Artificial intelligence in cancer therapy. *Science, 367*, 982–983.

Hood, L., & Auffray, C. (2013). Participatory medicine: A driving force for revolutionizing healthcare. *Genome Medicine, 5*, 110.

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer, 18*, 500–510.

Hossain, M. A., Islam, S. M. S., Quinn, J. M. W., Huq, F., & Moni, M. A. (2019). Machine learning and bioinformatics models to identify gene expression patterns of ovarian cancer associated with disease progression and mortality. *Journal of Biomedical Informatics, 100*, 103313.

Houssami, N., Kirkpatrick-Jones, G., Noguchi, N., & Lee, C. I. (2019). Artificial Intelligence (AI) for the early detection of breast cancer: A scoping review to assess AI's potential in breast screening practice. *Expert Review of Medical Devices, 16*, 351–362.

Hu, L., Bell, D., Antani, S., Xue, Z., Yu, K., Horning, M. P., Gachuhi, N., Wilson, B., Jaiswal, M. S., Befano, B., Long, L. R., Herrero, R., Einstein, M. H., Burk, R. D., Demarco, M., Gage, J. C., Rodriguez, A. C., Wentzensen, N., & Schiffman, M. (2019). An observational study of deep learning and automated evaluation of cervical images for cancer screening. *JNCI: Journal of the National Cancer Institute, 111*, 923–932.

Isaksson, L. J., Pepa, M., Zaffaroni, M., Marvaso, G., Alterio, D., Volpe, S., Corrao, G., Augugliaro, M., Starzyńska, A., & Leonardi, M. C. (2020). Machine learning-based models for prediction of toxicity outcomes in radiotherapy. *Frontiers in Oncology, 10*, 790.

Jabbari, P., & Rezaei, N. (2019). Artificial intelligence and immunotherapy. *Expert Review of Clinical Immunology, 15*, 689–691.

Jedeszko, C., Paez-Ribes, M., Di Desidero, T., Man, S., Lee, C. R., Xu, P., Bjarnason, G. A., Bocci, G., & Kerbel, R. S. (2015). Postsurgical adjuvant or metastatic renal cell carcinoma therapy models reveal potent antitumor activity of metronomic oral topotecan with pazopanib. *Science Translational Medicine, 7*, 282ra50.

Jiang, X., Wells, A., Brufsky, A., & Neapolitan, R. (2019). A clinical decision support system learned from data to personalize treatment recommendations towards preventing breast cancer metastasis. *PLoS One, 14*, e0213292.

Klambauer, G., Hochreiter, S., & Rarey, M. (2019). Machine learning in drug discovery. *Journal of Chemical Information and Modeling, 59*, 945–946.

Koromina, M., Pandi, M.-T., & Patrinos, G. P. (2019). Rethinking drug repositioning and development with artificial intelligence, machine learning, and omics. *OMICS: A Journal of Integrative Biology, 23*, 539–548.

Leventakos, K., Helgeson, J., Mansfield, A. S., Deering, E., Schwecke, A., Adjei, A., Molina, J., Hocum, C., Halfdanarson, T., Marks, R., Parikh, K., Pomerleau, K., Coverdill, S., Rammage, M., & Haddad, T. (2019). Implementation of artificial intelligence (AI) for lung cancer clinical trial matching in a tertiary cancer center. *Annals of Oncology, 30*, ii74.

Levine, M. N., Alexander, G., Sathiyapalan, A., Agrawal, A., & Pond, G. (2019). Learning health system for breast cancer: Pilot project experience. *JCO Clinical Cancer Informatics*, 1–11.

Li, X., Hu, B., Li, H., & You, B. (2019). Application of artificial intelligence in the diagnosis of multiple primary lung cancer. *Thoracic Cancer, 10*, 2168–2174.

Li, Q., Qi, L., Feng, Q.-X., Liu, C., Sun, S.-W., Zhang, J., Yang, G., Ge, Y.-Q., Zhang, Y.-D., & Liu, X.-S. (2019). Machine learning-based computational models derived from large-scale radiographic-radiomic images can help predict adverse histopathological status of gastric cancer. *Clinical and Translational Gastroenterology*, *10*.

Liang, G., Fan, W., Luo, H., & Zhu, X. (2020). The emerging roles of artificial intelligence in cancer drug development and precision therapy. *Biomedicine & Pharmacotherapy, 128*, 110255.

Lin, L., Dou, Q., Jin, Y.-M., Zhou, G.-Q., Tang, Y.-Q., Chen, W.-L., Su, B.-A., Liu, F., Tao, C.-J., Jiang, N., Li, J.-Y., Tang, L.-L., Xie, C.-M., Huang, S.-M., Ma, J., Heng, P.-A., Wee, J. T. S., Chua, M. L. K., Chen, H., & Sun, Y. (2019). Deep learning for automated contouring of primary tumor volumes by MRI for nasopharyngeal carcinoma. *Radiology, 291*, 677–686.

Lin, A., Giuliano, C. J., Palladino, A., John, K. M., Abramowicz, C., Yuan, M. L., Sausville, E. L., Lukow, D. A., Liu, L., & Chait, A. R. (2019). Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Science Translational Medicine, 11*, eaaw8412.

Lind, A. P., & Anderson, P. C. (2019). Predicting drug activity against cancer cells by random forest models based on minimal genomic information and chemical properties. *PLoS One, 14*, e0219774.

Liu, C., Liu, X., Wu, F., Xie, M., Feng, Y., & Hu, C. (2018). Using artificial intelligence (Watson for oncology) for treatment recommendations amongst Chinese patients with lung cancer: Feasibility study. *Journal of Medical Internet Research, 20*, e11087.

Liu, C., Qi, L., Feng, Q.-X., Sun, S.-W., Zhang, Y.-D., & Liu, X.-S. (2019). Performance of a machine learning-based decision model to help clinicians decide the extent of lymphadenectomy (D1 vs. D2) in gastric cancer before surgical resection. *Abdominal Radiology, 44*, 3019–3029.

Lou, B., Doken, S., Zhuang, T., Wingerter, D., Gidwani, M., Mistry, N., Ladic, L., Kamen, A., & Abazeed, M. E. (2019). An image-based deep learning framework for individualising radiotherapy dose: A retrospective analysis of outcome prediction. *The Lancet Digital Health, 1*, e136–e147.

Martins, J., Magalhães, C., Rocha, M., & Osório, N. S. (2019). Machine learning-enhanced T cell neoepitope discovery for immunotherapy design. *Cancer Informatics, 18*, 1176935119852081.

Mcdonald, J. F. (2018). Back to the future — The integration of big data with machine learning is re-establishing the importance of predictive correlations in ovarian cancer diagnostics and therapeutics. *Gynecologic Oncology, 149*, 230–231.

Menden, M. P., Iorio, F., Garnett, M., Mcdermott, U., Benes, C. H., Ballester, P. J., & Saez-Rodriguez, J. (2013). Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties. *PLoS One, 8*, e61318.

Meng, C., Hu, Y., Zhang, Y., & Guo, F. (2020). PSBP-SVM: A machine learning-based computational identifier for predicting polystyrene binding peptides. *Frontiers in Bioengineering and Biotechnology, 8*, 245.

Meyer, P., Noblet, V., Mazzara, C., & Lallement, A. (2018). Survey on deep learning for radiotherapy. *Computers in Biology and Medicine, 98*, 126–146.

Nascimento, A. C. A., Prudêncio, R. B. C., & Costa, I. G. (2019). A drug-target network-based supervised machine learning repurposing method allowing the use of multiple heterogeneous information sources. *Methods in Molecular Biology, 1903*, 281–289.

Oyaga-Iriarte, E., Insausti, A., Sayar, O., & Aldaz, A. (2019). Prediction of irinotecan toxicity in metastatic colorectal cancer patients based on machine learning models with pharmacokinetic parameters. *Journal of Pharmacological Sciences, 140*, 20–25.

Paik, E. S., Lee, J.-W., Park, J.-Y., Kim, J.-H., Kim, M., Kim, T.-J., Choi, C. H., Kim, B.-G., Bae, D.-S., & Seo, S. W. (2019). Prediction of survival outcomes in patients with epithelial ovarian cancer using machine learning methods. *Journal of Gynecologic Oncology, 30*.

Pantuck, A. J., Lee, D.-K., Kee, T., Wang, P., Lakhotia, S., Silverman, M. H., Mathis, C., Drakaki, A., Belldegrun, A. S., Ho, C.-M., & Ho, D. (2018). Modulating BET bromodomain inhibitor ZEN-3694 and enzalutamide combination dosing in a metastatic prostate cancer patient using CURATE.AI, an artificial intelligence platform. *Advanced Therapeutics, 1*, 1800104.

Peng, H., Dong, D., Fang, M.-J., Li, L., Tang, L.-L., Chen, L., Li, W.-F., Mao, Y.-P., Fan, W., Liu, L.-Z., Tian, L., Lin, A.-H., Sun, Y., Tian, J., & Ma, J. (2019). Prognostic value of deep learning PET/CT-based radiomics: Potential role for future individual induction chemotherapy in advanced nasopharyngeal carcinoma. *Clinical Cancer Research, 25*, 4271–4279.

Pereira, J. C., Caffarena, E. R., & Dos Santos, C. N. (2016). Boosting docking-based virtual screening with deep learning. *Journal of Chemical Information and Modeling, 56*, 2495–2506.

Printz, C. (2017). Artificial intelligence platform for oncology could assist in treatment decisions. *Cancer, 123*, 905–905.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*, 1347–1358.

Rashid, M. B. M. A., Toh, T. B., Hooi, L., Silva, A., Zhang, Y., Tan, P. F., Teh, A. L., Karnani, N., Jha, S., & Ho, C.-M. (2018). Optimizing drug combinations against multiple myeloma using a quadratic phenotypic optimization platform (QPOP). *Science Translational Medicine, 10*, eaan0941.

Rocque, G. B., Lennes, I. T., Rhoads, K. F., & Yu, P. P. (2019). Quality improvement at the health system level: Measurement, risk stratification, disparity populations, and governance. *American Society of Clinical Oncology Educational Book, 39*, 388–398.

Sesen, M. B., Nicholson, A. E., Banares-Alcantara, R., Kadir, T., & Brady, M. (2013). Bayesian networks for clinical decision support in lung cancer care. *PLoS One, 8*, e82349.

Sharma, A., & Rani, R. (2020). Ensembled machine learning framework for drug sensitivity prediction. *IET Systems Biology, 14*, 39–46.

Sherbet, G. V., Woo, W. L., & Dlay, S. (2018). Application of artificial intelligence-based technology in cancer management: A commentary on the deployment of artificial neural networks. *Anticancer Research, 38*, 6607–6613.

Simon, A. B., Vitzthum, L. K., & Mell, L. K. (2020). Challenge of directly comparing imaging-based diagnoses made by machine learning algorithms with those made by human clinicians. *Journal of Clinical Oncology, 38*, 1868–1869.

Smaïl-Tabbone, M., & Rance, B. (2019). Contributions from the 2018 literature on bioinformatics and translational informatics. *Yearbook of Medical Informatics, 28*, 190–193.

Stanzione, A., Cuocolo, R., Del Grosso, R., Nardiello, A., Romeo, V., Travaglino, A., Raffone, A., Bifulco, G., Zullo, F., Insabato, L., Maurea, S., & Mainenti, P. P. (2021). Deep myometrial infiltration of endometrial cancer on MRI: A radiomics-powered machine learning pilot study. *Academic Radiology, 28*, 737–744.

Steele, S. R., Bilchik, A., Johnson, E. K., Nissan, A., Peoples, G. E., Eberhardt, J. S., Kalina, P., Petersen, B., Brücher, B., Protic, M., Avital, I., & Stojadinovic, A. (2014). Time-dependent estimates of recurrence and survival in colon cancer: Clinical decision support system tool development for adjuvant therapy and oncological outcome assessment. *The American Surgeon, 80*, 441–453.

Sun, R., Limkin, E. J., Vakalopoulou, M., Dercle, L., Champiat, S., Han, S. R., Verlingue, L., Brandao, D., Lancia, A., Ammari, S., Hollebecque, A., Scoazec, J.-Y., Marabelle, A., Massard, C., Soria, J.-C., Robert, C., Paragios, N., Deutsch, E., & Ferté, C. (2018). 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. *The Lancet Oncology, 19*, 1180–1191.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. The MIT Press.

Tan, S., Li, D., & Zhu, X. (2020). Cancer immunotherapy: Pros, cons and beyond. *Biomedicine & Pharmacotherapy, 124*, 109821.

Tang, X., Huang, Y., Lei, J., Luo, H., & Zhu, X. (2019). The single-cell sequencing: New developments and medical applications. *Cell & Bioscience, 9*, 53.

Taninaga, J., Nishiyama, Y., Fujibayashi, K., Gunji, T., Sasabe, N., Iijima, K., & Naito, T. (2019). Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: A case-control study. *Scientific Reports, 9*, 12384.

Tartar, A., Akan, A., & Kilic, N. (2014). A novel approach to malignant-benign classification of pulmonary nodules by using ensemble learning classifiers. *IEEE*, 4651–4654.

Trebeschi, S., Drago, S. G., Birkbak, N. J., Kurilova, I., Călin, A. M., Delli Pizzi, A., Lalezari, F., Lambregts, D. M. J., Rohaan, M. W., Parmar, C., Rozeman, E. A., Hartemink, K. J., Swanton, C., Haanen, J. B. A. G., Blank, C. U., Smit, E. F., Beets-Tan, R. G. H., & Aerts, H. J. W. L. (2019). Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Annals of Oncology, 30*, 998–1004.

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., & Zhao, S. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery, 18*, 463–477.

Van der Waal, I. (2018). Skin cancer diagnosed using artificial intelligence on clinical images. *Oral Diseases, 24*, 873–874.

Walsh, S., De Jong, E. E. C., van Timmeren, J. E., Ibrahim, A., Compter, I., Peerlings, J., Sanduleanu, S., Refaee, T., Keek, S., Larue, R. T. H. M., Van Wijk, Y., Even, A. J. G., Jochems, A., Barakat, M. S., Leijenaar, R. T. H., & Lambin, P. (2019). Decision support systems in oncology. *JCO Clinical Cancer Informatics*, 1–9.

Wang, Y., Wang, Z., Xu, J., Li, J., Li, S., Zhang, M., & Yang, D. (2018). Systematic identification of non-coding pharmacogenomic landscape in cancer. *Nature Communications, 9*, 1–15.

Watson, O. P., Cortes-Ciriano, I., Taylor, A. R., & Watson, J. A. (2019). A decision-theoretic approach to the evaluation of machine learning algorithms in computational drug discovery. *Bioinformatics, 35*, 4656–4663.

Wei, J. W., Tafe, L. J., Linnik, Y. A., Vaickus, L. J., Tomita, N., & Hassanpour, S. (2019). Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. *Scientific Reports, 9*, 3358.

West, J., You, L., Zhang, J., Gatenby, R. A., Brown, J. S., Newton, P. K., & Anderson, A. R. A. (2020). Towards multidrug adaptive therapy. *Cancer Research, 80*, 1578–1589.

Witteveen, A., Nane, G. F., Vliegen, I. M. H., Siesling, S., & Ijzerman, M. J. (2018). Comparison of logistic regression and Bayesian networks for risk prediction of breast cancer recurrence. *Medical Decision Making, 38*, 822–833.

