Journal of Advanced Health Informatics Research (JAHIR)
Vol. 1, No. 3, December 2023, pp. 128-141
ISSN: 2985-6124, DOI: 10.59247/jahir.v1i3.167

Challenges and Opportunities: Integration of Data Science in Cancer Research Through A Literature Review Approach

Purwono Purwono [1], Ariefah Khairina Islahati [2], Yuslena Sari [3],
Dewi Astria Faroek [4], Muhammad Baballe Ahmad [5]
[1,2] *Universitas Harapan Bangsa, Purwokerto, Indonesia*
[3] *Universitas Lambung Mangkurat, Banjarmasin, Indonesia*
[4] *Universitas Muhammadiyah Sorong, Sorong, Indonesia*
[5] *Nigerian Defence Academy, Kaduna, Nigeria*

ARTICLE INFO

Article history:

Received December 28, 2024
Revised January 21, 2024
Published January 22, 2024

Keywords:

Cancer;
Cancer Treatment;
Outcomes;
Artificial Intelligence;
Machine Learning;
Challenges

ABSTRACT

Several research articles in this journal relate to various aspects of cancer, such as treatment, patient outcomes, caregiver responsibilities, and the use of AI and liquid biopsy in cancer research. Covers a wide range of topics, including valuable insights into the latest developments in cancer research as well as potential future opportunities and issues. Several articles discuss the impact of non-coding RNA on gastric cancer, machine learning decision support systems for cancer survival factors, economic impact of cancer mortality, nausea in children diagnosed with cancer, protein-RNA variations in cancer clinical analysis, integration and proteomic data analysis in the context of cancer genomics, personalized cancer medicine, mass spectrometry-based clinical proteomics, cancer proteogenomics, subtype-based This journal provides an in-depth overview of various aspects of current cancer research and future research prospects.

Corresponding Author:

Purwono Purwono, Universitas Harapan Bangsa, Purwokerto, Indonesia
Email: purwono@uhb.ac.id

1. INTRODUCTION

Most research investment is still focused on basic science and clinical translational research, despite strong evidence showing that prevention is the primary defense against cancer. Assessment of patient experience has historically relied on qualitative research designs and standardized patient outcome instruments; both are susceptible to retrospective questioning and retrospective recall bias. Big data and AI will change our understanding of cancer from onset, screening, diagnosis, treatment, response, toxicity, recurrence and survival. Cancer research, one of the most complex fields of health science, has great potential to benefit from the application of Data Science in the exploration of the health field in today's information age. Through a thorough literature review, the journal aims to explore the challenges and prospects emerging from the application of data science in cancer research. By understanding the current research context, we can find ways to maximize the use of data science in cancer treatment[1][2][3].

Cancer research has undergone significant transformation over time, with technological advances and increasingly massive data collection. However, the extent to which Data Science has made a substantial contribution to this progress needs to be understood further. In this write-up, we will outline in detail critical challenges, such as data complexity and the need for better analytical models, as well as detail the enormous potential offered by Data Science integration, such as personalization of treatment and improved early diagnosis. The organization responsible for developing OMOP, Observational Health Data Sciences and Informatics (OHDSI), concentrates on the horizontal sharing of data sets that have similar characteristics[4].

The term "Big Data" refers to very large and complex data sets that are difficult to process traditionally. Some characteristics of Big Data include its large volume, speed of data collection, and diversity of formats. The use of big data in cancer research can aid in better research and therapy development decisions, discover important patterns, and provide deeper insights.

The characteristics of big data require a powerful and innovative parallel approach to extract critical information from large datasets in a scalable manner by leveraging the computational resources of distributed memory-based systems. Current standard analyzes show shortcomings in measuring the future economic costs and benefits of big data, primarily due to a lack of clarity about the units of measurement used to measure big data. This creates challenges and opportunities for optimizing the use of big data in cancer research, where a deep understanding of the characteristics and economic potential of big data is critical to addressing the problems facing medicine[5][6].

Basic cancer research investigates questions about the origins and development of cancer, producing useful theoretical knowledge without immediate results. These advances have left the cancer survivor population facing significant physical, social, and psychological problems in addition to the suffering associated with most cancer treatments[1][7].

This research uses the Literature review research method by collecting information obtained from previous research with search keywords regarding Understanding cancer, BioInformatics, Machine Learning, Integration Opportunities, Data Science, Data Science Integration, Cancer Research. Literature data searches were obtained using search engines via Science Direct, IEEE, and SpringerLink within the last 5 years.

Of the 230 papers found through literature searches regarding Understanding cancer, BioInformatics, Machine Learning, Integration Opportunities, Data Science, Data Science Integration, Cancer Research, researchers filtered them using certain criteria. The selection process is carried out carefully, considering the relevance and quality of each article. These criteria include aspects such as the Machine Learning method approach, the deep learning used, the accuracy of research results, as well as the unique contribution provided by each selected article. As a result, a number of selected references that best suited the objectives of this research were selected for further analysis.

2. UNDERSTANDING CANCER

Millions of people around the world have been affected by cancer, a deadly disease. Various types of cancer can affect human organs. some of the most common types of cancer and provides relevant information.

a. Breast Cancer: One of the most common types of cancer among women is breast cancer. If found at an early stage, this cancer can develop in breast tissue and has a higher cure rate. According to a study published by the American Cancer Society, therefore, breast cancer treatment is increasingly personalized by considering certain factors, such as clinical stage, biological characteristics of the tumor, including tumor subtype, and subtypes within it [8].
b. Lung Cancer: One of the most common causes of death due to cancer in the world is lung cancer. The World Health Organization (WHO) says that smoking and exposure to harmful substances such as cigarette smoke and air pollution increases the risk of developing lung cancer. However, references show that stopping smoking and reducing exposure to harmful substances can help reduce the risk of developing lung cancer. Data science and big data have transformed oncology, especially in the treatment of lung cancer, focusing on technical staging, better imaging, and the use of molecular markers to improve patient management [9].
c. Prostate Cancer: A study from the National Cancer Institute shows that things like age, family history, and lifestyle can influence the risk of developing prostate cancer. In the treatment of prostate cancer, early detection and appropriate treatment are essential. Men of African descent have the highest incidence and mortality rates. There is increasing understanding about the male gender which is more susceptible to various types of prostate cancer. Research into new treatment approaches for localized prostate cancer is limited. According to the latest WHO tumor classification, prostate cancer (PCa)

can be divided into several categories based on histological type, subtype, and growth pattern[10][11][12].

