Int. j. adv. multidisc. res. stud. 2025; 5(1):1344-1351

Received: 08-01-2025
Accepted: 18-02-2025

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Harnessing the Power of Health Data Analytics to Combat Non-Communicable Diseases: A Review of AI and Data Science Approaches

[1] Erica Afrihyia, [2] Olufunke Omotayo, [3] Ashiata Yetunde Mustapha, [4] Opeoluwa Oluwanifemi Akomolafe, [5] Ernest Chinonso Chianumba, [6] Adelaide Yeboah Forkuo

[1] Independent Researcher, Ohio, USA
[2] Independent Researcher, Alberta, Canada
[3] Kwara State Ministry of Health, Nigeria
[4] Independent Researcher, UK
[5] Department of Computer Science & Information Technology, School of Computing, Montclair State University, United States
[6] Independent Researcher, Ghana

DOI: https://doi.org/10.62225/2583049X.2025.5.1.4299

Corresponding Author: **Erica Afrihyia**

Abstract

Non-communicable diseases (NCDs) such as cardiovascular diseases, cancer, diabetes, and chronic respiratory conditions are leading causes of morbidity and mortality globally. The increasing availability of health data and advancements in artificial intelligence (AI) and data science offer significant opportunities to improve NCD management. This review explores key AI and data science approaches, including machine learning (ML) for predictive analytics, natural language processing (NLP) for analyzing unstructured health data, and big data technologies for real-time monitoring. Additionally, the role of AI in personalized medicine and treatment optimization is discussed. While AI-driven innovations have the potential to transform healthcare, challenges such as data privacy, algorithmic bias, and technical limitations must be addressed. The paper also provides recommendations for enhancing AI applications in NCD management, emphasizing collaboration between healthcare providers, technologists, and policymakers.

Keywords: Non-communicable Diseases (NCDs), Artificial Intelligence (AI), Machine Learning (ML), Health Data Analytics, Personalized Medicine, Data Privacy and Security

1. Introduction

1.1 Context and Importance

Non-communicable diseases (NCDs) pose a major global public health challenge. Unlike communicable diseases caused by infections, NCDs are chronic conditions that arise from a combination of genetic, environmental, and lifestyle factors (Coates *et al*., 2020) [6]. The most common NCDs include cardiovascular diseases, cancers, chronic respiratory diseases, and diabetes. According to the World Health Organization (WHO), NCDs are responsible for over 70% of global deaths annually, claiming more than 41 million lives. Of these, 15 million deaths occur in individuals between the ages of 30 and 69, making NCDs not just a health issue but also a socio-economic one as they disproportionately affect younger populations in their most productive years (Ramesh & Kosalram, 2023) [34].

NCDs place immense pressure on health systems, particularly in low- and middle-income countries where 85% of premature deaths from NCDs occur. These diseases also lead to substantial economic losses due to long-term treatment costs, loss of productivity, and the need for sustained healthcare support (Ganju *et al*., 2020) [10]. Increasing urbanization, sedentary lifestyles, poor dietary choices, tobacco use, and the harmful use of alcohol fuel the global rise of NCDs. These risk factors contribute to the long-term development of these diseases, making prevention, early detection, and effective management crucial. However, traditional methods of addressing NCDs have been reactive, focusing on treatment rather than prevention or personalized care.

This has led to a need for innovative approaches that can offer solutions for managing NCDs more effectively (Haque *et al*., 2020) [12].

1.2 Health Data Analytics and AI
The healthcare industry is witnessing a transformative shift due to the growing availability of health data and advancements in artificial intelligence (AI) and data science. Fueled by large volumes of clinical, genomic, behavioral, and environmental data, health data analytics has created unprecedented opportunities for improving the prevention, diagnosis, and treatment of diseases, especially NCDs. Integrating AI techniques such as machine learning (ML), deep learning, and natural language processing (NLP) has further enhanced the ability to harness this data for predictive and prescriptive purposes (Paramesha, Rane, & Rane, 2024) [31].