Wong, C. H., Siah, K. W., & Lo, A. W. (2019). Estimation of clinical trial success rates and related parameters. *Biostatistics, 20*, 273–286.

Zarrinpar, A., Lee, D.-K., Silva, A., Datta, N., Kee, T., Eriksen, C., Weigle, K., Agopian, V., Kaldas, F., & Farmer, D. (2016). Individualizing liver transplant immunosuppression using a phenotypic personalized medicine platform. *Science Translational Medicine, 8*, 333ra49.

Zhang, J., Cunningham, J. J., Brown, J. S., & Gatenby, R. A. (2017). Integrating evolutionary dynamics into treatment of metastatic castrate-resistant prostate cancer. *Nature Communications, 8*, 1–9.

Zhavoronkov, A., Ivanenkov, Y. A., Aliper, A., Veselov, M. S., Aladinskiy, V. A., Aladinskaya, A. V., Terentiev, V. A., Polykovskiy, D. A., Kuznetsov, M. D., & Asadulaev, A. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nature Biotechnology, 37*, 1038–1040.

Zhu, H. (2020). Big data and artificial intelligence modeling for drug discovery. *Annual Review of Pharmacology and Toxicology, 60*, 573.

Zhu, X., Lin, M. C. M., Fan, W., Tian, L., Wang, J., Ng, S. S., Wang, M., Kung, H., & Li, D. (2012). An intronic polymorphism in GRP78 improves chemotherapeutic prediction in non-small cell lung cancer. *Chest, 141*, 1466–1472.

Chapter 14
AI-Pathway Companion in Clinical Decision Support: Enabling Personalized and Standardized Care Along Care Pathways in Oncology

Meshack Bida, Mpho Kgomo, Bahoueli Gaudji, Demetra Demetriou, and Zodwa Dlamini

Abstract The practice of medicine generates data in the management of patients to make decisions. Often health care providers rely on their problem-solving skills, judgements and experience. AI-Pathway companion is now the foundation of precision medicine and provides advanced analytics to manage the vast amount of data. In the past, practising medicine was focused on general solutions that allow safe treatments of most patients with similar symptoms. Thus, the working methods within the medical community were on a generalized basis. The experimental and experience-based approach will be replaced by the evidence-based approach with improved prognosis, analysis, diagnosis and treatment methods. Disruptive technologies, such as genome sequencing and advanced biotechnology, generate vast amount of data, impossible for an average mind to remember everything, thus requiring AI. These technologies allow patients to participate in the decision-making process about the management of their condition, using hand-held devices. Deep

M. Bida (✉)
Department of Anatomical Pathology, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
e-mail: meshack.bida@nhls.ac.za

M. Kgomo
Department of Gastroenterology, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
e-mail: mpho.kgomo@up.ac.za

B. Gaudji
Department of Urology, Faculty of Health Sciences, Steve Biko Academic Hospital, University of Pretoria, Pretoria, South Africa
e-mail: guy.gaudji@up.ac.za

D. Demetriou · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: dd.demetriou@up.ac.za; zodwa.dlamini@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_14

learning algorithms assist physicians in patient care. However, there is a need to combine these algorithms with human expertise. By combining the knowledge of deep learning systems and the human factor, efficacy is increased. It is important to combine knowledge and AI in the medical field.

Keywords Clinical decision · Support systems · A1-Pathway companion · Personalized care · Standardized pathways · Oncology

14.1 Clinical Decision Support Systems

14.1.1 Defining Clinical Decision Support Systems (CDSS)

Clinical decision support systems are computer-based programs. These programs can analyse data within the electronic health records (EHRs) that can aid medical professionals through the evidence-based approach (Häyrinen et al., 2008). The functions by CDSS include an alarm system, diagnostics, prescription control, disease management and drug control, to name a few. They can be demonstrated as reminders and alerts, guidelines, reports, order sets, workflow tools or templates (Sutton et al., 2020). There are various types of clinical decision support systems and can include diagnostic support systems like quick medical reference (QMR) and MYCIN, Arden Syntax for alerts or reminders and patient management systems (Martin Pusic & Ansermino, 2004). The key components to clinical decision support systems include a knowledge-base, an inference mechanism, for example, an algorithm and patient data from various sources, see Fig. 14.1.

The patients' history is the first step of data collection which is then prepared for modelling before the application of predictive analytics decision support. Critical decisions can be made from the predicted outcomes. Predicted outcomes are compared with the actual outcomes to determine the accuracy of machine learning process. The number of patient encounters can be recorded with each prediction made until the predicted outcomes match closely the actual outcome, see Fig. 14.1.

In low-resource settings, there are decreased amounts of patient-to-physician ratios, poor health infrastructure and a limited specialist. This causes a strain on the already over-burdened healthcare systems. Clinical decision support systems can alleviate this burden and support physicians in their daily routines. The system that is suited for the low-resource setting has been predominantly focused on infection control and maternal care which does not require deep learning (Dani Kiyasseh et al., 2022). This highlights the need for further investigation to include various health conditions and to improve patient outcomes. The CDSSs can be electronic or manual. Regarding the electronic system, the results show clinical application and in a particular order (see Fig. 14.2). Throughout this stage, multiple electronic systems are outlined.

There are two main clinical decision support systems: manual and electronic. The manual system consists of two categories, viz. the Early Warning scores and the Plate-based algorithms. The electronic system consists of multiple disease

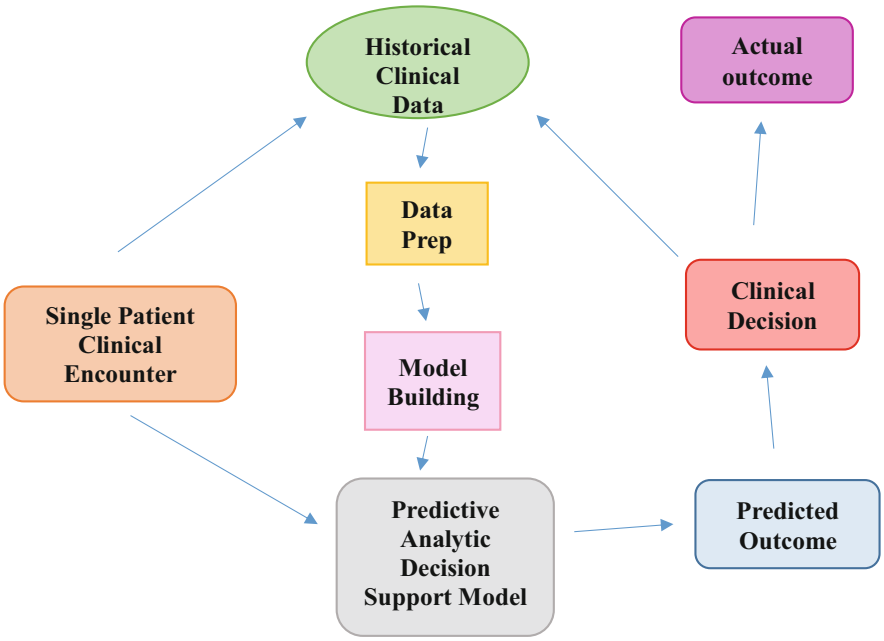


Fig. 14.1 Clinical decision support systems

categories, such as bacteria, parasitic, the antimicrobial resistance and viral diseases. The electronic systems category will provide electronic alerts and messages, questionnaires and digital scoring.

14.1.2 Clinical Decision Support Systems in Clinical Practice and Clinical Trials

Clinical decision support has been successfully used to systematically review clinical trials in identifying features critical to success (Kawamoto et al., 2005). A nationwide multi-institutional audit assessment can be done using quality indicators to pick up preventable medical errors. Thus, adverse events can be adequately monitored to prevent deaths using these clinical decision support systems. To address the issues in healthcare, organizations are increasingly using clinical decision support systems to assist with recommendations for clinical decision-making and to provide patient-specific assessments. These support systems may be computer-based or manual and can provide reminders to patient's charts. Computerized systems can provide patient-specific recommendations (Steinbrook, 2009). These systems reduce medication errors and improve prescribing practices, delivery services and adherence standard care. Compared with other approaches for improving practice, these systems are more effective and will be long-lasting.

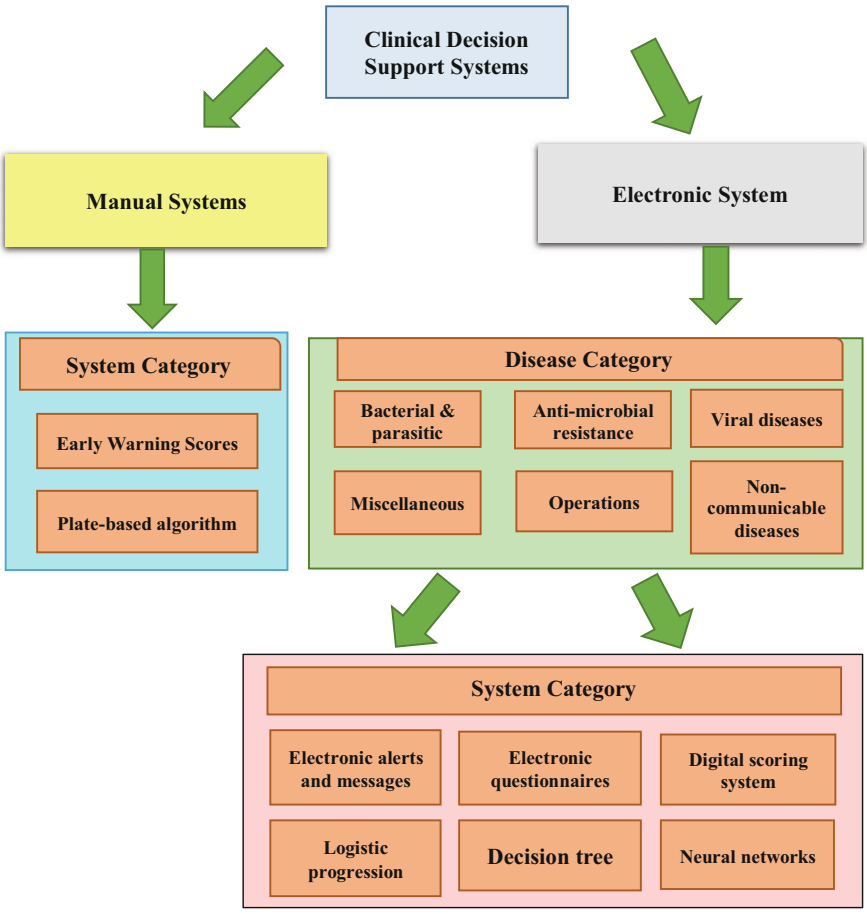


Fig. 14.2 Chart that delineates clinical decision support systems (CDSSs) and how it presents the findings

14.1.3 Clinical Decision Support Systems Feature in Clinical Practice

Computer-based systems improve clinical practice regarding the following:

- Automatic provision.
- Recommendations and not only assessments.
- Decision support regarding location and time.
- Computer-based decision support (Holroyd-Leduc et al., 2011).

14.1.4 Clinical Decision Support Systems in Pathology

The manual review of pathology test results for missed diagnoses is laborious, inaccurate and time-consuming. An automated solution has decreased the errors involved and improved the accuracy for clinical decision support. The system focused on microbiology test results that included the strain information and antibiotic sensitivities. Both factors play a role in clinical care and ongoing patient safety. The system is highly effective at identifying abnormal test results. Furthermore, the system used information of discharged summaries to identify patient follow-ups. This shows the positive results in using the system. The system increased efficiency, accuracy and supports patient safety by allowing diagnosis and accurate treatment (Blumenthal, 2010; Holroyd-Leduc et al., 2011).

There are two aspects regarding decision-making in anatomical pathology, namely decision support systems and decision analysis. Record keeping is problematic and to be prepared for all possibilities is impossible. The analysis of the various data from clinical trials and other genomic studies requires specialized, multidisciplinary knowledge. Pathologists have minimal training in this regard. The use of decision-making systems and AI decreases the associated challenges. Light microscopy is a key level in cancer management and histopathology expertise will be of importance (Hendrickson & Balzer, 2011).

The two main types of CDSS are evidence-based (knowledge-base) and non-evidence-based. Various scientists and medical members like pathologists are interested in the application of automated clinical decision support to laboratory medicine and pathology. The application has enormous ability to optimize laboratory test selection, analysis and correlation with existing data. This will transform laboratory medicine from an observational field to a specialized field allowing precise and comprehensive diagnosis. Anatomic pathology can improve diagnostic information (Jason M. Baron et al., 2014).

This transformation of pathology practice does not only assist the reduction of waste or errors often seen in test selection and result interpretation but also improves diagnostic precision. Advances in next-generation sequencing, laboratory automation, mass spectrometry and other technologies will improve diagnosis, prognosis and treatment. Substantial improvements in the diagnosis came from the use, results and analysis of data from using existing and traditional technologies. However, these can be improved with the use of AI. Clinical decision support avoids unwanted testing and ensures that the correct tests are used. This allows misinterpretation of test results to be avoided. The application of AI to clinical and laboratory data can show important insight and patterns that overpower manual interpretation. Many of the barriers in evolving the clinical application of NGS rely on analysis and interpretation of the data.

Transforming the focus of pathology to enhance data analysis and precision medicine will improve the value of laboratory testing. Data generation can become routine, but accurate data extraction will become more complex. For pathology and laboratory services to be cost-effective, the focus will need to be adjusted to the

generating of diagnostic, prognostic and therapeutic information. Laboratories will also have to be involved with test selection. Computational data analysis and decision support systems will aid in integrating this evolution in laboratory and pathology diagnosis, prognosis and treatment (Jason M. Baron et al., 2014). The symbiosis between computers and humans will optimize diagnostic efficiency.

CDSSs show a paradigm shift in healthcare in modern times. CDSSs are used to aid clinicians in the decision-making processes that can be very complexed. CDSSs have evolved since their first use in the 1980s. They are commonly used through electronic medical records and other computerized clinical workflows. Despite the benefits of CDSS, there are still unknown factors, for example the use and cost effects. Numerous publications show the success of CDSS, but challenges have also been recorded (Holroyd-Leduc et al., 2011).

14.2 AI-Pathway Companion in CDSS: Cancer Control and Prevention Includes Awareness, Screening and Early Diagnosis

14.2.1 AI-Pathway Companion in Clinical Decision Support

AI-Pathway Companion clinical decision support solution allows standard and precision care in infectious diseases, oncology and cardiology. The CDSS aims to facilitate treatment and diagnosis to improve patient outcomes and increase the survival rate (Blake et al., 2011). The system shows important patient data without the need to access multiple files and databases. The key information is immediately available for consultation purposes or collaboration with other specialists. AI technologies provide further insights for the patient and the medical staff. The AI-Pathway Companion compares the patient's medical status against known guidelines and assists in the next steps in diagnosis and therapy (Henkel et al., 2022a, 2022b) (Fig. 14.3).

The patients are at the centre of adverse changes around them. Diseases and sample amounts have been on the rise. The decrease in sample analysis results from the use of outdated diagnostic tools, manual and subjective analysis, burdened patients, overworked pathologists, decreased number of specialists and increased sample volumes. These factors influence the result turnover time. The rising risk of disease and sample volume is major cause of disequilibrium. This results in a negative cascade around the centre as there are overworked pathologists, a shortage of specialists and increased patient numbers.