d. Colorectal Cancer: According to the American Cancer Society, colorectal cancer, also known as colon cancer, is a type of cancer that affects the intestines and rectum. Regular check-ups and following a diet rich in fiber can help reduce the risk of developing colorectal cancer. While downregulation of ANP32B inhibits proliferation and migration of colorectal cancer cells, upregulation of ANP32B increases proliferation and migration of colorectal cancer cells. Cost-of-illness (COI) studies aid the allocation of limited resources in healthcare and provide important information for decision making about therapy options[13][14].
e. Cervical Cancer: Cervical cancer is now the fourth most common malignancy in women throughout the world, and is a serious cancer. To improve the prognosis of cervical cancer patients, a thorough understanding of the tumor ecosystem is essential because the heterogeneity of the complex tumor ecosystem influences tumorigenesis, malignancy progression, and response to treatment. The use of single cell sequencing has grown rapidly and become more common. This has created a new paradigm in cancer research and provided a deeper and more comprehensive understanding of cancer[15].

It is hoped that a better understanding of cancer can increase awareness of the importance of preventing and detecting cancer early. To provide the latest information about types of cancer and how to prevent them, international health organizations such as WHO and the American Cancer Society are very helpful.

3. ARTIFICIAL INTELLIGENCE IN CANCER

There is the potential to change the paradigm for treating this disease. Artificial intelligence can help in various aspects of cancer research by exploiting the ability to process and analyze very complex data. Artificial intelligence plays an important role in diagnosis. Using diverse datasets, machine teaching algorithms such as Random Forests or Support Vector Machines (SVM) can be trained to identify patterns and characteristics of cancer in medical data. This allows for faster and more accurate diagnosis.

In addition, Artificial intelligence algorithms are involved in the analysis of omics data, such as genomics and proteomics. This helps understand the molecular and genetic aspects of cancer, supporting the development of appropriate medicines. Artificial intelligence algorithms are expected to help cancer research gain a better understanding of the complexity of the disease. Additionally, Artificial intelligence can enable the development of better therapies and personalization of treatments, which in turn will improve the prognosis and quality of life of patients suffering from cancer.

Cancer uses computational models and algorithms to study clinical and molecular data. Machine Learning methods can also be used for classification of cancer types, discovering potential biomarkers, and improving understanding of tumor heterogeneity. This technology can also help in predicting cancer risk, understanding response to therapy, and developing personalized treatment strategies. In cancer research, the use of machine learning can improve the utilization of medical data and open new opportunities for the understanding, prevention, and treatment of this disease.

Polarimetric imaging systems that combine Machine Learning can help the diagnosis process and intervention decision making for cancer detection and staging. In this journal, we thoroughly study various commonly used machine teaching methods and present empirical comparisons between these methods on curated datasets of varying sizes and complexities. This method, through a review of the literature, provides a comprehensive overview of the difficulties and prospects of incorporating Data Science in cancer research[19][20].

4. DATA SCIENCE IN CANCER

The Data Science chapter in the context of cancer research is an essential component that has a major impact on the understanding and treatment of this disease. The application of Data Science technology opens wide doors to revolutionize various stages of cancer research, from the detection stage to the development of more effective treatment strategies. The main focus of this chapter is to detail in greater depth the challenges and opportunities that arise from the integration of Data Science within the cancer research paradigm. The significant challenges faced involve the multi-dimensional complexity of cancer data, managing large and complex data volumes, and the need for analytical models capable of addressing the heterogeneity of the disease.

By using data science, there are many opportunities to improve early diagnosis, determine individualized treatment, and gain a deeper understanding of risk factors. The latest developments in the field of data science

show inventive data analysis techniques and the most relevant research results. In this chapter, this in-depth analysis aims to provide a comprehensive understanding of how data science can help advance cancer research. The most worrying potential effect of reduced cancer screening is increased cancer mortality[21].

Strategic recommendations based on literature findings to guide future research and development. In doing so, readers will be provided with an in-depth look at the crucial role of Data Science in advancing the understanding and treatment of cancer, as well as exploring optimal ways to fully exploit the potential of Data Science integration in the treatment and management of cancer.

5. INTEGRATION OF DATA SCIENCE AND CANCER RESEARCH

The use of data science is critical in cancer research. This allows researchers to overcome the difficulties of analyzing complex health data and discover new patterns that can improve the understanding and treatment of cancer. However, the current dominant paradigm that views Complementary Therapies (TM) only as a complementary and alternative use option cannot encourage the latest research and development regarding the potential of traditional therapies (TM) as well as serendipitous discoveries that can bring innovation in the field of cancer. The Use of Epidemiology and Clinical Statistics in Cancer Research[22].

Epidemiology addresses the distribution of diseases and their causal factors in populations, while Clinical Statistics analyzes individual data regarding diagnosis, treatment, and patient outcomes. This combination enables a better understanding of disease from the population to the individual level, strengthens holistic strategies to improve population health, and provides effective clinical care.

Research on prognosis and therapy are two subfields of clinical epidemiology that need attention. Prognosis research involves research into the prognosis of a disease to improve the quality of health care and ensure positive outcomes. Prognosis research can help in analyzing the prognosis of cancer by using statistical and modeling techniques to predict patient health outcomes based on clinical data and prognostic factors.

Biology, health and medicine are fields that have developed from data science. Data science enables cancer research to comprehensively analyze genetic information, risk factors, response to therapy, and clinical patterns. By using methods such as machine learning, data science can help discover complex patterns that may not be detected with conventional analysis methods. This is especially important for cancer research because researchers can create creative data-analysis algorithms with programming, especially with languages like Python. In addition, data management through databases allows for easy storage and access to large amounts of health data, which is critical for cancer research. Table 1 shown the previous research.