AI-driven health data analytics has emerged as a powerful tool in combating NCDs by enabling the processing and analysis of large-scale health data to uncover hidden patterns, trends, and associations. For example, machine learning algorithms can analyze electronic health records (EHRs) to predict patient outcomes, identify high-risk individuals, and personalize treatment plans. Similarly, wearable devices, mobile health apps, and remote monitoring systems generate continuous streams of data that, when analyzed, can provide real-time insights into a patient's health status, helping to detect early signs of diseases or exacerbations (Tahir & Khan, 2023) [41].

Furthermore, AI tools offer significant promise in the area of precision medicine, where treatment is tailored to the individual characteristics of each patient, taking into account genetic, environmental, and lifestyle factors. By leveraging health data analytics, healthcare providers can optimize treatment regimens, improve disease management strategies, and deliver more effective and efficient care. This shift from reactive to proactive healthcare holds great potential to reduce the burden of NCDs by enhancing early detection, promoting personalized interventions, and improving long-term patient outcomes (Bawack, Fosso Wamba, & Carillo, 2021) [5].

1.3 Objective
This paper aims to provide a comprehensive review of the AI and data science approaches currently being utilized to combat non-communicable diseases. It aims to explore the role of these technologies in addressing key challenges associated with NCDs, such as early diagnosis, personalized treatment, and ongoing management. Specifically, the paper will focus on the use of machine learning, natural language processing, big data analytics, and real-time health monitoring systems to advance the fight against NCDs.

The review will examine how these approaches can be leveraged to improve predictive accuracy, identify at-risk populations, and optimize treatment plans. In addition, the paper will highlight the challenges associated with implementing AI and data-driven solutions in healthcare, including concerns around data privacy, algorithmic bias, and the need for robust infrastructure. Ethical considerations will also be discussed to ensure that AI applications in healthcare are developed and deployed fairly, transparently, and beneficial to all patients, particularly those in vulnerable populations.

Ultimately, this paper seeks to underscore the transformative potential of AI and health data analytics in the fight against NCDs. Reviewing the current state of research and development in this area aims to provide insights into future directions for innovation and collaboration between healthcare professionals, researchers, and policymakers. The objective is to highlight the benefits of these technologies and propose recommendations for overcoming barriers to their widespread adoption and ensuring that they are used responsibly to improve patient outcomes and reduce the global burden of NCDs.

2. Overview of Non-Communicable Diseases (NCDs)
2.1 Key NCDs
Non-communicable diseases (NCDs) are chronic conditions not transmitted from person to person, typically developing over long periods and progressing slowly. They account for a substantial portion of the global disease burden. Among the most prevalent NCDs are cardiovascular diseases (CVDs), diabetes, cancers, and chronic respiratory diseases. These four categories collectively cause the majority of NCD-related deaths and disabilities worldwide (Luna & Luyckx, 2020) [19].

Cardiovascular diseases (CVDs), which include heart attacks, strokes, and hypertension, are the leading cause of death globally. The World Health Organization (WHO) estimates that CVDs account for nearly 17.9 million deaths each year, representing approximately 31% of global deaths. Risk factors for CVDs include high blood pressure, elevated cholesterol levels, smoking, physical inactivity, and poor dietary habits (Roth, Mensah, & Fuster, 2020) [37].

Diabetes is another major NCD, affecting over 422 million people globally. High blood sugar levels characterize it and, if left unmanaged, can lead to severe complications such as kidney failure, blindness, and limb amputations. Type 2 diabetes, which is largely preventable, accounts for the vast majority of diabetes cases. The increasing rates of obesity and sedentary lifestyles have contributed to the global surge in diabetes prevalence (Rao, 2020) [35].

Cancers are a diverse group of diseases that can affect nearly any part of the body. According to WHO, cancer accounted for 9.6 million deaths in 2018, with lung, colorectal, stomach, liver, and breast cancers being the most common types. The causes of cancer are multifaceted, involving genetic, environmental, and lifestyle factors. Early detection is critical in improving survival rates, yet many cases are diagnosed at late stages due to a lack of access to screening and treatment (Ngoma & Ngoma, 2021) [24].

Chronic respiratory diseases, such as chronic obstructive pulmonary disease (COPD) and asthma, are prominent NCDs. COPD, often caused by long-term exposure to harmful substances like tobacco smoke, is responsible for approximately 3 million deaths annually. Asthma, though not as deadly, significantly impairs quality of life and affects over 262 million people worldwide. These diseases considerably burden healthcare systems and economies, particularly in low- and middle-income countries, where resources for management and treatment are often limited (Gould *et al*., 2023) [11].