14.2.2 What Is an AI-Pathway Companion?

AI-Pathway Companion consists of several medical devices and health products in development. In some countries several such disease-specific products of AI-Pathway Companion already exist which cover the following:

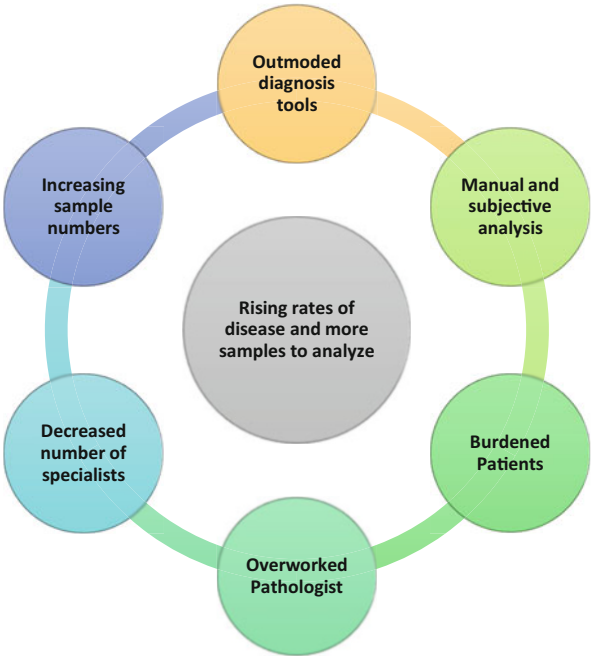


Fig. 14.3 Factors influencing the increase of disease and result turnover time

14.2.3 AI-Pathway Prostate Cancer

The AI-Pathway Companion Prostate Cancer VA10B (Siemens Healthcare GmbH) manufactured in Erlangen, Germany, has demonstrated success in providing patient information in comprehensive dashboards on information effectiveness, quality and satisfaction (Henkel et al., 2022a, 2022b). The variable indicators showed remarkable improvement in this software tool. This system provides recommendations and supports diagnosis or therapeutic options for prostate cancer patients utilizing clinical guidelines in correlation with the patient's current disease condition. Using the CDSS connector system that is integrated with the patient information system automatically collects all the relevant patient data from all the necessary source of information systems (Fig. 14.4).

PACS is imported into the landing zone. In the landing zone, it is converted into a uniformed data model and sent to the AIPC database. The database provides the patient graph that is imported into AIPC applications. The applications provide the report that is stored as the electronic health report. The electronic health report is used with PACS and the process starts all over again. The system is suited to handle large volume of data. The implementation of this software has significantly reduced consultation preparation times in prostate cancer management and effectively improved the decision-making process and customer satisfaction. However, mapping the more complex patient pathways, such as the follow-up treatment of prostate cancer, still requires refinement and is subject to further research, especially investigating the effect in post-therapeutic prostate cancer management.

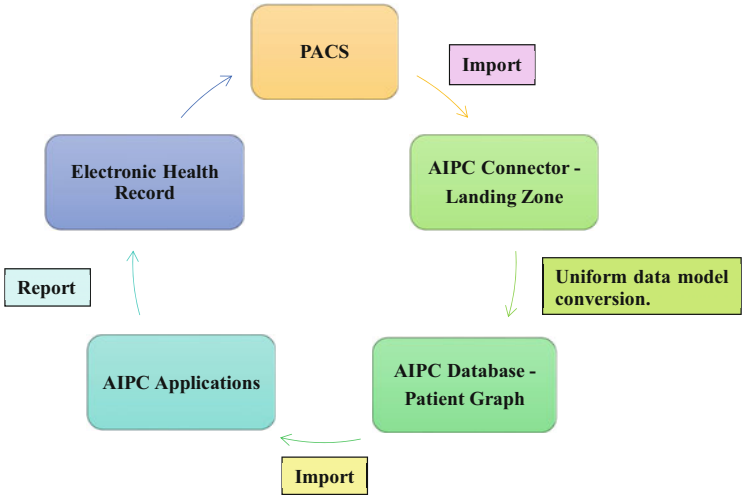


Fig. 14.4 The web-based AI-Pathway companion product suite of several components consists of a picture archiving and communication system (PACS), AI-Pathway Connector, AIPC database and AIPC Applications

14.2.4 AI-Pathway Companion Breast Cancer

The possibility of digitizing whole-slide images of tissue has simplified the incorporation of AI using machine learning software tools in digital pathology. The sensitivity and specificity of this application magnify the subvisual morphometric phenotypes and ultimately improve pathologists in diagnosis, prognostication and therefore patient management (Kaustav Bera et al., 2019). The use of AI in the detection of breast cancer using mammography has also been tested but was found to be not sufficiently specific (Karoline Freeman et al., 2021).

14.2.5 AI-Pathway Companion Coronary Artery Disease

Siemens Healthiness is currently in collaboration with a UK-based group in the National Health System (NHS) to develop a prototype AI-Pathway Companion application for patients with suspected or known CAD. It is envisaged that the AI application will automatically interpret scans and help identify patients at low risk to shorten hospital stays. This is being complemented by creating computer-aided software to aid clinical decisions and diagnosis. The tool will foster individualized precision medicine and improve clinical and operational outcomes in the NHS (LONDON, K.S.C., n.d.).

14.2.6 AI-Pathway Companion Infectious Diseases

The implementation of AI predictive analytics tools in the Ministry of Health (MoH) in Malaysia focused on controlling infectious diseases. This has for the first time provided the MoH with the capability to predict diseases breakout and spreading (USAID, n.d.). The MoH fully digitized and integrated EMRs from health facilities across the country which enhances its analysis further by using AI. Thus, using the AI tools, the MoH was able to map health issues and disease outbreaks occurring across the country through the application of AI to identify correlations among multiple variables across complex data sets to identify risk factors and predict the spread of diseases (Agrebi & Larbi, 2020).

Non-health data was also inserted to assist in the prediction of future outbreaks by using natural language processing algorithms from social media and news reports. Through the application of AI to newly digitized data, the MoH was enabled to view, analyse, interpret and react to health data in real time. The predictive ML algorithms enabled the MoH to exact the date and geolocation of the next disease outbreak months in advance. The ML-powered algorithms also enabled the MoH to decide which control interventions would be most effective and to plan the intervention rollouts.

14.3 Clinical Uses of AI-Pathway Companion

- The companion correlates, aggregates and visualizes relevant patient-specific information and other disease pathways.
- The companion enables automatic patient-specific mapping using evidence-based guidelines.
- The companion facilitates objective decision-making using data from multidisciplinary specialists on correlated patient preferences and data.
- The companion offers transparent diagnosis and treatment insights into time (Blake et al., 2011).

14.4 Cancer Prevention and Control Using COMPAS

COMPAS is a computer-based algorithm that is used in criminal cases to predict the likelihood that a defendant will re-offend. The application of AI in cancer control programs must be investigated using a similar concept of algorithm, viz. COMPASS (Leatherdale & Lee, 2019). Understanding the risk factors that start cancer is key for reducing the future burden of cancer. Most current cancer control insights are concluded from existing cohort studies and modern large-scale population laboratories (Welch & Kawamoto, 2013). Big data assets can be changed to have a greater

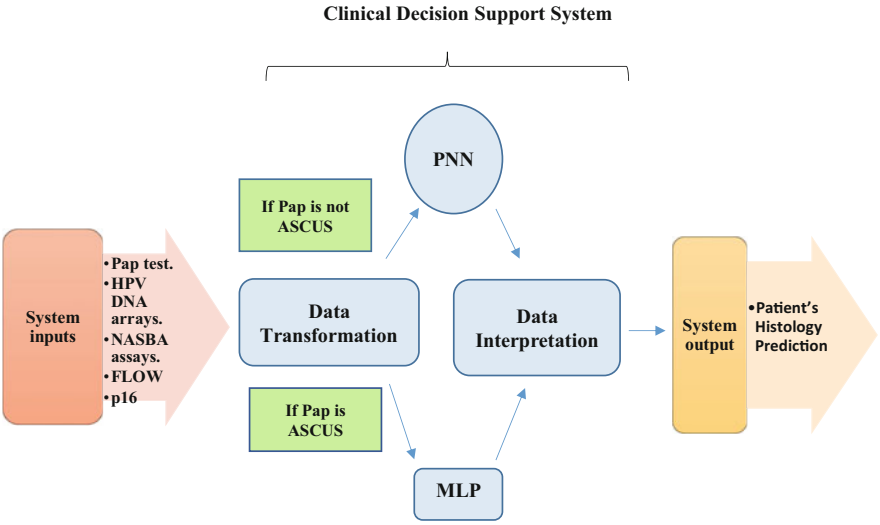


Fig. 14.5 Schematic block diagram of the proposed decision support system; Nucleic Acid Based Amplification (NASBA); Abnormal Squamous Cell of Undetermined Significance (ASCUS); Probabilistic Neural Network (PNN), Multilayer Perceptron Network (MLP)

impact on the future cancer burden by focusing on primary prevention efforts that use AI and ML. ML automatically learns patterns and can create complex models and algorithms that aid in the prediction in big data, revealing unexpected new relationships. AI has been successful in several field but the potential application in cancer prevention is unknown (Leatherdale & Lee, 2019). An example of a practical application of ML in cancer prevention is used in the Mayo clinic for cervical pap smears, see Fig. 14.5.

14.5 Overview of Clinical Decision Support Systems

The clinical decision support system consists of three modules:

- Guideline engine.
- Data module.
- Free-text processor.

The data module is used for patient information searching from the EMR. The information is kept as an edible file for guideline engines and depends on the free-text processor to interpret free-text such as Papanicolaou (Pap) reports. The free-text processor and guideline engine are basically rule-based (Kavishwar B. Wagholikar et al., 2012).

The results of the Pap test, HPV DNA assay, NASBA assay, FLOW and p16 are added into the system. The data undergoes a transformation. If the Pap test is not

ASCUS, it will result in PNN. If the Pap test is ASCUS, it will result in MLP. The data is interpreted and results in the system's output. The system can be adjusted to suit any HPV testing algorithm even in remote rural locations.

Another useful Clinical Decision Support System is the one used in community-based screening programs (see Fig. 14.5). A community-based screening program using WHO-approved algorithms of either Pap test or HPV DNA testing can be constructed. The medical data is changed into processing data by the multilayer perceptron network (MLP) or probabilistic neural network (PNN) subsystems. Depending on the Pap test value, the patient's data is promoted to the PNN of the MLP subsystem. If the Pap test is ASCUS, the data is promoted to the MLP; otherwise, the data will be promoted to the PNN. The output of each network is transformed into useful medical information. So, the Clinical Decision Support System provides predictions regarding the actual cervical status of each patient.

14.6 Machine Learning Tools

Artificial intelligence can assist with cancer detection using machine learning tools, furthermore assist in decision-making, including treatment approaches (Meskó et al., 2017). According to the National Cancer Institute (NCI), ML, AI and deep learning can be used to improve patient outcomes, survival and healthcare (Blake et al., 2011). Integration of AI can improve the speed and accuracy and diagnosis, prognosis, decision-making and survival. In low-income settings, AI-guided clinical care can play a key role in reducing health disparities. Medical professionals have shared the use of AI capabilities in cancer detection. Researchers at the Tulane University discovered that AI can accurately diagnose and detect colorectal cancer by analysing tissue scans (Singh & Graber, 2010; Singh et al., 2013). The researchers gathered approximately 13,000 images of colorectal cancer from 13 independent cancer centres in Germany, China and the USA and 8803 patients. Using random images, the researchers built a machine learning program. The program can recognize images of colorectal cancer. After creation of a performance measurement tool, the researchers compared the pathology work to the machine learning technique and models. The study showed that the average pathologist scored around 0.969 for accuracy, while the program scored 0.98. This has shown that the ML techniques are more accurate compared to manual data work by pathologists.

The researchers stated that the study will encourage pathologists to use modern technologies and more pre-screening technology to speed up diagnosis. AI can detect cancer earlier and improve the detection accuracy. New York University created an AI program that is trained to identify patterns among thousands of images to assist medical staff in diagnosis, prognosis, treatment and decision-making. AI increased the accuracy of breast cancer by 37% when tested on 44,755 ultrasound exams. The AI tool also assisted to reduce the tissue sample amount and biopsies necessary to confirm tumours by 27% (Suh et al., 2013). AI can assist radiologists to read breast ultrasounds to focus on positive cases and to avoid benign verification via

biopsy. AI can also improve existing technology regarding patient survival and outcomes. Medical professionals can use AI technology to accurately and efficiently sort through MRIs to differentiate between cancer and health patients.

14.7 Predictive Models Assist in Decision-making

By identifying risk factors, predictive models have become an important component in cancer surveillance due to the ability in determining the likelihood of cancer development (Overbeek et al., 2010). In this way, patients can be encouraged to apply preventive care strategies. According to studies by the University of Hawaii, deep learning can distinguish between the mammograms of women who are at risk to develop cancer in the future. The technology can also predict breast cancer risk by measuring the breast density. Denser breasts are associated with a higher cancer risk. However, other unknown factors can also contribute to the risk. The deep learning model is thus used to find finer details that can be linked to increased cancer risks (Bright et al., 2012).

14.8 Developing Treatment Responses

AI can be used to predict treatment responses of patients to predict adverse medication reaction. This predictive information is critical for patients and physicians when deciding treatment options to prevent life-threatening adverse reactions (Boonstra & Broekhuis, 2010). In one collaborative study, biopsy samples collected from three large were analysed in randomized clinical trials to create personalized treatment for patients with aggressive types of prostatic cancer using genetic test scores. Two-thirds of prostate cancer deaths occur in high-risk prostate cancer patients. Therefore, balancing the quality of life and survival risk is key in treatment selection. According to researchers, biomarkers can be used to identify treatment-patient benefits, create precision medicine and treatment guidelines (Meskó et al., 2017). Nguyen et al. used the Decipher biopsy test that analysed the activity of 22 genes in prostate tumours to create a scoring system that can show the aggressiveness of the cancer. The score is calculated using RNA extracted from biopsy samples during clinical trials. The score and long-term outcomes were compared and analysed (Blake et al., 2011). Using predictive analysis, the genetic signatures showed the likelihood of cancer metastases and whether the patient would die from the cancer or from external factors. It is key to acknowledge that patients will respond differently to medications (Petterson et al., 2012). ML and predictive analysis can decrease toxicity from cancer treatments that may be ineffective for the individual. The Georgia Institute of Technology and various Cancer Institutes used ML algorithms to determine drug response using cancer-fighting drugs (Brodiea et al., 2021; Milap Shah et al., 2020).

14.9 Limitations in the Application of AI in Precision Medicine

Although AI has potential, the challenges include ethical and legal factors such as responsibility for false results, the safety procedures and maintenance, economic implications and job security. Further research is required to provide accurate analysis of the challenges. AI also has limitations in healthcare. Prediction and forecasting methods are based on precedence through ML, but algorithms can underperform in some cases, for example in drug resistance. Therefore, AI cannot replace the human expertise. Data analysis must be supportive to the physicians' skills.

14.10 Conclusion

AI-Pathway companion enables personalized patient care in clinical decision support, given the vast amount of clinical data that needs to be processed by health care practitioners. It can also be used as a benchmark of standardized care for oncology patients who present in different cancer stages, various molecular targets of cancer treatment and resistance along care pathways in oncology. The strategic application of AI-Pathway companion in disease prevention, diagnosis and treatment with appropriate patient selection has the potential to revolutionize the healthcare industry in modern times.