Table 1. Previous Research

Title	Method	Results	Advantages and disadvantages
Behavioral Research in Cancer Prevention and Control: Emerging Challenges and Opportunities[23]	literature analysis or literature review methods to collect, evaluate, and synthesize previous research findings. As mentioned in the quote, they also use a multilevel and interdisciplinary approach.	Results include increased knowledge about the relationship between behavior and cancer risk, cancer management, and better prevention efforts. There may be specific findings regarding the COVID-19 pandemic, advances in techniques and technology, and an emphasis on the unique needs of the population and the increasing number of cancer patients.	Disadvantages: Interdisciplinary approach, use of data from multiple sources, and emphasis on sustainable outcomes for public health are some of the advantages of this journal. However, shortcomings may stem from limitations in the data or techniques used. Additionally, keep in mind that certain contextual variables may influence the results.
Exploring approaches for predictive cancer patient digital twins: Opportunities for collaboration and innovation[24]	High-Performance Computing (HPC) and Computational Modeling: Development and implementation of Patient Digital Twins (CPDTs) using high-performance computing (HPC)	demonstrated advances in in-depth phenotyping, treatment planning, and monitoring of treatment response	Advantages include the possibility of becoming more precise in cancer prevention and treatment, while disadvantages include the need for more long-term patient data and the difficulty of overcoming potential

	technology and computational modeling. Observational Data from Multiple Scales and Modalities: Studies construct CPDTs with multiple observational data from multiple scales and modalities. Transdisciplinary Research Community: This research investigates the development and implementation of CPDTs through a transdisciplinary research community at the intersection of high computing and cancer research.		limitations of CPDT implementation.
Application of informatics in cancer research and clinical practice: Opportunities and challenges[25]	data management, data science, data standards, terminology, high-speed omics data mining, machine learning algorithms, artificial intelligence in scanning, and intelligent radiation.	Informatics challenges in treatment decision making and patient outcomes are emphasized, while cross-disciplinary collaboration is considered key to future research and clinical applications.	The advantages include major benefits for research and treatment, the promise of earlier detection and more precise treatment, and optimization of cancer data acquisition and storage. Weaknesses include a lack of specific details about the methods and results of information technology applications, as well as a lack of clarity about aspects of the problem at hand.
Quality of Life in Breast Cancer Patients: A Systematic Review of the Qualitative Studies [26]	. Qualitative literature review on quality of life of people with breast cancer Searches were found in PubMed, Scopus, and Web of Science. Data are reported by theme for women without cancer and metastatic patients.	provides a deeper understanding of various aspects of quality of life (QOL) of women suffering from breast cancer. It is hoped that the information presented in this article will assist in the creation of better interventions that help patients and their families cope with life's difficulties.	Its advantage lies in the ability to develop more relevant interventions by understanding the social and cultural context. However, the generalizability of the findings, the subjectivity of methodological assessments, and data synthesis problems lead to limitations. Reliability and validity of research must be guaranteed through the consistent incorporation of results.
Arsenic and Environmental Health: State of the Science and Future Research Opportunities[27]	Literature review system	The document points to the adverse effects of arsenic exposure on health and the need for additional research, especially as it relates to dietary exposure. For more efficient research, methods are needed to reduce and integrate omics data.	Research on arsenic and environmental health is essential to understand the profound impact that arsenic has on society and to help create protective policies. However, the main obstacles are limited data, research complexity, and cost issues.
Integrative transcriptome analysis identifies MYBL2 as a poor prognosis marker for osteosarcoma and a	Using integrative transcriptome analysis, this study discovered MYBL2 as a marker of immune infiltration in various types of cancer and as a	In this article, the MYBL2 gene is discussed as a potential marker of osteosarcoma and its association with immune	provide valuable information for studying biomarkers and prognostic factors of osteosarcoma with clinical relevance, lack of explanation regarding the analytical

pan-cancer marker of immune infiltration[28]	deleterious marker for osteosarcoma.	infiltration in various types of cancer.	methods used, as well as lack of information regarding independent validation of research results
Integrative omics analysis reveals effective stratification and potential prognosis markers of pan-gastrointestinal cancers[29]	SNV analysis, somatic copy number alterations, DNA methylation data, mRNA expression, integrative clustering using the iClusterPlus algorithm, mutation signature analysis, and evaluation of the prognostic power of biomarkers	contribute to the understanding of pan-GI cancer pathophysiology and clinical therapy. The study also covers a variety of cancer-related topics, including molecular classification, gene expression analysis, tumor evolution, DNA methylation, and potential prognostic biomarkers	use of multi-omics data that allows identification of pan-gastrointestinal cancer subtypes with distinct survival and molecular features, The use of the iClusterPlus algorithm also allows accurate integrative clustering and identification of mutase signatures, limitations of the original omics data that may influence the statistical results, especially due to differences in sample sizes between cancers from the TCGA and ICGC databases, Additionally, the prognostic biomarker findings need to be further verified through in vitro and in vivo experiments to confirm their impact on pan-gastrointestinal cancers.
Operational Ontology for Oncology (O3): A Professional Society-Based, Multistakeholder, Consensus-Driven Informatics Standard Supporting Clinical and Research Use of Real-World Data From Patients Treated for Cancer[30]	multi-stakeholder collaboration and transparency in the development of operational frameworks for radiotherapy	O3 was developed in collaboration with various professional institutions and stakeholders, with the aim of facilitating the creation of comprehensive and interoperable datasets to support clinical quality improvement and research activities in the field of oncology. This ontology is also integrated into clinical, research, and professional society processes to improve data interoperability, as well as support the development of standard automated tools for data curation and analysis	There is a lack of research data available at this time, so readers cannot see empirical evidence of the implementation of the Operational Ontology for Radiotherapy (O3) in clinical practice. The advantage is a collaborative approach involving multiple stakeholders, including health care system vendors, government agencies, specialist groups in professional societies, institutions, and the public, engaging with a publicly accessible website to collect comments and direct engagement with representatives to ensure incorporation of a multi-stakeholder perspective showing that the research methods used involve multi-stakeholder collaboration and transparency in the development of operational frameworks for radiotherapy
Mechanistic insights into zearalenone-accelerated colorectal cancer in mice using integrative multi-omics approaches[31]	The methods used in this research were H&E staining, IHC, and TUNEL fluorescent staining for histology analysis and cell death detection, sample	revealed changes in the tumor transcriptome, serum metabolome, and gut microbiota, with pathways and metabolites potentially contributing to ZEA-induced	The strength of this study is the use of a comprehensive multi-omics approach, including histological analysis, metabolomics, and amino profiles, thus

	homogenization using a blade homogenizer and SCFA analysis using GC-MS, metabolomics analysis using UHPLC-QqQ-MS/MS for samples serum, amino acid profiling analysis using LC-MS/MS	tumor growth in colorectal cancer	providing a deep understanding of the mechanism of exposure to Zearalenone (ZEA) on tumor development in colo cancer. The weakness of this study is that there is no explanation in-depth regarding the potential side effects or toxicity of using ZEA, and there is no further research involving human populations to validate the findings in this study.
Circulating tumor DNA as liquid biopsy in lung cancer: Biological characteristics and clinical integration[32]	methods such as deep sequencing, fragment size analysis, and methylation profiling, as well as discussing the potential of ctDNA as a biomarker for personalized cancer treatment.	. These studies highlight the potential of ctDNA in predicting survival, monitoring prognosis, and detecting minimal residual disease in non-small cell lung cancer (NSCLC) patients, as well as highlighting the advantages and limitations of ctDNA as a liquid biopsy tool.	The advantage is its ability to determine complete tumor heterogeneity with a single blood draw thereby showing that liquid biopsy can be more efficient in understanding tumor heterogeneity by monitoring changes in ctDNA using liquid biopsy samples, which is beneficial for the majority of lung cancer patients, lacking in ctDNA analysis. One of them is the need for detailed information about the tumor genome for targeted approaches such as AS-PCR