Acknowledgements We would like to thank the South African Medical Research Council (SAMRC), Grant Number 23108, and the National Research Foundation (NRF), Grant Number 138139, for funding this work.

References

Agrebi, S., & Larbi, A. (2020). Use of artificial intelligence in infectious diseases. *Artificial Intelligence in Precision Health*. 415–38. https://doi.org/10.1016/B978-0-12-817133-2.00018-5. Epub 2020 Mar 13. PMCID: PMC7153335. https://healthitanalytics.com/features/amp/top-opportunities-for-artificial-intelligence-to-improve-cancer-care

Baron, J. M., Dighe, A. S., Arnaout, R., Balis, U. J., Stephen Black-Schaffer, W., Carter, A. B., Henricks, W. H., Higgins, J. M., Jackson, B. R., Kim, J., Klepeis, V. E., Le, L. P., Louis, D. N., Mandelker, D., Mermel, C. H., Michaelson, J. S., Nagarajan, R., Platt, M. E., Quinn, A. M., Rao, L., Shirts, B. H., & Gilbertson, J. R. (2014, Jan). The 2013 symposium on pathology data integration and clinical decision support and the current state of field. *Journal of Pathology Informatics, 5*(1), 2. https://doi.org/10.4103/2153-3539.126145

Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V., & Madabhushi, A. (2019, November). Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nature Reviews. Clinical Oncology, 16*(11), 703–715.

Blake, P. M., Decker, D. A., Glennon, T. M., Liang, Y. M., Losko, S., Navin, N., et al. (2011). Toward an integrated knowledge environment to support modern oncology. *Cancer Journal, 17*(4), 257–263.

Blumenthal, D. (2010). Launching HITECH. *The New England Journal of Medicine, 362*(5), 382–385. https://doi.org/10.1056/NEJMp0912825

Boonstra, A., & Broekhuis, M. (2010). Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. *BMC Health Services Research, 10*, 231. https://doi.org/10.1186/1472-6963-10-231

Bright, T. J., Wong, A., Dhurjati, R., Bristow, E., Bastian, L., Coeytaux, R. R., et al. (2012). Effect of clinical decision-support systems: A systematic review. *Annals of Internal Medicine, 157*(1), 29–43.

Brodiea, A., Daia, N., Teohb, J. Y.-C., Decaesteckerc, K., Dasguptad, P., & Vasdeve, N. (2021). Artificial intelligence in urological oncology: An update and future applications. *Urologic Oncology: Seminars and Original Investigations, 39*, 379–399.

Freeman, K., Geppert, J., Stinton, C., Todkill, D., & Johnson, S. (2021). Use of artificial intelligence for image analysis in breast cancer screening programmes: Systematic review of test accuracy. *BMJ, 374*, n1872.

Häyrinen, K., Saranto, K., & Nykänen, P. (2008). Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics, 77*(5), 291–304.

Hendrickson, M., & Balzer, B. (2011). Decision analysis and decision support systems in anatomic pathology. In A. Marchevsky & M. Wick (Eds.), *Evidence based pathology and laboratory medicine*. Springer. https://doi.org/10.1007/978-1-4419-1030-1_10

Maurice Henkel, Tobias Horn, Francois Leboutte, Pawel Trotsenko, Sarah Gina Dugas, Sarah Ursula Sutter, Georg Ficht, Christian Engesser, Marc Matthias, Aurelien Stalder, Jan Ebbing, Philip Cornford, Helge Seifert, & Bram Stieltjes. (2022a, July). Christian Wetteraue: Initial experience with AI Pathway Companion: Evaluation of dashboard-enhanced clinical. *Research Article*.

Henkel M, Horn T, Leboutte F, Trotsenko P, Dugas SG, Sutter SU, Ficht G, Engesser C, Matthias M, Stalder A, Ebbing J, Cornford P, Seifert H, Stieltjes B, & Wetterauer C. Initial experience with AI Pathway Companion: Evaluation of dashboard-enhanced clinical decision making in prostate cancer screening. *PLoS One*, 2022b, Jul;17(7):e0271183. https://doi.org/10.1371/journal.pone.0271183. PMID: 35857753; PMCID: PMC9299327.

Holroyd-Leduc, J. M., Lorenzetti, D., Straus, S. E., Sykes, L., & Quan, H. (2011). The impact of the electronic medical record on structure, process, and outcomes within primary care: A systematic review of the evidence. *Journal of the American Medical Informatics Association, 18*(6), 732–737.

Kensaku Kawamoto, Caitlin A. Houlihan, E. Andrew Balas, & David F. Loubach. (2005). Improving clinical practice using clinical decision support system: A systematic review of trials to identify features critical to success: Information practice. *BMJ*, https://doi.org/10.1136/bmj.38398.500764.8F (published 14 March 2005).

Kiyasseh, D., Zhu, T., & Clifton, D. (2022, January). The promise of clinical decision support systems targetting low-resource settings. *IEEE Reviews in Biomedical Engineering, 15*, 354–371.

Leatherdale, S. T., & Lee, J. (2019 Jul). Artificial intelligence (AI) and cancer prevention: the potential application of AI in cancer control programming needs to be explored in population laboratories such as COMPASS. *Cancer Causes & Control, 30*(7), 671–675.

LONDON, K. S. C. London Medical Imaging and AI Centre for Value Based Healthcare [Online]. Retrieved September 15, 2022, from https://www.kcl.ac.uk/bmeis/research-impact/london-medical-imaging-and-ai-centre-for-value-based-healthcare

Martin Pusic, J., & Ansermino, M. (2004, June). Clinical decision support systems. *BCMJ, 46*(5), 236–239.

Meskó, B., Drobni, Z., Bényei, É., Gergely, B., & Gyorffy, Z. (2017). Digital health is a cultural transformation of traditional healthcare. *mHealth, 3*, 38. https://doi.org/10.21037/mhealth.2017.08.07

Overbeek, L. I., Hermens, R. P., Van Krieken, J. H., Adang, E. M., Casparie, M., Nagengast, F. M., et al. (2010). Electronic reminders for pathologists promote recognition of patients at risk for Lynch syndrome: cluster-randomised controlled trial. *Virchows Archive, 456*(6), 653–659.

Petterson, S. M., Liaw, W. R., Phillips, R. L., Rabin, D. L., Meyers, D. S., & Bazemore, A. W. (2012). Projecting US primary care physician workforce needs: 2010–2025. *Annals of Family Medicine, 10*(6), 503–509.

Shah, M., Naik, N., Somani, B. K., & Zeeshan Hameed, B. M. (2020). Artificial intelligence (AI) in urology-Current use and future directions: An iTRUE study: Turk. *The Journal of Urology, 46*(Supp. 1), S27–S39.

Singh, H., Giardina, T. D., Meyer, A. N., Forjuoh, S. N., Reis, M. D., & Thomas, E. J. (2013). Types and origins of diagnostic errors in primary care settings. *JAMA International Medicine, 173*(6), 418–425.

Singh, H., & Graber, M. (2010). Reducing diagnostic error through medical home-based primary care reform. *Journal of the American Medical Association, 304*(4), 463–464.

Steinbrook, R. (2009). Health care and the American recovery and reinvestment act. *The New England Journal of Medicine, 360*(11), 1057–1060.

Suh, K. S., Sarojini, S., Youssif, M., Nalley, K., Milinovikj, N., Elloumi, F., et al. (2013). Tissue banking, bioinformatics, and electronic medical records: The front end requirements for personalized medicine. *Journal of Oncology, 2013*, 368751.

Sutton, R. T., Pincock, D., Baumgart, D. C., et al. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *npj Digital Medicine, 3*, 17. https://doi.org/10.1038/s41746-020-0221-y

USAID. Artificial intelligence in global health: Defining a collective path forward. [Online]. USAID's Center for Innovation and Impact. Retrieved September 27, 2022, from https://www.usaid.gov/cii/ai-in-global-health

Wagholikar, K. B., MacLaughlin, K. L., Henry, M. R., Greenes, R. A., Hankey, R. A., Liu, H., & Chaudhry, R. (2012). Clinical decision support with automated text processing for cervical cancer screening. *Journal of the American Medical Informatics Association, 19*, 833–839.

Welch, B. M., & Kawamoto, K. (2013). Clinical decision support for genetically guided personalized medicine: A systematic review. *Journal of the American Medical Informatics Association, 20*(2), 388–400.

Chapter 15
AI Tools Offering Cancer Clinical Applications for Risk Predictor, Early Detection, Diagnosis, and Accurate Prognosis: Perspectives in Personalised Care

Richard Khanyile, Rahaba Marima, Mandisa Mbeje, Shingai Mutambirwa, Daniel Montwedi, and Zodwa Dlamini

Abstract Artificial intelligence (AI) is transforming the medical research and clinical workflow by enhancing oncology clinical applications. AI-based tools are emerging as key role players in advancing precision oncology by improving oncology clinical applications in cancer risk prediction, early detection and diagnosis and accurate prognosis. Although there are challenges with every newly developed technology, efforts and significant investments have been placed to ensure the success of this technology. Additionally, the introduction of sophisticated AI-medical devices demonstrates the fundamental role that AI holds to offer in oncology. Several AI-tools have illustrated high performance towards cancer care

R. Khanyile (✉) · M. Mbeje
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Department of Medical Oncology, Steve Biko Academic Hospital and University of Pretoria, Pretoria, South Africa
e-mail: richard.khayile@up.ac.za; mandisa.mbeje@up.ac.za

R. Marima · Z. Dlamini
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: rahaba.marima@up.ac.za; zodwa.dlamini@up.ac.za

S. Mutambirwa
Department of Urology, Sefako Makgatho Health Sciences University (SMU), Ga-Rankuwa, South Africa
e-mail: shingai.mutambirwa@smu.ac.za

D. Montwedi
Department of General Surgery, Steve Biko Academic Hospital and University of Pretoria, Pretoria, South Africa
e-mail: daniel.montwedi@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_15

and management in various parts of the world. While risk prediction, early detection, diagnosis and accurate prognosis are a work in progress in some cancer types, this remains a challenge in various cancers. However, AI-based tools can advance human efforts with the overall aim of improving oncology patient outcome through personalised care. This chapter will focus on AI-based tools in advancing oncology personalised care by improving risk prediction, early detection and diagnosis, and accurate prognosis. Challenges in the application of AI-based tools from bench to bedside will also be discussed, while providing an overview of AI-based tools for predicting clinically relevant parameters in advancing precision oncology.

Keywords Artificial intelligence · Deep learning (DL) · Precision oncology · Early detection · Diagnosis · Accurate prognosis · Clinical applications

15.1 Introduction

Significant advances have been made in the past decades in cancer screening diagnosis and management. However, there are still existing challenges in providing personalised cancer care. Thus, the convergence of Artificial Intelligence (AI) tools and precision oncology is a promising potent tool in overcoming these challenges. AI-based tools have the potential to improve high-risk populations, early detection and diagnosis, as well as accurate prognosis in advancing precision oncology (Farina et al., 2022). Early cancer diagnosis is still a global challenge and efficient screening strategies have been reported to be limited by various factors which may include financial support and public buy-in. Some of the initiatives have demonstrated inadequate covering of the majority of at-risk populations (Ahnen et al., 2014). On the other hand, enhancing screening initiatives with the lack of evidence-based indication can also lead to a waste of valuable resources and a substantial financial burden, particularly in resource-limited public health systems such as in the low-middle income countries (LMICs) (Verma et al., 2018). Although data science applications are underdeveloped in LMICs such as African countries, various catalytic factors are already in place. These factors may include cloud computing developments, significant investments in digitising health information and smartphone penetration robustness (Waljee et al., 2022; Holst et al., 2020). Furthermore, the United Nations (UN) promotes the centralisation of AI to achieve its Sustainable Development Goals (Parker et al., 2020). Additionally, the US National Institutes of Health has invested about US$75 million over a period of 5 years in advancing data science, health innovation and discoveries across Africa in its new programme, Health Discovery and Innovation in Africa (DSI-Africa) (1). AI technologies have advanced in healthcare, offering novel approaches in effective cancer screening, diagnosis and prognosis (Qiu et al., 2022; Zhu et al., 2020; Dlamini et al., 2020). Following cancer screening, timely and accurate diagnosis are keys in identifying suitable treatment for effective cancer care management.

AI is categorised into 3 classes: Generalised AI (with general human intellectual), Super AI (AI that exceeds human abilities) and Narrow AI (AI with narrow

specialised capabilities). Machine can learn from narrow AI to figure out the most complicated biological processes that may not be figured by humans. Compared to humans, narrow AI is designed to learn and is not driven by emotions (Iqbal et al., 2021). Typical examples of narrow AI are Siri by Apple, Google assistance, Alexa by Amazon and Microsoft's Cortana. Most of these narrow AI-tools process the human input (either by language or given data) and respond accordingly. These tools, Artificial Narrow Intelligent (ANI) tools, work within a pre-defined range. For example, when asking Siri what day it is, an accurate response is given because the task is within Siri's pre-defined AI range. Similarly, the advanced self-driving cars also operate within their artificial narrow intelligence range (Aron, 2011; Kepuska & Bohouta, 2018; Goksel Canbek & Mutlu, 2016; Brill et al., 2019). In health care, it has been reported that Nvidia, a leading US-based multinational technology company announced its building efforts of AI supercomputer for drug design and medical research (Buitrago et al., 2020; Yujuan et al., 2020; Kochanny & Pearson, 2021). Such efforts would require domain-specific expertise, such as 'precision oncology'. The AI-powered tools in clinical applications remain to be largely elucidated. This chapter will discuss AI-based tools in clinical oncology workflows that include risk prediction, early detection and diagnosis, accurate prognosis in advancing precision oncology.

15.2 AI-based Tools in Clinical Oncology Workflows

AI-based tools in health, medicine and clinical oncology are advancing. For example, the sophisticated medical devices such as robots in health care. These may include the care-bots that have been specifically trained to care for aged patients and assisting surgeons in surgery (Larson et al., 2014). Additionally, in Scotland, National Health Services (NHS) 24, an AI-based clinical assessment service is in testing at clinical phase. This tool is meant to assist the community with minor issues of health telephonically (McCartney, 2018). Furthermore, AI-based tools have also been reported to have high accuracy in the determination of infection-related tumours and recommended most appropriate therapeutic strategies (Leibovici et al., 2007). Figure 15.1 illustrates AI-based tools in offering key clinical oncology applications. These include risk prediction, screening, detection, early diagnosis and prognosis. These AI-tools depend on clinical data such as pathology, radiology and omics.

In oncology workflows, AI is mostly applicable in pathology and radiology imaging fields (Jha & Topol, 2016). In radiology, DL algorithms may be used in cancer detection, classification, segmentation, characterisation and monitoring (Bi et al., 2019; Hosny et al., 2018). In cancer screening studies, image classification is necessary. AI-based tools can aid radiologists in the classification of small lesions. AI can also help radiologists in creating improved oncology workflows by the determination of high priority report groups, thereby achieving better outcomes. For example, various radiology studies demonstrated improved mammography

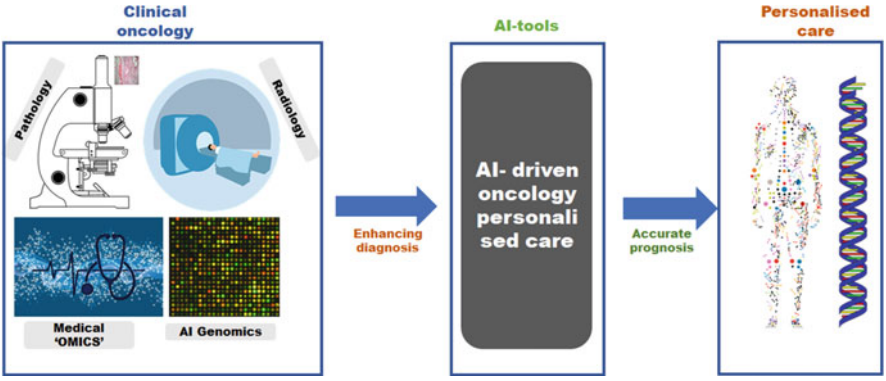


Fig. 15.1 Clinical applications of AI-tools in oncology. AI-tools can be applied in oncology imaging modalities such as pathology, radiology. AI-powered genomics and medical omics are key in AI-enhanced cancer early diagnosis and accurate prognosis in precision oncology

screening for breast cancer by combining with AI-based tools (Ghosh, 2019; Schaffter et al., 2020). AI-based tools in cancer screening are a growing field, demonstrating advances in various cancer types (Miotto et al., 2016; Nartowt et al., 2019, 2020; Hart et al., 2018; Muhammad et al., 2019; Roffman et al., 2018; Stark et al., 2019). With regard to detection, it has been reported that AI-based tools can aid to identify cancerous lesions that are highly likely to be missed by humans. For example, AI-based tools can be used in finding lung nodules and can distinguish cancerous from non-cancerous lesions and identify brain metastasis on MRI scans (Schultheiss et al., 2020). Thus, AI-based tools enhance physicians' medical imaging processing in cancer detection and are not meant to replace human effort.

Cancer segmentation can be challenging with traditional approaches. However, AI-based tools have been showed to alleviate this problem (Qiu et al., 2022). Cancer segmentation aids in classifying individual pixels based on their organs or lesions by particularly recognising lesions and accessing lesions' size and volume (Shaver et al., 2019). DL methods can be used to reveal rare cancer characteristics and patterns from medical images. Radiomics, a field that studies the extraction and analyses of large amounts of advanced quantitative image features with the intent of creating mineable databases from radiological images, is emerging in advancing precision oncology. Radiomics plays an important role in cancer characterisation and can inform ML models that can effectively predict treatment response (Avanzo et al., 2020). Furthermore, radiomics has been reported to be applied in various cancers such as liver, lung and brain tumours (Dreher et al., 2020; Kocher et al., 2020). Radiomics can be used to predict clinically relevant oncology parameters, as medical imaging can be used as a prognostic information source. Additionally, patient genomic data can also be used for prognostic purposes (Li Wen & Leech, 2020; Sakellaropoulos et al., 2019).

Furthermore, DL using brain MRI radiomic features has been reported to distinguish brain gliomas from brain metastasis, with a matched neuroradiologist trained performance. In addition, AI-based tools have great potential in cancer monitoring, due to AI's capability to detect hidden discriminative features that may not be read by humans (Bi et al., 2019). Generative adversarial networks (GANs) are unique AI-models that are able to generate a new set of images based on various data types. For example, GANs have been reported to generate synthetic CT images from MRI images. GAN holds potential in enhancing radiotherapy planning (Maspero et al., 2020). In addition, GAN has been proved valuable in prostate cancer, by automating dose distribution for intensity-modulated radiation therapy (IMRT) (Murakami et al., 2020). Additionally, autoencoders (VEs) and variational autoencoders (VAEs) generative networks have been proposed to have the potential of reducing radiation dose and intravenous contrast use by improving MRI and CT multimodality imaging (Haubold et al., 2021; Katsari et al., 2021).

15.3 AI-Models for Predicting Clinically Relevant Parameters in Advancing Precision Oncology

Unstructured data are a common data type used in recording qualitative and subjective information that may be acquired through imaging modalities or patient–practitioner interactions. Artificial neural networks (ANN) have been reported to be useful tools in the analysis of unstructured data (Wang et al., 2019). Contrarily, convolutional neural networks (CNNs) are used frequently in the imaging files. Various AI-models have been developed globally to advance oncology personalised care. However, valid steps must be followed when developing ML models that can be used in clinical practice. These include choosing the correct problems, collection of data and pre-processing (anonymity of data, for example), machine training, internal validation, testing and optimising, evaluation and external validation (Wiens et al., 2019), Fig. 15.2. Every step is critical to the optimal functioning of the ML model. Following the deployment of any model, the results and application must be continually monitored for drifting, loss of performance, etc., as quality check and ensuring consistency of the same ML model. Furthermore, the clinical applicability and utility of the newly developed ML models must be assessed and validated in possible clinical trials using the commonly applied metrics for task classifications in medicine. Receiver Operating Characteristic Curve (ROC curve) is the most commonly used metric in health and medicine. ROC plots the true positive rate and the false positive rate. The area under ROC (AUROC) then expresses the accuracy level. Additionally, the confusion matrix is used to assess the model's sensitivity, specificity and precision (Handelman et al., 2019; Ranschaert et al., 2019). This may be applicable in precision oncology, improving risk prediction in high-risk populations, early detection and diagnosis, and accurate prognosis.

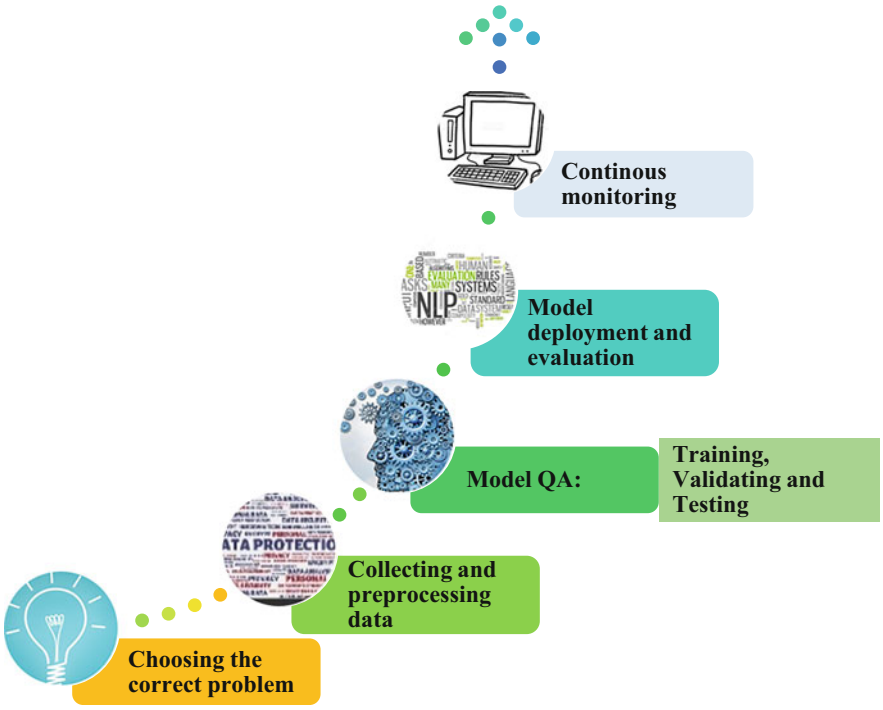


Fig. 15.2 Developing the ML model. The development of new ML models goes through various steps which may begin by choosing the correct problems, training, validating, testing and evaluating the model. Validation can be internal and external. Continuous monitoring for drifting and loss of performance purposes should be done following model deployment (https://depositphotos.com/2120339/stock-photo-desktop-computer-isolated.html)

15.4 AI-Enhanced Technologies in Early Cancer Detection, Diagnosis, Risk Stratification and Prognosis

AI-based tools have been reported to improve cancer diagnosis. For example, AI-enhanced colonoscopy has been demonstrated to be an effective intervention in the identification of benign polyps (Mori et al., 2020). This AI-enhanced colonoscopy approach has the potential of healthcare-cost reduction, as well as preventing invasive treatment approaches. In cervical cancer screening, ML algorithms have been reported to have high accuracy of precancerous lesions prediction. AI-based tools thus have the potential to minimise invasive treatment interventions, optimise cancer diagnosis and minimise over-treatment in patients and thus improve accurate prognosis (Hu et al., 2019; Shaffer, 2018). Notably, similar to other solid tumours, significant advances have been made in AI-powered CRC precision oncology. A study by van Rijn et al. (2006) showed that the miss rate for traditional colonoscopy for any size polyp was 22%. This miss rate was reported to significantly increase for smaller polyps (van Rijn et al., 2006). However, AI-aided colonoscopy has been

reported to improve the detection of even smaller polyps, characterise polyps and minimise variation that may be caused by different clinicians (Ozawa et al., 2020; Qadir et al., 2020; Urban et al., 2018; Fernández-Esparrach et al., 2016; Lee et al., 2021; Zhang et al., 2017; Tian et al., 2019; Takemura et al., 2012). Prediction models or statistical algorithms such as Multianalyte Assays with Algorithmic Analysis (MAAA) have been shown to detect pre-symptomatic longitudinal complete blood count (CBC) patterns that may be undetectable to clinicians in CRC (Colón-Franco et al., 2018). Together with demographic data, the MAAA models can be used to identify high-risk CRC patients (Schneider et al., 2020; Hornbrook et al., 2017; Kinar et al., 2016; Ayling et al., 2021). Accurate prognosis and prediction of precise treatment response outcome form an integral part of precision oncology (Huang et al., 2020), even though the AI-based tools in cancer prognosis are less common compared to the use of medical statistics. Different studies in several countries have reported the use of AI-powered prognosis prediction in various cancers. To date, AI-based tools in the advancement of precision oncology are evident, although associated challenges may exist. Yamada et al. (2019) developed an AI-diagnostic model with the aim of bridging disparities between different clinicians and improving early CRC detection. This system is reported to detect early CRC signs during colonoscopy. Additionally, the sensitivity and specificity of this model were reported to be 97.3% and 99%, with 0.975 AUC. This model has the ability to notify clinicians in real-time of even non-polypoid polyps, avoiding missed diagnosis. Tables 15.1, 15.2 and 15.3 provide a summary of AI applications in cancer screening, diagnosis and prognosis.

15.5 AI from Bench to Bedside: Challenges and Limitations

Although AI-based algorithms hold significant potential towards transforming cancer care, their translational application into an oncology clinical workflow still has a long road. These challenges include data collection and training, lack of potential clinical validation that can be attributed to the aforementioned challenges, ethical and legal regulatory guidelines, user education limitations and high cost associated with AI-models development in LMICs (Chua et al., 2021; Patel et al., 2020). However, AI-based tools already form part of daily routine tasks, e.g., the use of smartphone applications such as Siri. Such tools can be integrated into smartphones as risk assessment tools that can provide cancer risk assessment outcome to the public. These initiatives were particularly observed during the Covid-19 self-assessment applications. Patients with high-risk outcomes can be recommended for further medical care, even though lack of symptoms in early cancer detection and diagnosis may be a challenge. The 'Big Data' generated by advancing cancer research initiatives sometimes possess as a challenge to clinicians as they try to apply such recommendations in their clinical workflows. Furthermore, it has been demonstrated that captured data from oncology health care workers is complex. This includes the Doctor's notes (typed or hand-written), imaging data, pathology data, laboratory

Table 15.1 Summary of AI applications for CRC screening (Qiu et al., 2022)

AI-Model	Year	Task	Dataset	Performance	Reference
LightGBM DNN	2021	Detection of high-risk patients	Microarray data from 111 patients with 22,278 features included.	100% accuracy	Nazari et al. (2021)
SSD	2021	Polyp classification	47,555 images taken from endoscopies of 24 patients.	Accuracy: 0.9067, precision: 0.9744, recall: 0.9067, F1: 0.9393	Lee et al. (2021)
Random Tree, RF, LMT, SVM	2021	Detection of serum biomarker	186 blood serum samples made up of 90 CRC, 39 advanced adenomas and 57 healthy individuals.	75% accuracy	Pan et al. (2021)
RF, LR, SVM, DT, Gaussian NB, and extremely randomised trees	2020	Detection of serum biomarker	263 blood serum protein samples where 213 samples were obtained from individuals undergoing screening endoscopy and 50 non-metastatic CRC.	AUC: 0.75, 70% sensitivity 89% specificity	Ivancic et al. (2020)
CNN	2020	Detection and classification of polyp	27,508 endoscopy images.	Detection: Sensitivity—0.92, PPV—0.86 Classification: Sensitivity—0.83, PPV—0.81	Ozawa et al. (2020)
RetinaNet	2020	Polyp localisation	EAD2019, CVC-ClinicDB, ETIS-Larib, in-house dataset, Kvasir-SEG	Precision: 0.537	Kayser et al. (2020)
Faster R-CNN, SSD	2020	Detection of polyp	ASU-Mayo Clinic, CVC-CLINIC, CVC-ClinicVideoDB	Precision: 0.8154, Sensitivity: 0.9086, F1: 0.8595	Qadir et al. (2020)
ResNet50, RetinaNet	2019	Detection and classification of polyp	871 images taken from endoscopies of 218 patients.	F1: 0.6872, F2: 0.6607	Tian et al. (2019)
CNN	2018	Detection of polyp	8641 endoscopy images.	90.0% sensitivity, 63.3 specificity, 76.5% accuracy	Urban et al. (2018)

(continued)

Table 15.1 (continued)

AI-Model	Year	Task	Dataset	Performance	Reference
		Polyp segmentation	CVC-ColonDB	74.8% specificity, 99.3% sensitivity, 97.7% accuracy	Akbari et al. (2018)
Colonflag	2018	Prediction of high-risk patient	Colon cancer screening centre data (EMRs)	The odds of Colonflag and normal colonoscopies: 2.0	Hilsden et al. (2018)
CNN	2017	Classification of polyp	1930 NBI images	85.9% accuracy, 87.3% precision, 87.6% recall rate	Zhang et al. (2017)
Colonflag	2017	Detection of high-risk patient	112,584,133 US community-based insured data	AUC: 0.80 ± 0.01	Hornbrook et al. (2017)
Mescore	2017	Detection of a high-risk patient	17,095 patients from KPNW (EMRs)	Top 3% score > 97.02 Top 1% score > 99.38	Kinar et al. (2017)
Energy map	2016	Polyp detection	24 videos of endoscopies	AUC: 0.79, 70.4% sensitivity, 72.4% specificity	Fernández-Esparrach et al. (2016)
Mescore	2016	Detection of a high-risk patient	606,403 Israelis and 25,613 UK dataset (EMRs)	AUC: 0.82 ± 0.01 and 0.81 for validation sets	Kinar et al. (2016)
HuPAS version 3.1	2012	Polyp classification	1890 NBI endoscopic images	98.7% accuracy	Takemura et al. (2012)

data and patient generated health data. Obtaining relevant meaning from such data is depended on sufficient data extraction, interpretation, analysis and integration into the clinical workflow. Thus, the human brain capacity can be enhanced by AI-tools to process such robust information. Data processing, analysis and storage can be done by AI-based tools on the already available data. AI-tools in cancer diagnosis is a growing field with around a 100 recently registered clinical trials in AI-cancer diagnosis (Dong et al., 2020). With every new technology, specifically in health care, there are challenges and limitations that need to be overcome. Nonetheless, AI-based tools hold promising interventions in early cancer detection, diagnosis and accurately predicting prognosis, Fig. 15.3.