6. DISCUSSION

Collecting, evaluating, and synthesizing previous research results is carried out through literature analysis or literature review techniques. In this process, an interdisciplinary and multilevel approach is used. Results include increased knowledge about the relationship of behavior to cancer risk, cancer management, and better prevention strategies. Specific findings may include information about the COVID-19 pandemic, advances in science and technology, as well as a focus on the unique needs of the population and the increasing number of cancer patients. There may be shortcomings in this journal because the data or methods used are limited. Additionally, results may be influenced by several contextual factors. However, its interdisciplinary approach, use of data from multiple sources, and emphasis on sustainable outcomes for public health are some of the advantages of this journal.

This study focuses on the use of computational modeling and High-Performance Computing (HPC) technology to develop and implement Patient Digital Twins (CPDTs). CPDTs were created to enable in-depth phenotyping, treatment planning, and monitoring of treatment response with observational data from multiple scales and modalities. This research promises greater accuracy in cancer prevention and treatment through a transdisciplinary research community at the intersection of advanced computing and cancer research. However, there are several issues that need to be addressed, such as the need for more long-term patient data and the difficulty of addressing potential limitations that may occur when implementing CPDT.

The study emphasizes data management, data science, data standards, terminology, high-speed omics data mining, machine learning algorithms, artificial intelligence in scanning, and intelligent radiation in relation to patient treatment decision making and outcomes. Cross-disciplinary collaboration is considered essential for future research and clinical applications, while advantages include major benefits for research and treatment, the promise of earlier detection and more precise treatment, and optimization of cancer data acquisition and storage. However, the weakness lies in the lack of clarity about the elements of the problem at hand and the lack of specific details about the techniques and results of information technology applications.

A better understanding of various aspects of quality of life (QOL) of women with breast cancer was gained from a review of qualitative literature on the quality of life of people with breast cancer. It is hoped that the creation of better interventions will help patients and their families face life's challenges. The ability to

develop more relevant interventions by understanding the social and cultural context is an advantage. Limitations arise due to the generalizability of findings, the subjectivity of methodological assessments, and data synthesis issues. To ensure the validity and credibility of the research, the results must be combined.

A person can experience adverse effects on their health if exposed to arsenic, especially if exposed through food. For more efficient research, additional research is needed to reduce and integrate omics data. Environmental health and arsenic studies are critical to understanding the impact of arsenic on society and helping to create protective policies. However, major obstacles include the limited amount of data, research complexity, and cost issues. Therefore, efforts are needed to ensure that research is valid and reliable by incorporating consistent findings.

Using integrative transcriptome analysis, this study identified MYBL2 as a marker of immune infiltration in various types of cancer and as a poor marker for osteosarcoma. The article discusses the role of the MYBL2 gene as a potential marker for osteosarcoma and its relationship with immune infiltration in various types of cancer. This study has shortcomings, such as a lack of explanation of the analytical techniques used and a lack of information about independent validation of research results. However, they also provide useful information for understanding osteosarcoma biomarkers and prognostic factors with clinical relevance.

Multi-omics analysis, such as SNVs, somatic copy number alterations, DNA methylation data, and mRNA expression, was performed in this study, using the iClusterPlus algorithm for integrative clustering. These studies cover topics such as molecular classification, gene expression analysis, tumor evolution, DNA methylation, and potential prognostic biomarkers. It also helps understand pan-gastrointestinal cancer pathophysiology and clinical therapy. With multi-omics data, pan-gastrointestinal cancer subtypes with multiple survival and molecular features can be identified, while the iClusterPlus algorithm enables accurate integrative clustering and mutation signature discovery. However, the statistical results were affected by the lack of original omics data, especially the difference in cancer sample sizes between the TCGA and ICGC databases. To confirm its impact on pan-gastrointestinal cancers, the results of prognostic biomarkers should be further verified through in vitro and in vivo tests.

The Operational Ontology for Radiotherapy (O3) was developed through collaboration involving various stakeholders and professional institutions in the field of oncology. O3 supports clinical quality improvement, encourages research activities, and encourages the creation of comprehensive and interoperable datasets. To improve data interoperability, these ontologies are used in clinical processes, research, and professional communities. Additionally, it helps the development of conventional automated tools for data curation and analysis. Although its strengths include working collaboratively with multiple stakeholders, readers were unable to find empirical evidence of the application of O3 in clinical practice. This research method shows that creating an operational framework for radiotherapy requires collaboration and transparency from various stakeholders.

The impact of Zearalenone (ZEA) exposure on colorectal cancer was assessed by histology, SCFA, metabolomics and amino profile analysis. The results showed changes in the gut microbiota, serum metabolome, and tumor transcriptome. The advantage of this research is the multi-omics approach which provides in-depth understanding. However, drawbacks include explaining the potential side effects of ZEA and the fact that further research is needed to validate the findings in human populations.

This study evaluates the possibility of ctDNA as a biomarker in personalized cancer treatment using deep sequencing methods, fragment size analysis and methylation profiling. This study shows that ctDNA helps predict survival, monitor prognosis, and find minimal residual disease in non-small cell lung cancer (NSCLC) patients. Its advantage is its ability to determine complete tumor heterogeneity with a single blood draw, allowing a more efficient understanding of changes in ctDNA. However, the drawback lies in ctDNA analysis, which requires detailed information about the tumor genome, especially for targeted approaches such as AS-PCR.

7. TREND ANALYSIS OF DATA SCIENCE INTEGRATION IN CANCER RESEARCH

Therefore, we suggest conducting a bibliometric analysis of the literature published in the last ten years regarding the relationship between physical activity or exercise and breast cancer to identify trends in publication rates and to determine the most common research topics in this area [33]. In most countries, incidence and mortality trends are more pronounced among the elderly [34]. Self-care is an important component that can improve health and quality [35] the lives of cancer survivors.