Table 15.2 Summary of AI-based tools in cancer diagnosis (Hunter et al., 2022)

AI-model	Type	Description	Reference
Random forest	Regression, Classification	Grouping prediction of random decision trees.	Xiao et al. (2017)
Logistic regression	Regression	Prediction of categorical outcomes using logistic function.	Chhatwal et al. (2009)
Convolutional neural network	Regression, Classification	Detecting image features using kernels.	Suh et al. (2020)
Support vector machine	Regression, Classification	Constructing hyperplanes using to optimise data separation.	Zhang et al. (2013)
Extreme gradient boosting	Regression, Classification	Similar to random forest. Chronological errors minimised by descending gradient.	Liew et al. (2021)
Naïve Bayes	Classification	Uses Bayesian probability which includes priors for classification.	Olatunji et al. (2021)
Artificial neural network	Regression, Classification	Input multiplication by weights and biases to predict outcome.	Muhammad et al., (2019)

15.6 Conclusions

Although AI methods show great promise towards improving cancer health care, in cancer diagnosis, accurate prognosis, predicting risk and response to treatment, how precisely AI improves cancer health is not fully demonstrated. Some of the challenges that face the transition of AI-models into real-world settings include fairness and bias, socio-environmental factors and data safety and privacy. Regarding fairness and bias, it has been documented that health data can sometimes be bias. For example, missing values and lack of diverse sampling. AI-models trained on such data may augment the bias, thus making unfavourable decisions and possibly discriminating against a particular group of people and exacerbating cancer health disparities. This can be mitigated by diversifying data sources. It has been reported that AI-models developed in one geographical region with unique socio-economic factors may not respond and thus display similar levels of accuracy when placed in a different socio-environmental setting. This may significantly impair the initial validated model's high standards. Additionally, AI-models developed in a particular race group may sub-optimally perform in other race groups. Furthermore, data safety and privacy issues must also be addressed. Data sharing and availability is key to any AI-model. This includes data such as medical history, genomics, behaviours data and social data that cover patients' daily lives. However, building a well-controlled and safe ecosystem for data storage, sharing and management is crucial (Johnson et al., 2021). Overall, AI-aided screening methods are most likely to detect clinically relevant parameters such as polyps in CRC which may otherwise be missed by humans and may be precancerous lesions (Hilsden et al., 2018; van Rijn et al., 2006). AI-technology also may be able to detect clinically relevant biomarkers to improve cancer patient outcome (Issa & Noureddine, 2017; van Rijn et al., 2006; Yamada et al., 2019). However, it has been reported that this technology may also increase over diagnosis of early cancer stages with no potential of malignancy, which may

Table 15.3 Summary of AI-based tools in cancer prognosis (Huang et al., 2020)

Cancer Type	AI-Method	Year	Number of Patients	Performance	Geographic Region	Reference
Breast Cancer	Multimodal DNN	2018	1980	/	China	Sun et al. (2018)
	Semi-supervised Learning Model	2013	162,500	/	USA	Park et al. (2013)
	ANN and DT	2005	433,272	Accuracy: DT (93.6%), ANN (91.2%)	USA	Delen et al. (2005)
	Dynamic Gradient BoostingMachine with GA	2019	82,707	Accuracy Improved (28%)	USA	Lu et al. (2019)
Prostate Cancer	Fuzzy Neural Network	2015	100	/	Taiwan	Kuo et al. (2015)
	SVM model	2017	/	Average Accuracy (66%)	USA	Zhang et al. (2017)
Lung Cancer	GBM, SVM	2017	10,442	RMSE (32,15.05) for GBM, SVM	USA	Lynch et al. (2017)
	SVM with RFE and RF	2018	101	Accuracy (71%, 59%)	France	Sepehri et al. (2018)
	Naive Bayes, SVM with Gaussian, etc.	2016	168	/	Italy	Yu et al. (2016)
Colorectal Cancer	Six Neural Networks	1997	334	Accuracy (> 80%), Mean Sensitivity (60%), Mean Specificity (88%)	UK	Bottaci et al. (1997)
	Semi-random Regression Tree	2019	1568	/	China	Wang et al. (2019)
	LSTM, Naïve Bayes, SVM	2018	641	Hazard Ratio (2.3); CI (95%,1.79–3.03), AUC (0.69)	Finland	Bychkov et al. (2018)
Glioblastoma	Neural Network	2018	215	Accuracy: DT (89.2%)	India	Vasudevan and Murugesan (2018)
Ovarian Cancer	SVM	2019	84	HR (0.644), CI (95%,0.436–0.952)	Taiwan	Lu et al. (2019)
	Unsupervised HierarchicalClustering	2019	364	RPV: A Novel Prognostic Signature Discovered	UK	Lu et al. (2019)
	Fuzzy Forest	2018	469	Accuracy (80.60 ± 0.5%), Sensitivity (81.40%), Specificity (76.30%)	Singapore & Malaysia	Acharya et al. (2018)

(continued)

Table 15.3 (continued)

Cancer Type	AI-Method	Year	Number of Patients	Performance	Geographic Region	Reference
Bladder Cancer	Statistical Analysis	2019	115	NEDD8: Poor Prognosis Found	China	Tian et al. (2019)
	KNN, RF, etc.	2019	3503	Sensitivity & Specificity (> 70%)	USA	Hasnain et al. (2019)

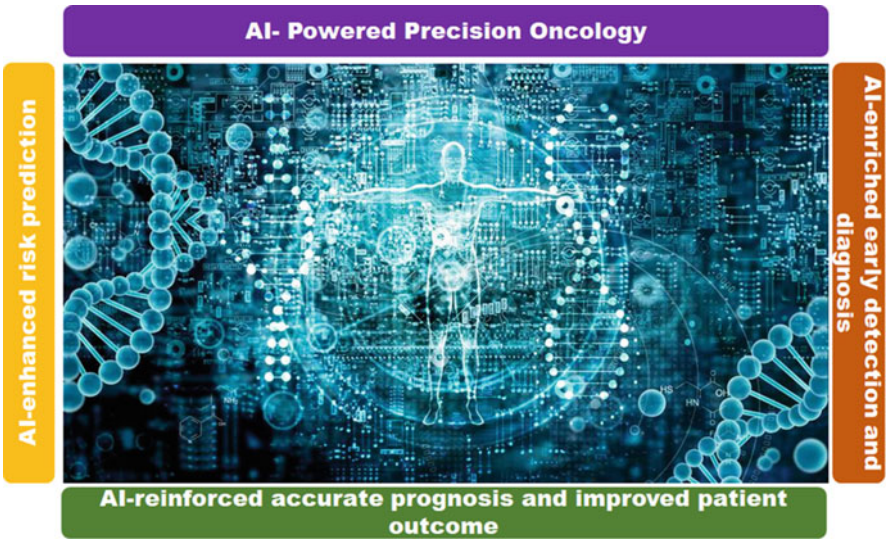


Fig. 15.3 AI-tools and precision oncology intersection and reinforcement. AI-empowers precision oncology through enhanced risk prediction, early detection and diagnosis, reinforced prognosis, and improved patient outcome (https://www.dreamstime.com/photos-images/biology.html)

negatively impact on patients and medical resources (Ozawa et al., 2020; Mori et al., 2020). The benefits and the pitfalls of AI in personalised cancer care remain to be accurately evaluated. Nonetheless, these benefits hold promising applications in oncology personalised care, breaking and creating unique inter- and intra-population barriers in advancing precision oncology and improving overall patient outcome, Fig. 15.4.

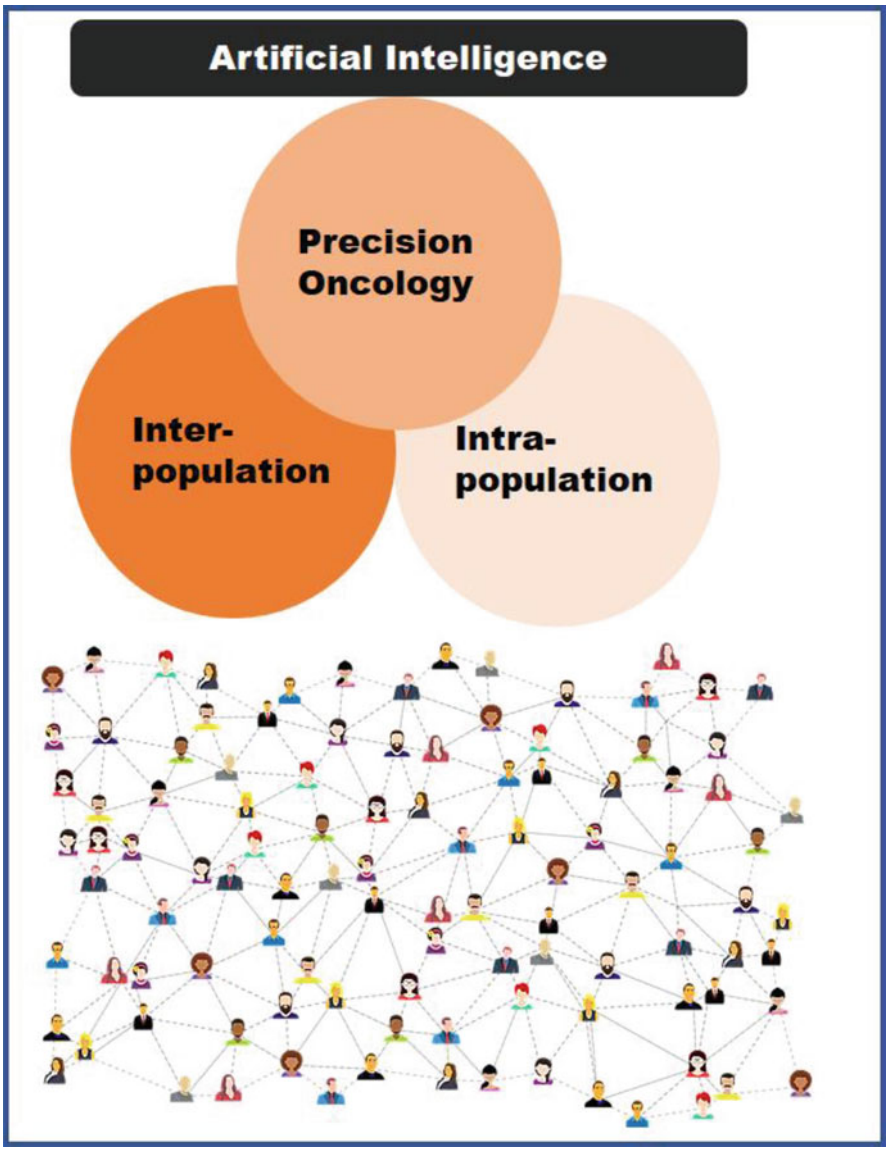


Fig. 15.4 AI-based tools hold great potential in breaking inter- and intra-population boundaries, while connecting global communities in advancing precision oncology (https://pixabay.com/images/search/biology/)

References

Acharya, U. R., Akter, A., Chowriappa, P., Dua, S., Raghavendra, U., Koh, J. E. W., Tan, J. H., Leong, S. S., Vijayananthan, A., Hagiwara, Y., Ramli, M. T., & Ng, K. H. (2018). Use of

nonlinear features for automated characterization of suspicious ovarian tumors using ultrasound images in fuzzy forest framework. *International Journal of Fuzzy Systems, 20*, 1385–1402.

Ahnen, D. J., Wade, S. W., Jones, W. F., Sifri, R., Mendoza Silveiras, J., Greenamyer, J., Guiffre, S., Axilbund, J., Spiegel, A., & You, Y. N. (2014). The increasing incidence of young-onset colorectal cancer: A call to action. *Mayo Clinic Proceedings, 89*, 216–224.

Akbari, M., Mohrekesh, M., Nasr-Esfahani, E., Soroushmehr, S. M. R., Karimi, N., Samavi, S., & Najarian, K. (2018). Polyp segmentation in colonoscopy images using fully convolutional network. *Annual International Conference of the IEEE Engineering in Medical Biology Society, 2018*, 69–72.

Aron, J. (2011). How innovative is Apple's new voice assistant, Siri? *New Scientist, 212*, 24.

Avanzo, M., Stancanello, J., Pirrone, G., & Sartor, G. (2020). Radiomics and deep learning in lung cancer. *Strahlentherapie und Onkologie, 196*, 879–887.

Ayling, R. M., Wong, A., & Cotter, F. (2021). Use of ColonFlag score for prioritisation of endoscopy in colorectal cancer. *BMJ Open Gastroenterology, 8*.

Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., Allison, T., Arnaout, O., Abbosh, C., Dunn, I. F., Mak, R. H., Tamimi, R. M., Tempany, C. M., Swanton, C., Hoffmann, U., Schwartz, L. H., Gillies, R. J., Huang, R. Y., & Aerts, H. (2019). Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA: A Cancer Journal for Clinicians, 69*, 127–157.

Bottaci, L., Drew, P. J., Hartley, J. E., Hadfield, M. B., Farouk, R., Lee, P. W., Macintyre, I. M., Duthie, G. S., & Monson, J. R. (1997). Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions. *Lancet, 350*, 469–472.

Brill, T. M., Munoz, L., & Miller, R. J. (2019). Siri, Alexa, and other digital assistants: A study of customer satisfaction with artificial intelligence applications. *Journal of Marketing Management, 35*, 1401–1436.

Buitrago, P. A., Nystrom, N. A., Gupta, R., & Saltz, J. (2020). Delivering scalable deep learning to research with Bridges-AI. In J. L. Crespo-Mariño & E. Meneses-Rojas (Eds.), *High performance computing* (pp. 200–214). Springer.

Bychkov, D., Linder, N., Turkki, R., Nordling, S., Kovanen, P. E., Verrill, C., Walliander, M., Lundin, M., Haglund, C., & Lundin, J. (2018). Deep learning based tissue analysis predicts outcome in colorectal cancer. *Scientific Reports, 8*, 3395.

Chhatwal, J., Alagoz, O., Lindstrom, M. J., Kahn, C. E., Shaffer, K. A., & Burnside, E. S. (2009). A logistic regression model based on the national mammography database format to aid breast cancer diagnosis. *American Journal of Roentgenology, 192*, 1117–1127.

Chua, I. S., Gaziel-Yablowitz, M., Korach, Z. T., Kehl, K. L., Levitan, N. A., Arriaga, Y. E., Jackson, G. P., Bates, D. W., & Hassett, M. (2021). Artificial intelligence in oncology: Path to implementation. *Cancer Medicine, 10*, 4138–4149.

Colón-Franco, J. M., Bossuyt, P. M. M., Algeciras-Schimnich, A., Bird, C., Engstrom-Melnyk, J., Fleisher, M., Kattan, M., & Lambert-Messerlian, G. (2018). Current and emerging multianalyte assays with algorithmic analyses – Are laboratories ready for clinical adoption? *Clinical Chemistry, 64*, 885–891.

Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: A comparison of three data mining methods. *Artificial Intelligence in Medicine, 34*, 113–127.