In recent years, the integration of data science in cancer research has become increasingly significant. This is mainly due to advances in information and computing technology that enable more complex and in-depth data analysis. Here are some of the key trends in data science integration in cancer research:

a. Utilization of Big Data in Cancer Analysis
Many branches of biomedical health science are influenced by big data analysis. However, managing and storing large amounts of data and analyzing and interpreting the results quickly on local computers is a big challenge. In cancer analysis, big data has become a major trend. Data scientists can analyze very large and complex health data, such as genetic, clinical, and population data. Researchers can discover patterns that might not have been detected previously using methods such as machine learning and big data analysis. This provides new insights into risk factors, disease progression, and response to cancer therapy[36][37].
b. Predictive Model Development
Data science has enabled the development of predictive models for cancer. By utilizing machine learning techniques, researchers can build predictive models to predict the risk of cancer development, response to therapy, and patient prognosis. This helps in the development of more personalized therapies and more precise preventative measures.
c. Genomic and Proteomic Analysis
The integration of data science has enabled deeper analysis of genomic and proteomic data in the context of cancer. With advanced data analysis techniques, researchers can identify genetic mutations and changes at the protein level that are associated with the development of cancer. This helps in further understanding of the biological mechanisms of cancer and the development of more targeted therapies.
d. Development of Cancer Data Analysis Tools
A local knowledge assessment tool aimed at women could be used to change the approach to early detection of prostate cancer (PCa) in pharmacy practice. Values clarification tools may be a valuable additional tool for use in Shared Decisions in this particular medical context, with its complex and high-stakes life-or-death decisions. Additionally, research looking at creating prioritization tools and conversation aids found that patients became overwhelmed if asked to sacrifice more than four health outcomes[38][39][40]. The development of new tools to analyze cancer data is also a trend. Data scientists can create more sophisticated cancer data analysis algorithms, as well as data visualization tools, which make it possible to better explore and understand cancer data. help make better decisions about cancer diagnosis, therapy, and control. The integration of data science in cancer research has opened up new opportunities in understanding and treating this disease. As data science technology and methodologies continue to develop, it is hoped that this integration will continue to make a significant contribution to efforts to prevent, diagnose and treat cancer.

8. CHALLENGES OF DATA SCIENCE INTEGRATION IN CANCER RESEARCH

Big data curation and management typically includes data capture, quality control, and validation to ensure that data is accurate, complete, and reliable and meets ethical and legal requirements. identify candidate patients who may respond to immunotherapy to monitor treatment outcomes and discover mechanisms of resistance. As a result, this endangers a person's life and health[41][42].

Although there is no definitive definition, “cancer big data” usually refers to the collection of data from multiple sources:

a. Data Diversity: Many types of data are used in cancer research, including genomic data, clinical data, and medical image data. Addressing this diversity and integrating data from multiple sources are key challenges in achieving a broader understanding of cancer.
b. Privacy and Security: When personal health data is used in research, there are privacy and security concerns. Developing an effective data security strategy and ensuring regulatory compliance is critical.
c. Algorithm Scalability: Ensuring that data science algorithms can efficiently manage the enormous and complex volumes of data generated by cancer research is a technical challenge that must be resolved.
d. Validation and Replication of Results: Research results should be validated and replicated using independent datasets to ensure that the results are reliable. Due to limited access to relevant datasets, this is often a challenge.

Cancer, a serious, fatal disease caused by uncontrolled cell growth, is the biggest challenge facing the world of medicine and health. However, the challenge associated with miRNA-based therapy is the delivery of miRNA-based therapy to target cells, which inevitably requires optimization[43][44][45].

Among cancer patients, the burden of care is largely related to uncertainty of diagnosis, longer course of disease, recurrent acute phase of disease, doubtful prognosis, longer treatment requirements, repeated hospital stays, and more experience with side effects of polyvalent treatment options[46][47][48].

9. OPPORTUNITIES TO INTEGRATE DATA SCIENCE IN CANCER RESEARCH

It is fortunate that there is increasing interest in establishing research programs related to the use of Ayurveda as supportive therapy, rehabilitation, and palliative care. On the contrary, critical analysis of the literature opens the door to the anti-cancer potential of Withania somnifera (WS), which may play an important role in cancer prevention. Additionally, there are many opportunities with artificial intelligence (AI). This could be used to improve the quality of liquid biopsy testing and be incorporated into clinical workflows in the future. Furthermore, the drug sensitivity profile of organoid cultures of circulating tumor cells (CTCs) was associated with response to treatment. This creates exciting opportunities to integrate data science in cancer research, increase public understanding of cancer, and improve the efficiency of cancer diagnosis, therapy, and management[49][50][51][52][53].

a. Genomic Mapping and Personalized Therapies: Data science enables better genomic mapping, which enables the development of personalized therapies that can improve treatment effectiveness.
b. Prediction and Early Detection: To predict cancer risk or detect cancer at an early stage, machine teaching algorithms can be used to analyze patterns in clinical and genomic data.
c. Clinical Research Optimization: Data science integration can accelerate the clinical research process, from patient recruitment to outcome analysis. This can speed up the development of drugs and therapies.
d. Discovery of New Knowledge: Data Science can help discover new things and discover relationships and patterns that may be missed in traditional research by analyzing data from various sources thoroughly (building).

10. Implications and Benefits

Expression of protein-coding genes can be inhibited or enhanced by lncRNA transcriptional control. By enabling healthcare providers to use decision support systems, they can simplify financial and administrative processes and increase the efficiency and cost-effectiveness of treatment. This has had a significant impact on the health sector. Evaluate the current and projected economic impact of cancer, as well as other diseases. Interventions targeting clustered symptoms may be more beneficial than treating nausea alone because correlations between individual symptoms may have a synergistic negative impact on clinical outcomes[54][55][56][57]. But great efforts are being made to add layers of proteomics, which provide further functional information about tumor behavior and downstream effects of mutations. The goal of this effort is to improve our understanding of the biological characteristics of tumors and the impact of the genetic changes that occur[58][59][60][61][62].

The results of the study significantly showed that the low-risk group of patients had a better prognosis and a better response to immune checkpoint blockade therapy. However, there is still much debate about the value of a single surgeon's operative volume in the context of centralized thoracic oncology services[63][64].

These findings create interesting dynamics to draw on in future research in the journal entitled "Challenges and Opportunities: Integrating Data Science in Cancer Research Through a Literature Review Approach." Understanding more about the relationship between cancer risk factors, response to therapy, and how the number of operations performed by a surgeon can influence patient outcomes is challenging. Instead, it is possible that data science can be used to analyze literature and investigate factors holistically.