Dlamini, Z., Francies, F. Z., Hull, R., & Marima, R. (2020). Artificial intelligence (AI) and big data in cancer and precision oncology. *Computational and Structural Biotechnology Journal, 18*, 2300–2311.

Dong, J., Geng, Y., Lu, D., Li, B., Tian, L., Lin, D., & Zhang, Y. (2020). Clinical trials for artificial intelligence in cancer diagnosis: A cross-sectional study of registered trials in ClinicalTrials. gov. *Front Oncologia, 10***,** 1629.

Dreher, C., Linde, P., Boda-Heggemann, J., & Baessler, B. (2020). Radiomics for liver tumours. *Strahlentherapie und Onkologie, 196*, 888–899.

Farina, E., Nabhen, J. J., Dacoregio, M. I., Batalini, F., & Moraes, F. Y. (2022). An overview of artificial intelligence in oncology. *Future Science OA, 8*, FSO787.

Fernández-Esparrach, G., Bernal, J., López-Cerón, M., Córdova, H., Sánchez-Montes, C., Rodríguez de Miguel, C., & Sánchez, F. J. (2016). Exploring the clinical potential of an automatic colonic polyp detection method based on the creation of energy maps. *Endoscopy, 48*, 837–842.

Ghosh, A. (2019). Artificial intelligence using open source BI-RADS data exemplifying potential future use. *Journal of the American College of Radiology, 16*, 64–72.

Goksel Canbek, N., & Mutlu, M. E. (2016). On the track of artificial intelligence: Learning with intelligent personal assistants. *Journal of Human Sciences, 13*, 592–601.

Handelman, G. S., Kok, H. K., Chandra, R. V., Razavi, A. H., Huang, S., Brooks, M., Lee, M. J., & Asadi, H. (2019). Peering into the black box of artificial intelligence: Evaluation metrics of machine learning methods. *AJR American Journal of Roentgenology, 212*, 38–43.

Hart, G. R., Roffman, D. A., Decker, R., & Deng, J. (2018). A multi-parameterized artificial neural network for lung cancer risk prediction. *PLoS One, 13*, e0205264.

Hasnain, Z., Mason, J., Gill, K., Miranda, G., Gill, I. S., Kuhn, P., & Newton, P. K. (2019). Machine learning models for predicting post-cystectomy recurrence and survival in bladder cancer patients. *PLoS One, 14*, e0210976.

Haubold, J., Hosch, R., Umutlu, L., Wetter, A., Haubold, P., Radbruch, A., Forsting, M., Nensa, F., & Koitka, S. (2021). Contrast agent dose reduction in computed tomography with deep learning using a conditional generative adversarial network. *European Radiology, 31*, 6087–6095.

Hilsden, R. J., Heitman, S. J., Mizrahi, B., Narod, S. A., & Goshen, R. (2018). Prediction of findings at screening colonoscopy using a machine learning algorithm based on complete blood counts (ColonFlag). *PLoS One, 13*, e0207848.

Holst, C., Sukums, F., Radovanovic, D., Ngowi, B., Noll, J., & Winkler, A. S. (2020). Sub-Saharan Africa-the new breeding ground for global digital health. *Lancet Digit Health, 2*, e160–e162.

Hornbrook, M. C., Goshen, R., Choman, E., O'Keeffe-Rosetti, M., Kinar, Y., Liles, E. G., & Rust, K. C. (2017). Early colorectal cancer detected by machine learning model using gender, age, and complete blood count data. *Digestive Diseases and Sciences, 62*, 2719–2727.

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. (2018). Artificial intelligence in radiology. *Nature Reviews. Cancer, 18*, 500–510.

Hu, L., Bell, D., Antani, S., Xue, Z., Yu, K., Horning, M. P., Gachuhi, N., Wilson, B., Jaiswal, M. S., Befano, B., Long, L. R., Herrero, R., Einstein, M. H., Burk, R. D., Demarco, M., Gage, J. C., Rodriguez, A. C., Wentzensen, N., & Schiffman, M. (2019). An observational study of deep learning and automated evaluation of cervical images for cancer screening. *Journal of the National Cancer Institute, 111*, 923–932.

Huang, S., Yang, J., Fong, S., & Zhao, Q. (2020). Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. *Cancer Letters, 471*, 61–71.

Hunter, B., Hindocha, S., & Lee, R. W. (2022). The role of artificial intelligence in early cancer diagnosis. *Cancers (Basel), 14*.

Iqbal, M. J., Javed, Z., Sadia, H., Qureshi, I. A., Irshad, A., Ahmed, R., Malik, K., Raza, S., Abbas, A., Pezzani, R., & Sharifi-Rad, J. (2021). Clinical applications of artificial intelligence and machine learning in cancer diagnosis: Looking into the future. *Cancer Cell International, 21*, 270.

Issa, I. A., & Noureddine, M. (2017). Colorectal cancer screening: An updated review of the available options. *World Journal of Gastroenterology, 23*, 5086–5096.

Ivancic, M. M., Megna, B. W., Sverchkov, Y., Craven, M., Reichelderfer, M., Pickhardt, P. J., Sussman, M. R., & Kennedy, G. D. (2020). Noninvasive detection of colorectal carcinomas using serum protein biomarkers. *The Journal of Surgical Research, 246*, 160–169.

Jha, S., & Topol, E. J. (2016). Adapting to artificial intelligence: Radiologists and pathologists as information specialists. *JAMA, 316*, 2353–2354.

Johnson, K. B., Wei, W. Q., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., Zhao, J., & Snowdon, J. L. (2021). Precision medicine, AI, and the future of personalized health care. *Clinical and Translational Science, 14*, 86–93.

Katsari, K., Penna, D., Arena, V., Polverari, G., Ianniello, A., Italiano, D., Milani, R., Roncacci, A., Illing, R. O., & Pelosi, E. (2021). Artificial intelligence for reduced dose 18F-FDG PET examinations: A real-world deployment through a standardized framework and business case assessment. *EJNMMI Physics, 8*, 25.

Kayser, M., Soberanis-Mukul, R. D., Zvereva, A.-M., Klare, P., Navab, N., & Albarqouni, S. (2020). Understanding the effects of artifacts on automated polyp detection and incorporating that knowledge via learning without forgetting. *arXiv preprint arXiv:2002.02883*.

Kepuska, V., & Bohouta, G. (2018). Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, 99–103.

Kinar, Y., Akiva, P., Choman, E., Kariv, R., Shalev, V., Levin, B., Narod, S. A., & Goshen, R. (2017). Performance analysis of a machine learning flagging system used to identify a group of individuals at a high risk for colorectal cancer. *PLoS One, 12*, e0171759.

Kinar, Y., Kalkstein, N., Akiva, P., Levin, B., Half, E. E., Goldshtein, I., Chodick, G., & Shalev, V. (2016). Development and validation of a predictive model for detection of colorectal cancer in primary care by analysis of complete blood counts: A binational retrospective study. *Journal of the American Medical Informatics Association, 23*, 879–890.

Kochanny, S. E., & Pearson, A. T. (2021). Academics as leaders in the cancer artificial intelligence revolution. *Cancer, 127*, 664–671.

Kocher, M., Ruge, M. I., Galldiks, N., & Lohmann, P. (2020). Applications of radiomics and machine learning for radiotherapy of malignant brain tumors. *Strahlentherapie und Onkologie, 196*, 856–867.

Kuo, R. J., Huang, M. H., Cheng, W. C., Lin, C. C., & Wu, Y. H. (2015). Application of a two-stage fuzzy neural network to a prostate cancer prognosis system. *Artificial Intelligence in Medicine, 63*, 119–133.

Larson, J. A., Johnson, M. H., & Bhayani, S. B. (2014). Application of surgical safety standards to robotic surgery: Five principles of ethics for nonmaleficence. *Journal of the American College of Surgeons, 218*, 290–293.

Lee, K.-S., Son, S.-H., Park, S.-H., & Kim, E. S. (2021). Automated detection of colorectal tumors based on artificial intelligence. *BMC Medical Informatics and Decision Making, 21*, 33.

Leibovici, L., Paul, M., Nielsen, A. D., Tacconelli, E., & Andreassen, S. (2007). The TREAT project: Decision support and prediction using causal probabilistic networks. *International Journal of Antimicrobial Agents, 30*(Suppl 1), S93–S102.

Li Wen, Y., & Leech, M. (2020). Review of the role of radiomics in tumour risk classification and prognosis of cancer. *Anticancer Research, 40*, 3605–3618.

Liew, X. Y., Hameed, N., & Clos, J. (2021). An investigation of XGBoost-based algorithm for breast cancer classification. *Machine Learning with Applications, 6*, 100154.

Lu, H., Arshad, M., Thornton, A., Avesani, G., Cunnea, P., Curry, E., Kanavati, F., Liang, J., Nixon, K., Williams, S. T., Hassan, M. A., Bowtell, D. D. L., Gabra, H., Fotopoulou, C., Rockall, A., & Aboagye, E. O. (2019). A mathematical-descriptor of tumor-mesoscopic-structure from computed-tomography images annotates prognostic- and molecular-phenotypes of epithelial ovarian cancer. *Nature communications* [Online], 10. https://doi.org/10.1038/s41467-019-08718-9 [Accessed 2019/02//].

Lu, T. P., Kuo, K. T., Chen, C. H., Chang, M. C., Lin, H. P., Hu, Y. H., Chiang, Y. C., Cheng, W. F., & Chen, C. A. (2019). Developing a prognostic gene panel of epithelial ovarian cancer patients by a machine learning model. *Cancers (Basel), 11*.

Lu, H., Wang, H., Yoon, S. W., Won, D., & Park, S. (2019). Dynamic predictive modeling of solder paste volume with real time memory update in a stencil printing process. *Procedia Manufacturing, 38*, 108–116.

Lynch, C. M., Abdollahi, B., Fuqua, J. D., De Carlo, A. R., Bartholomai, J. A., Balgemann, R. N., van Berkel, V. H., & Frieboes, H. B. (2017). Prediction of lung cancer patient survival via supervised machine learning classification techniques. *International Journal of Medical Informatics, 108*, 1–8.

Maspero, M., Bentvelzen, L. G., Savenije, M. H. F., Guerreiro, F., Seravalli, E., Janssens, G. O., van den Berg, C. A. T., & Philippens, M. E. P. (2020). Deep learning-based synthetic CT generation for paediatric brain MR-only photon and proton radiotherapy. *Radiotherapy and Oncology, 153*, 197–204.

Mccartney, M. (2018). Margaret McCartney: AI in medicine must be rigorously tested. *BMJ, 361*, k1752.

Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep patient: An unsupervised representation to predict the future of patients from the electronic health records. *Scientific Reports, 6*, 26094.

Mori, Y., Kudo, S. E., East, J. E., Rastogi, A., Bretthauer, M., Misawa, M., Sekiguchi, M., Matsuda, T., Saito, Y., Ikematsu, H., Hotta, K., Ohtsuka, K., Kudo, T., & Mori, K. (2020). Cost savings in colonoscopy with artificial intelligence-aided polyp diagnosis: An add-on analysis of a clinical trial (with video). *Gastrointestinal Endoscopy, 92*, 905–911.e1.

Muhammad, W., Hart, G. R., Nartowt, B., Farrell, J. J., Johung, K., Liang, Y., & Deng, J. (2019). Pancreatic cancer prediction through an artificial neural network. *Frontiers Artificial Intelligence, 2*, 2.

Murakami, Y., Magome, T., Matsumoto, K., Sato, T., Yoshioka, Y., & Oguchi, M. (2020). Fully automated dose prediction using generative adversarial networks in prostate cancer patients. *PLoS One, 15*, e0232697.

Nartowt, B. J., Hart, G. R., Muhammad, W., Liang, Y., Stark, G. F., & Deng, J. (2020). Robust machine learning for colorectal cancer risk prediction and stratification. *Frontiers in Big Data, 3*, 6.

Nartowt, B. J., Hart, G. R., Roffman, D. A., Llor, X., Ali, I., Muhammad, W., Liang, Y., & Deng, J. (2019). Scoring colorectal cancer risk with an artificial neural network based on self-reportable personal health data. *PLoS One, 14*, e0221421.

Nazari, E., Aghemiri, M., Avan, A., Mehrabian, A., & Tabesh, H. (2021). Machine learning approaches for classification of colorectal cancer with and without feature selection method on microarray data. *Gene Reports, 25*, 101419.

Olatunji, S. O., Alotaibi, S., Almutairi, E., Alrabae, Z., Almajid, Y., Altabee, R., Altassan, M., Basheer Ahmed, M. I., Farooqui, M., & Alhiyafi, J. (2021). Early diagnosis of thyroid cancer diseases using computational intelligence techniques: A case study of a Saudi Arabian dataset. *Computers in Biology and Medicine, 131*, 104267.

Ozawa, T., Ishihara, S., Fujishiro, M., Kumagai, Y., Shichijo, S., & Tada, T. (2020). Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks. *Therapeutic Advances in Gastroenterology, 13*, 1756284820910659.

Pan, Y., Zhang, L., Zhang, R., Han, J., Qin, W., Gu, Y., Sha, J., Xu, X., Feng, Y., Ren, Z., Dai, J., Huang, B., Ren, S., & Gu, J. (2021). Screening and diagnosis of colorectal cancer and advanced adenoma by Bionic Glycome method and machine learning. *American Journal of Cancer Research, 11*, 3002–3020.

Park, K., Ali, A., Kim, D., An, Y., Kim, M., & Shin, H. (2013). Robust predictive model for evaluating breast cancer survivability. *Engineering Applications of Artificial Intelligence, 26*, 2194–2205.

Parker, R. K., Mwachiro, M. M., Ranketi, S. S., Mogambi, F. C., Topazian, H. M., & White, R. E. (2020). Curative surgery improves survival for colorectal cancer in rural Kenya. *World Journal of Surgery, 44*, 30–36.

Patel, S. K., George, B., & Rai, V. (2020). Artificial intelligence to decode cancer mechanism: Beyond patient stratification for precision oncology. *Frontiers in Pharmacology, 11*, 1177.

Qadir, H. A., Balasingham, I., Solhusvik, J., Bergsland, J., Aabakken, L., & Shin, Y. (2020). Improving automatic polyp detection using CNN by exploiting temporal dependency in colonoscopy video. *IEEE Journal of Biomedical and Health Informatics, 24*, 180–193.

Qiu, H., Ding, S., Liu, J., Wang, L., & Wang, X. (2022). Applications of artificial intelligence in screening, diagnosis, treatment, and prognosis of colorectal cancer. *Current Oncology, 29*, 1773–1795.

Ranschaert, E. R., Morozov, S. P., & Algra, P. R. (2019). *Artificial intelligence in medical imaging: Opportunities, applications and risks*. Springer.

Roffman, D., Hart, G., Girardi, M., Ko, C. J., & Deng, J. (2018). Predicting non-melanoma skin cancer via a multi-parameterized artificial neural network. *Scientific Reports, 8*, 1701.

Sakellaropoulos, T., Vougas, K., Narang, S., Koinis, F., Kotsinas, A., Polyzos, A., Moss, T. J., Piha-Paul, S., Zhou, H., Kardala, E., Damianidou, E., Alexopoulos, L. G., Aifantis, I., Townsend, P. A., Panayiotidis, M. I., Sfikakis, P., Bartek, J., Fitzgerald, R. C., Thanos, D., Mills Shaw, K. R., Petty, R., Tsirigos, A., & Gorgoulis, V. G. (2019). A deep learning framework for predicting response to therapy in cancer. *Cell Reports, 29*, 3367–3373.e4.