This research hopes to provide a broader picture of current trends and status in the application of science in cancer research with a focus on a review of the literature. As a result, it is hoped that the results of this research will provide a strong basis for finding innovative solutions and harmonizing the problems faced in efforts to optimize cancer management. In addition, the application of science in cancer research does not only have a direct impact on the development of cancer research.

CONCLUSION

The takeaway from this summary of cancer research articles is that cancer research today covers a lot of ground, such as treatment, patient outcomes, caregiver responsibilities, and the use of advanced technologies such as AI and liquid biopsies. These articles provide valuable insight into recent developments in cancer research as well as future opportunities and obstacles. These articles show that the integration of data science in cancer research has great potential to change people's understanding of and treatment of cancer. However,

they face problems because the data is complex and requires better analysis models. Additionally, this study emphasizes the use of multi-omics methods to understand the effects of zearalenone exposure on colorectal cancer as well as the possibility of ctDNA as a biomarker in the treatment of non-small cell lung cancer. This study provides a comprehensive overview of various aspects of current cancer research and potential research directions in the future as the integration of data science in cancer research has become important in recent years. It also enables more complex and in-depth data analysis.

REFERENCES

[1] A. Schmutz et al., "Mapping the European cancer prevention research landscape: A case for more prevention research funding," Eur. J. Cancer, vol. 195, no. April, p. 113378, 2023, doi: 10.1016/j.ejca.2023.113378.

[2] C. Paterson, L. Armitage, and M. Turner, "Current Landscape of Ecological Momentary Assessment (Real-Time Data) Methodology in Cancer Research: A Systematic Review,"Semin. Oncol. Nurs., vol. 39, no. 6, p. 151514, 2023, doi: 10.1016/j.soncn.2023.151514.

[3] X. Wu, W. Li, and H. Tu, "Big data and artificial intelligence in cancer research,"Trends in Cancer, pp. 1–14, 2023, doi: 10.1016/j.trecan.2023.10.006.

[4] E. Parimbelli et al., "A review of AI and Data Science support for cancer management," Artif. Intel. Med., vol. 117, no. August 2020, p. 102111, 2021, doi: 10.1016/j.artmed.2021.102111.

[5] RR Expósito and J. González-Domínguez, "BigDEC: A multi-algorithm Big Data tool based on the k-mer spectrum method for scalable short-read error correction," Future. Gener. Comput. Syst., Jan. 2024, doi: 10.1016/j.future.2024.01.011.

[6] A. Nani, "Valuing big data: An analysis of current regulations and proposals of frameworks,"Int. J. Accounts. Inf. Syst., vol. 51, p. 100637, Dec. 2023, doi: 10.1016/j.accinf.2023.100637.

[7] M. Resell, M. Stranden, G. Qvigstad, D. Chen, and C.-M. Zhao, "Gaps between needs of patients and public involvement and interests of researchers on pancreatic cancer," Patient Educ. Couns., vol. 120, no. July 2023, p. 108125, 2023, doi: 10.1016/j.pec.2023.108125.

[8] G. Curigliano et al., "Understanding breast cancer complexity to improve patient outcomes : The St Gallen International Consensus Conference for the Primary Therapy of Individuals with Early Breast Cancer 2023," no. 11, 2023, doi: 10.1016/j.annonc.2023.08.017.

[9] C. Ladbury et al., "Integration of artificial intelligence in lung cancer : Rise of the machine," Cell Reports Med., vol. 4, no. 2, p. 100933, 2023, doi: 10.1016/j.xcrm.2023.100933.

[10] BF Darst et al., "Evaluating approaches for constructing polygenic risk scores for prostate cancer in men of African and European ancestry," Am. J Hum. Genet., vol. 110, no. 7, pp. 1200–1206, 2023, doi: 10.1016/j.ajhg.2023.05.010.

[11] G. Marra et al., "Impact of Epithelial Histological Types, Subtypes, and Growth Patterns on Oncological Outcomes for Patients with Nonmetastatic Prostate Cancer Treated with Curative Intent: A Systematic Review," Eur. Urol., vol. 84, no. 1, pp. 65–85, 2023, doi: 10.1016/j.eururo.2023.03.014.

[12] V. Zamora et al., "Comparative effectiveness of new treatment modalities for localized prostate cancer through patient-reported outcome measures," Clin. Transl. Radiat. Oncol., vol. 44, no. October 2023, p. 100694, 2024, doi: 10.1016/j.ctro.2023.100694.

[13] B. Balkhi, A. Alghamdi, S. Alqahtani, M. Al Najjar, A. Al Harbi, and T. Bin Traiki, "Colorectal cancer-related resource utilization and healthcare costs in Saudi Arabia," Saudi Pharm. J., vol. 31, no. 11, p. 101822, 2023, doi: 10.1016/j.jsps.2023.101822.

[14] LL Yang, M. Li, W. Huang, PT Ren, QH Yan, and YH Hao, "ANP32B promotes colorectal cancer cell progression and reduces cell sensitivity to PRAP1 inhibitor through up-regulating HPF1," Heliyon, vol. 10, no. 1, p. e23829, 2024, doi: 10.1016/j.heliyon.2023.e23829.

[15] S. Yue, Q. Wang, J. Zhang, Q. Hu, and C. Liu, "Understanding cervical cancer at single-cell resolution,"Cancer Lett., vol. 576, no. September, p. 216408, 2023, doi: 10.1016/j.canlet.2023.216408.

[16] S. Colucci, F. M. Donini, and E. Di Sciascio, "Logical comparison over RDF resources in bio-informatics," J. Biomed. Inform., vol. 76, pp. 87–101, Dec. 2017, doi: 10.1016/j.jbi.2017.11.004.

[17] G. Dall' Agnese et al., "Role of condensates in modulating DNA repair pathways and its implications for chemoresistance," J. Biol. Chem., vol. 299, no. 6, p. 104800, Jun. 2023, doi: 10.1016/j.jbc.2023.104800.

[18] U. Neri et al., “Expansion of the global RNA virome reveals diverse clades of bacteriophages,” Cell, vol. 185, no. 21, pp. 4023-4037.e18, Oct. 2022, doi: 10.1016/j.cell.2022.08.023.

[19] T.-T.-H. Pham, T.-N. Luu, T.-V. Nguyen, N.-T. Huynh, Q.-H. Phan, and T.-H. Le, “Polarimetric imaging combining optical parameters for classification of mice non-melanoma skin cancer tissue using machine learning,” Heliyon, vol. 9, no. 11, p. e22081, Nov. 2023, doi: 10.1016/j.heliyon.2023.e22081.

[20] W. Tizi and A. Berrado, “Machine learning for survival analysis in cancer research: A comparative study,”Sci. African, vol. 21, p. e01880, Sep. 2023, doi: 10.1016/j.sciaf.2023.e01880.

[21] UA Emirates, “A potential rise of breast cancer risk in the UAE post-COVID-19 lockdown : A call for action,” vol. 79, no. June, pp. 8–11, 2022, doi: 10.1016/j.amsu.2022.103976.

[22] M. Soler, E. Lu, R. Masch, K. Alfaro, J.R. Anderson, and M. Cremer, “Affordable cancer technologies: Lessons learned from the design and implementation of two randomized clinical trials to develop innovative treatments for cervical precancer,” Contemp. Clin. Trials Commun., vol. 36, p. 101224, Dec. 2023, doi: 10.1016/j.conctc.2023.101224.

[23] WMP Klein et al., “Behavioral Research in Cancer Prevention and Control : Emerging Challenges and Opportunities,” vol. 114, pp. 179–186, 2022, doi: 10.1093/jnci/djab139.

[24] EA Stahlberg et al., “Exploring approaches for predictive cancer patient digital twins : Opportunities for collaboration and innovation,” no. October, pp. 1–10, 2022, doi: 10.3389/fdgth.2022.1007784.

[25] N. Hong and J. Wang, “Application of informatics in cancer research and clinical practice : Opportunities and challenges,” no. April, pp. 80–91, 2022, doi: 10.1002/cai2.9.

[26] Z. Heidary, M. Ghaemi, BH Rashidi, OK Gargari, and A. Montazeri, "Quality of Life in Breast Cancer Patients : A Systematic Review of the Qualitative Studies," vol. 30, pp. 1–10, 2023, doi: 10.1177/10732748231168318.

[27] DJ Carlin et al., “Arsenic and Environmental Health : State of the Science and Future Research Opportunities,” vol. 890, no. 7, pp. 890–899, 2016.

[28] X. Qiu et al., “Integrative transcriptome analysis identifies MYBL2 as a poor prognosis marker for osteosarcoma and a pan-cancer marker of immune infiltration,” Genes Dis., vol. 11, no. 3, p. 101004, May 2024, doi: 10.1016/j.gendis.2023.04.035.

[29] H. Jiangzhou et al., “Integrative omics analysis reveals effective stratification and potential prognosis markers of pan-gastrointestinal cancers,” iScience, vol. 24, no. 8, p. 102824, Aug. 2021, doi: 10.1016/j.isci.2021.102824.

[30] CS Mayo et al., “Operational Ontology for Oncology (O3): A Professional Society-Based, Multistakeholder, Consensus-Driven Informatics Standard Supporting Clinical and Research Use of Real-World Data From Patients Treated for Cancer,” Int. J. Radiat. Oncol., vol. 117, no. 3, pp. 533–550, Nov. 2023, doi: 10.1016/j.ijrobp.2023.05.033.

[31] EKK Lo et al., “Mechanistic insights into zearalenone-accelerated colorectal cancer in mice using integrative multi-omics approaches,” Comput. Struct. Biotechnol. J., vol. 21, pp. 1785–1796, 2023, doi: 10.1016/j.csbj.2023.02.048.

[32] C. Li, J. Shao, P. Li, J. Feng, J. Li, and C. Wang, “Circulating tumor DNA as liquid biopsy in lung cancer: Biological characteristics and clinical integration,” Cancer Lett., vol. 577, p. 216365, Nov. 2023, doi: 10.1016/j.canlet.2023.216365.

[33] L. Luis and R. Escamilla-ugarte, “Heliyon Trends in physical activity research for breast cancer - A bibliometric analysis of the past ten years,” vol. 9, no. September, 2023, doi: 10.1016/j.heliyon.2023.e22499.

[34] B. Zhou et al., "Global burden and temporal trends in incidence and mortality of oesophageal cancer," J. Adv. Res., vol. 50, pp. 135–144, 2023, doi: 10.1016/j.jare.2022.10.007.

[35] J. Yeom, I. Yeom, H. Park, and S. Lim, "European Journal of Oncology Nursing Cultural factors affecting the self-care of cancer survivors : An integrative review," Eur. J. Oncol. Nurs., vol. 59, no. June, p. 102165, 2022, doi: 10.1016/j.ejon.2022.102165.

[36] A. Dhasmana et al., “Integrative big transcriptomics data analysis implicates crucial role of MUC13 in pancreatic cancer,” Comput. Struct. Biotechnol. J., vol. 21, pp. 2845–2857, 2023, doi: 10.1016/j.csbj.2023.04.029.

[37] T. Amano, I. Kinoshita, and M. Yoichi Ito, "P65-1 Development of genome data repository system for big data utilization and secure storage," Ann. Oncol., vol. 34, p. S1464, 2023, doi: 10.1016/j.annonc.2023.09.419.

[38] E. Wiafe, KB Mensah, IF Adjei, F. Oosthuizen, and V. Bangalee, "A psychometric analysis of a Ghanaian tool (PCaKAB-Gh) for studying prostate cancer knowledge, attitudes and beliefs among women," Sci. African, vol. 21, p. e01800, Sep. 2023, doi: 10.1016/j.sciaf.2023.e01800.

[39] JLP van Gurp, LGG van Lent, N. Stoel, CCD van der Rijt, JCM van Weert, and J. Hasselaar, "Accentuating patient values in shared decision-making: A mixed methods development of an online value clarification tool and communication training in the context of early phase clinical cancer trials," Patient Educ. Couns., vol. 119, p. 108075, Feb. 2024, doi: 10.1016/j.pec.2023.108075.

[40] PAL Seghers et al., "Development and testing of the Outcome Prioritization Tool adjusted for older patients with cancer: A pilot study," J. Geriatr. Oncol., vol. 14, no. 8, p. 101590, Nov. 2023, doi: 10.1016/j.jgo.2023.101590.

[41] H. Ao, Z. Xin, and Z. Jian, "Liquid biopsy to identify biomarkers for immunotherapy in hepatocellular carcinoma,"Biomark. Res., vol. 9, no. 1, p. 91, Dec. 2021, doi: 10.1186/s40364-021-00348-y.