Schaffter, T., Buist, D. S. M., Lee, C. I., Nikulin, Y., Ribli, D., Guan, Y., Lotter, W., Jie, Z., Du, H., Wang, S., Feng, J., Feng, M., Kim, H. E., Albiol, F., Albiol, A., Morrell, S., Wojna, Z., Ahsen, M. E., Asif, U., Jimeno Yepes, A., Yohanandan, S., Rabinovici-Cohen, S., Yi, D., Hoff, B., Yu, T., Chaibub Neto, E., Rubin, D. L., Lindholm, P., Margolies, L. R., Mcbride, R. B., Rothstein, J. H., Sieh, W., Ben-Ari, R., Harrer, S., Trister, A., Friend, S., Norman, T., Sahiner, B., Strand, F., Guinney, J., Stolovitzky, G., Mackey, L., Cahoon, J., Shen, L., Sohn, J. H., Trivedi, H., Shen, Y., Buturovic, L., Pereira, J. C., Cardoso, J. S., Castro, E., Kalleberg, K. T., Pelka, O., Nedjar, I., Geras, K. J., Nensa, F., Goan, E., Koitka, S., Caballero, L., Cox, D. D., Krishnaswamy, P., Pandey, G., Friedrich, C. M., Perrin, D., Fookes, C., Shi, B., Cardoso Negrie, G., Kawczynski, M., Cho, K., Khoo, C. S., Lo, J. Y., Sorensen, A. G., & Jung, H. (2020). Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. *JAMA Network Open, 3*, e200265.

Schneider, J. L., Layefsky, E., Udaltsova, N., Levin, T. R., & Corley, D. A. (2020). Validation of an algorithm to identify patients at risk for colorectal cancer based on laboratory test and demographic data in diverse, community-based population. *Clinical Gastroenterology and Hepatology, 18*, 2734–2741.e6.

Schultheiss, M., Schober, S. A., Lodde, M., Bodden, J., Aichele, J., Müller-Leisse, C., Renger, B., Pfeiffer, F., & Pfeiffer, D. (2020). A robust convolutional neural network for lung nodule detection in the presence of foreign bodies. *Scientific Reports, 10*, 12987.

Sepehri, S., Upadhaya, T., Desseroit, M.-C., Visvikis, D., Cheze le Rest, C., & Hatt, M. (2018). Comparison of machine learning algorithms for building prognostic models in non-small cell lung cancer using clinical and radiomics features from 18F-FDG PET/CT images. *Journal of Nuclear Medicine, 59*, 328.

Shaffer, K. (2018). Can machine learning be used to generate a model to improve management of high-risk breast lesions? *Radiology, 286*, 819–821.

Shaver, M. M., Kohanteb, P. A., Chiou, C., Bardis, M. D., Chantaduly, C., Bota, D., Filippi, C. G., Weinberg, B., Grinband, J., Chow, D. S., & Chang, P. D. (2019). Optimizing neuro-oncology imaging: A review of deep learning approaches for glioma imaging. *Cancers (Basel), 11*.

Stark, G. F., Hart, G. R., Nartowt, B. J., & Deng, J. (2019). Predicting breast cancer risk using personal health data and machine learning models. *PLoS One, 14*, e0226765.

Suh, Y. J., Jung, J., & Cho, B. J. (2020). Automated breast cancer detection in digital mammograms of various densities via deep learning. *Journal of Personalized Medicine, 10*.

Sun, D., Wang, M., & Li, A. (2018). A multimodal deep neural network for human breast cancer prognosis prediction by integrating multi-dimensional data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*.

Takemura, Y., Yoshida, S., Tanaka, S., Kawase, R., Onji, K., Oka, S., Tamaki, T., Raytchev, B., Kaneda, K., Yoshihara, M., & Chayama, K. (2012). Computer-aided system for predicting the histology of colorectal tumors by using narrow-band imaging magnifying colonoscopy (with video). *Gastrointestinal Endoscopy, 75*, 179–185.

Tian, Y., Pu, L. Z., Singh, R., Burt, A. D., & Carneiro, G. (2019). One-stage five-class polyp detection and classification. *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019). IEEE*, 70–73.

Tian, D. W., Wu, Z. L., Jiang, L. M., Gao, J., Wu, C. L., & Hu, H. L. (2019). Neural precursor cell expressed, developmentally downregulated 8 promotes tumor progression and predicts poor prognosis of patients with bladder cancer. *Cancer Science, 110*, 458–467.

Urban, G., Tripathi, P., Alkayali, T., Mittal, M., Jalali, F., Karnes, W., & Baldi, P. (2018). Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology, 155*, 1069–1078.e8.

van Rijn, J. C., Reitsma, J. B., Stoker, J., Bossuyt, P. M., van Deventer, S. J., & Dekker, E. (2006). Polyp miss rate determined by tandem colonoscopy: A systematic review. *The American Journal of Gastroenterology, 101*, 343–350.

Vasudevan, P., & Murugesan, T. (2018). Cancer subtype discovery using prognosis-enhanced neural network classifier in multigenomic data. *Technology in Cancer Research & Treatment, 17*, 1533033818790509.

Verma, V., Sprave, T., Haque, W., Simone, C. B., 2nd, Chang, J. Y., Welsh, J. W., & Thomas, C. R., Jr. (2018). A systematic review of the cost and cost-effectiveness studies of immune checkpoint inhibitors. *Journal for Immunotherapy of Cancer, 6*, 128.

Waljee, A. K., Weinheimer-Haus, E. M., Abubakar, A., Ngugi, A. K., Siwo, G. H., Kwakye, G., Singal, A. G., Rao, A., Saini, S. D., Read, A. J., Baker, J. A., Balis, U., Opio, C. K., Zhu, J., & Saleh, M. N. (2022). Artificial intelligence and machine learning for early detection and diagnosis of colorectal cancer in sub-Saharan Africa. *Gut, 71*, 1259–1265.

Wang, F., Casalino, L. P., & Khullar, D. (2019). Deep learning in medicine-promise, progress, and challenges. *JAMA Internal Medicine, 179*, 293–294.

Wang, Y., Wang, D., Ye, X., Wang, Y., Yin, Y., & Jin, Y. (2019). A tree ensemble-based two-stage model for advanced-stage colorectal cancer survival prediction. *Information Sciences, 474*, 106–124.

Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., Ossorio, P. N., Thadaney-Israni, S., & Goldenberg, A. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine, 25*, 1337–1340.

Xiao, L. H., Chen, P. R., Gou, Z. P., Li, Y. Z., Li, M., Xiang, L. C., & Feng, P. (2017). Prostate cancer prediction using the random forest algorithm that takes into account transrectal ultrasound findings, age, and serum levels of prostate-specific antigen. *Asian Journal of Andrology, 19*, 586–590.

Yamada, M., Saito, Y., Imaoka, H., Saiko, M., Yamada, S., Kondo, H., Takamaru, H., Sakamoto, T., Sese, J., Kuchiba, A., Shibata, T., & Hamamoto, R. (2019). Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Scientific Reports, 9*, 14465.

Yu, K. H., Zhang, C., Berry, G. J., Altman, R. B., Ré, C., Rubin, D. L., & Snyder, M. (2016). Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nature Communications, 7*, 12474.

Yujuan, J., Xiangyang, L., & Binlai, A. (2020). AI based supercomputer: Opportunities and challenges. In Q. Yu (Ed.), *Space information networks* (pp. 47–55). Springer.

Zhang, F., Kaufman, H. L., Deng, Y., & Drabier, R. (2013). Recursive SVM biomarker selection for early detection of breast cancer in peripheral blood. *BMC Medical Genomics, 6*(Suppl 1), S4.

Zhang, S., Xu, Y., Hui, X., Yang, F., Hu, Y., Shao, J., Liang, H., & Wang, Y. (2017). Improvement in prediction of prostate cancer prognosis with somatic mutational signatures. *Journal of Cancer, 8*, 3261–3267.

Zhang, R., Zheng, Y., Mak, T. W., Yu, R., Wong, S. H., Lau, J. Y., & Poon, C. C. (2017). Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain. *IEEE Journal of Biomedical and Health Informatics, 21*, 41–47.

Zhu, W., Xie, L., Han, J., & Guo, X. (2020). The application of deep learning in cancer prognosis prediction. *Cancers (Basel), 12*.

Chapter 16
Conclusion and Insights into the Future of AI in Precision Oncology

Zodwa Dlamini, Demetra Demetriou, and Thulo Molefi

Abstract Precision oncology is the main goal in the medical field. Personalized patient- and health care will decrease the burden of cancer and improve diagnosis, prognosis, treatment, and survival. Artificial intelligence (AI) is fast changing how medical research is done with resultant improvements in patient care. By analyzing large data sets, AI allows for the prediction of disease risk, early detection of disease progression, and treatment outcomes for an individual patient on a particular therapy. AI is also used as a novel approach for exploring ccfNAs that can be used as potential drug targets or biomarkers. AI has led to improved data collection, analysis, and interpretation of microbiomics, epigenomics, and metabolomics. AI has improved the capabilities of medical imaging by enhancing digital pathology and giving rise to the new field of radiogenomics. AI is being used to assist in the development and application of nanomedicine to precision oncology. AI has improved drug design to be more successful, less expensive, more target specific and create drugs with reduced toxicity. The application of AI to the field of medical devices has led to the introduction of sophisticated AI-based medical devices. Additionally, AI will be used to solve many yet unforeseen challenges faced daily in the medical field. Precision oncology through the application of AI to medicine allows the best possible care for patients as it is personalized to their specific needs.

Z. Dlamini (✉) · D. Demetriou
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa
e-mail: zodwa.dlamini@up.ac.za; dd.demetriou@up.ac.za

T. Molefi
SAMRC Precision Oncology Research Unit (PORU), DSI/NRF SARChI Chair in Precision Oncology and Cancer Prevention (POCP), Pan African Cancer Research Institute (PACRI), University of Pretoria, Pretoria, South Africa

Department of Medical Oncology, Steve Biko Academic Hospital and University of Pretoria, Pretoria, South Africa
e-mail: thulo.molefi@up.ac.za

Z. Dlamini (ed.), *Artificial Intelligence and Precision Oncology*,
https://doi.org/10.1007/978-3-031-21506-3_16

Keywords Artificial Intelligence · Precision oncology · Biomarkers · ccfNAs · Omics · Nanomedicine · Drug design · AI-medical devices · Patient care · Treatment

16.1 Conclusion

Improved understanding of the pathogenesis of cancer, the tumor microenvironment, and metastatic pathways has led to a more precise approach to how malignancies are managed. The aim is to provide treatment that is accurate and less toxic. AI has allowed this aim to be reached. AI is also used as a novel approach for exploring ccfNAs in personalized clinical diagnosis and prognosis. In addition, the information obtained from large omics data sets has led to improved screening and monitoring of cancer in the form of radiogenomics. Epigenomics reflects how an individual's environment and lifestyle can alter their cancer risk. It also provides data concerning the predisposition to develop cancer at premalignant stages in the cancer pathogenesis sequence. Epigenomics is considered big data and its analysis and interpretation can only be realistically undertaken using AI and machine learning. Metabolomics, on the other hand, is defined as the comprehensive analysis of metabolites in a biological specimen and holds a long-awaited promise to inform the practice of precision medicine. Metabolites are also used to diagnose complex metabolic diseases, such as inborn errors of metabolism. The human body is also inhabited by a vast number of microorganisms that are known as the microbiome. The microbiome offers substantially more genetic diversity, and hence more flexibility, than the human genome. Given the diversity of both the microbiome and cancer, together with recent progress in multi-omics studies, it is inevitable that machine learning and AI algorithms will be incorporated as essential tools required for the accurate interpretation of big data and used to enhance decision-making systems and advance precision oncology. AI can also be applied to new and established technologies in the medical fields to improve their application and development (Fig. 16.1).

Nanomedicine has offered innovative solutions for some of the world's most pressing problems, particularly in the health space. AI applications have also been applied to nanotechnology in the form of nanomedicine. AI-based algorithms for nanomedicine hold promise for improving precision care. AI-based medical devices promise to revolutionize medical imaging tools and lead to advancement in the diagnosis and management of cancer patients. The development of new drugs is expensive, time-consuming, and often results in failure. These problems can partially be solved using AI to identify drug targets, search for ligands for these targets,

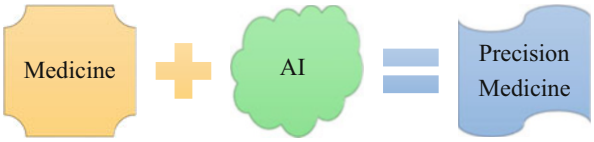


Fig. 16.1 Medicine combined with Artificial Intelligence will lead to precision medicine and the next-generation medicine with improved healthcare

and then model the interactions of the drug and its target while determining the physiochemical properties of the drug. Cancer immunotherapy refers to the manipulation of the patient's immune system to fight cancer. AI can monitor the use of cancer immunotherapy, predict patient tolerance, and optimize treatment response. The experience-based, one-size-fits-all approach of medicine is being replaced by more precise approaches, which can be individualized with more accuracy. Disruptive technologies, such as genome sequencing and advanced biotechnology, generate vast amounts of data. Data which is so vast, that it would be impossible for the human mind to keep up with and remember in totality, thus requiring AI. The introduction of sophisticated AI-medical devices demonstrates the fundamental role that AI holds to offer in oncology. Several AI-tools have illustrated high performance toward cancer care and management in various parts of the world.

The integration and analysis of data from various sources, such as "multi-omics" data, medical images, medical imaging reports, electronic and hand-written medical records, is only possible in a practical manner using AI techniques. AI-based tools can enhance human potential by identifying changes in disease patterns. These tools will identify variances in patient data and help to identify predictors of unforeseen patients' response to therapy and enable personalized patient care. The need for personalized care is needed more in oncology than any other field in medicine. AI can be used for the screening of common malignancies in high-risk individuals and allow for early and accurate diagnosis while assisting to rule out other benign differential diagnoses. The development of AI-based technologies for biomedical applications opened a new era in the field of personalized clinical diagnosis and prognosis. Combining digital pathology, radiogenomics, and AI can improve workflow and lead to advanced diagnostics. AI-aided screening methods are more likely to detect premalignant features within colon polyps which may otherwise have been missed by humans. Epigenomic data can be used on its own or integrated with other omics data to provide a clear view of the molecular landscape within signaling pathways involved in the development and progression of cancer. Integrating metabolomics technologies with AI indeed offers more effective cancer care strategies, as it aids in overcoming challenges that either technique cannot overcome on its own. Advances in nanomedicine, coupled with novel computational methods of intelligent analysis, will be a game-changer in the development of healthcare and precision medicine. The ability of these devices to use machine learning to improve their abilities makes them autonomous. Therefore, they do not require external upgrades and as they learn they will in future be able to operate independently without human manipulation. AI can be used to help solve these challenges and lead to shorter development times for drugs. AI has allowed for the stratification of patients into responders and non-responders so that only patients who will benefit from cancer immunotherapy are treated accordingly. The strategic application of a patient-centered AI-Pathway companion in disease prevention, diagnosis, and treatment has the potential to revolutionize the healthcare industry as we know it.