[42] J. Zhou et al., "Guidelines for the Diagnosis and Treatment of Hepatocellular Carcinoma (2019 Edition)," Liver Cancer, vol. 9, no. 6, pp. 682–720, 2020, doi: 10.1159/000509424.

[43] J. Akter, KD Konlan, M. Nesa, and A. Ispriantari, "Factors influencing cancer patients' caregivers' burden and quality of life: An integrative review," Heliyon, vol. 9, no. 11, p. e21243, Nov. 2023, doi: 10.1016/j.heliyon.2023.e21243.

[44] S. Bhattacharya, RK Mahato, S. Singh, GK Bhatti, SS Mastana, and JS Bhatti, "Advances and challenges in thyroid cancer: The interplay of genetic modulators, targeted therapies, and AI-driven approaches," Life Sci., vol. 332, p. 122110, Nov. 2023, doi: 10.1016/j.lfs.2023.122110.

[45] SRU Din, S. Saeed, SU Khan, FM Arbi, G. Xuefang, and M. Zhong, "Bacteria-driven cancer therapy: Exploring advances and challenges," Crit. Rev. Oncol. Hematol., vol. 191, p. 104141, Nov. 2023, doi: 10.1016/j.critrevonc.2023.104141.

[46] J.-Y. Tan, A. Molassiotis, M. Lloyd-Williams, and J. Yorke, "Burden, emotional distress and quality of life among informal caregivers of lung cancer patients: An exploratory study," Eur. J. Cancer Care (Engl)., vol. 27, no. 1, p. e12691, Jan. 2018, doi: 10.1111/ecc.12691.

[47] SG Eskin, R. Yuksel, and H. Bulut, "Investigation of the Relationship between Care Giving Burden and Quality of Life of Caregivers of Cancer Patients," J. Evol. Med. Dent. Sci., vol. 10, no. 9, pp. 567–572, March. 2021, doi: 10.14260/jemds/2021/123.

[48] Z. Liu, C. Heffernan, and J. Tan, "Caregiver burden: A concept analysis,"Int. J. Nurs. Sci., vol. 7, no. 4, pp. 438–445, Oct. 2020, doi: 10.1016/j.ijnss.2020.07.012.

[49] JT Arnold, "Integrating ayurvedic medicine into cancer research programs part 2: Ayurvedic herbs and research opportunities," J. Ayurveda Integr. Med., vol. 14, no. 2, p. 100677, Mar. 2023, doi: 10.1016/j.jaim.2022.100677.

[50] A. Saggam et al., "Withania somnifera (L.) Dunal: A potential therapeutic adjuvant in cancer," J. Ethnopharmacol., vol. 255, p. 112759, Jun. 2020, doi: 10.1016/j.jep.2020.112759.

[51] Y.-C. Zhong, J.-W. Cheng, P.-X. Wang, J. Fan, J. Zhou, and X.-R. Yang, "Opportunities and challenges of liquid biopsy in liver cancer," Clin. Surg. Oncol., vol. 2, no. 4, p. 100026, Dec. 2023, doi: 10.1016/j.cson.2023.100026.

[52] O. Ginghina et al., "Liquid Biopsy and Artificial Intelligence as Tools to Detect Signatures of Colorectal Malignancies: A Modern Approach in Patient's Stratification," Front. Oncol., vol. 12, March. 2022, doi: 10.3389/fonc.2022.856575.

[53] Y.-H. Woo et al., "Correlation between drug sensitivity profiles of circulating tumor cell-derived organoids and clinical treatment response in patients with pancreatic ductal adenocarcinoma," Eur. J. Cancer, vol. 166, pp. 208–218, May 2022, doi: 10.1016/j.ejca.2022.01.030.

[54] A. Ghorbani et al., "The prognostic, diagnostic, and therapeutic impact of Long noncoding RNAs in gastric cancer," Cancer Genet., vol. 282–283, pp. 14–26, Apr. 2024, doi: 10.1016/j.cancergen.2023.12.006.

[55] AZ Dag, M. Johnson, E. Kibis, S. Simsek, B. Cankaya, and D. Delen, "A machine learning decision support system for determining the primary factors impacting cancer survival and their temporal effect," Health c. Anal., vol. 4, p. 100263, Dec. 2023, doi: 10.1016/j.health.2023.100263.

[56] M. De Camargo Cancela et al., "The economic impact of cancer mortality among working-age individuals in Brazil from 2001 to 2030," Cancer Epidemiol., vol. 86, p. 102438, Oct. 2023, doi: 10.1016/j.canep.2023.102438.

[57] M. van den Brink et al., "Nausea and nausea-related symptoms in children with cancer: Presence, severity, risk factors and impact on quality of life during the first year of treatment," EJC Paediatr. Oncol., vol. 2, p. 100128, Dec. 2023, doi: 10.1016/j.ejcped.2023.100128.

[58] G. Arad and T. Geiger, "Functional Impact of Protein–RNA Variation in Clinical Cancer Analyses," Mol. Cell. Proteomics, vol. 22, no. 7, p. 100587, Jul. 2023, doi: 10.1016/j.mcpro.2023.100587.

[59] P. Woo et al., "Integration and Analysis of CPTAC Proteomics Data in the Context of Cancer Genomics in the cBioPortal," Mol. Cell. Proteomics, vol. 18, no. 9, pp. 1893–1898, Sept. 2019, doi: 10.1074/mcp.TIR119.001673.

[60] S. Doll, F. Gnad, and M. Mann, "The case for proteomics and phospho-proteomics in personalized cancer medicine,"PROTEOMICS--Clinical Appl., vol. 13, no. 2, p. 1800113, 2019.

[61] A. Macklin, S. Khan, and T. Kislinger, "Recent advances in mass spectrometry based clinical proteomics: applications to cancer research," Clin. Proteomics, vol. 17, pp. 1–25, 2020.

[62] DR Mani et al., "Cancer proteogenomics: current impact and future prospects," Nat. Rev. Cancer, vol. 22, no. 5, pp. 298–313, May 2022, doi: 10.1038/s41568-022-00446-5.

[63] X. Zhang et al., "The implications of integrative multiple RNA modification-based subtypes in gastric cancer immunotherapy and prognosis," iScience, p. 108897, Jan. 2024, doi: 10.1016/j.isci.2024.108897.

[64] CM Maxwell et al., "Comprehensive value implications of surgeon volume for lung cancer surgery: Use of an analytic framework within a regional health system," JTCVS Open, Nov. 2023, doi: 10.1016/j.xjon.2023.11.010.