2.2 Risk Factors
The development of NCDs is strongly linked to a set of modifiable risk factors, which, if addressed, could prevent a significant proportion of these diseases. The most prominent risk factors include tobacco use, unhealthy diets, physical

inactivity, and excessive alcohol consumption. These behaviors contribute to metabolic and physiological changes in the body, such as elevated blood pressure, obesity, high blood glucose, and abnormal lipid profiles, which are precursors to many NCDs (Le Foll *et al*., 2022) [17].

Tobacco use is one of the most preventable causes of death worldwide, responsible for approximately 8 million deaths annually. Smoking is a major risk factor for CVDs, cancers (particularly lung cancer), and respiratory diseases like COPD. Unhealthy diets, often characterized by high intake of processed foods, sugars, and trans fats, contribute to obesity and increase the risk of diabetes, CVDs, and certain cancers. Physical inactivity, driven by increasingly sedentary lifestyles, further exacerbates the risk of developing NCDs, as regular exercise is critical for maintaining healthy body weight, regulating blood pressure, and improving cardiovascular health (He *et al*., 2022) [14].

Excessive alcohol consumption also plays a significant role in the development of NCDs. Alcohol is a risk factor for many diseases, including liver cirrhosis, certain cancers, and hypertension. Moreover, the harmful use of alcohol contributes to injuries and accidents, adding to the overall disease burden. In addition to these behavioral risk factors, environmental factors, such as air pollution, have been increasingly recognized as contributors to the rise of NCDs, particularly respiratory diseases like asthma and COPD (Åberg, Byrne, Pirola, Männistö, & Sookoian, 2023) [2].

Addressing these risk factors requires a multifaceted approach that involves public health campaigns to promote healthy lifestyles, policy interventions to regulate the availability and marketing of harmful products, and efforts to create supportive environments for physical activity. Early detection and prevention efforts are essential in reducing the incidence of NCDs. Often aided by health data analytics and AI, predictive measures can identify individuals at high risk of developing NCDs, allowing for targeted interventions before the diseases fully manifest. This shift toward prevention rather than treatment is critical in managing the growing NCD burden.

2.3 Current Challenges

Despite advances in healthcare, several challenges hinder effective NCD management. One of the most significant challenges is the delayed diagnosis of NCDs. Many NCDs, such as cancer and diabetes, are often asymptomatic in their early stages, leading to diagnosis only after significant progression. Late-stage diagnoses complicate treatment options and reduce the chances of positive outcomes. In resource-limited settings, where access to healthcare services is restricted, the lack of routine screening further exacerbates this issue. This highlights the urgent need for more accessible and affordable diagnostic tools to detect NCDs at earlier stages (McGarvey, Gitlin, Fadli, & Chung, 2022) [21].

Another challenge is the lack of personalized care in managing NCDs. While traditional treatment approaches rely on standardized guidelines, these may not be effective for all patients, as individual responses to treatment can vary widely based on genetic, environmental, and lifestyle factors. Personalized medicine, which tailors interventions to an individual's unique biological and clinical profile, holds promise for improving NCD outcomes, but its widespread implementation remains limited due to the complexity and costs involved in gathering and analyzing large amounts of patient data (Scherrenberg *et al*., 2023) [38].

Health inequities also pose a substantial barrier to reducing the global NCD burden. Low- and middle-income countries, which bear the brunt of NCD-related morbidity and mortality, often lack the healthcare infrastructure and resources needed for effective prevention, diagnosis, and treatment. In these settings, access to essential medicines and technologies for NCD management is often limited, leading to poorer outcomes. Addressing these disparities requires global collaboration and investment in strengthening healthcare systems, ensuring equitable access to care, and building NCD prevention and management capacity in underserved regions (Jakovljevic, Westerman, Sharma, & Lamnisos, 2021) [15].

The rapid rise of NCDs also financially strains health systems and economies. As people live longer with chronic conditions, the demand for long-term care and ongoing management increases, driving up healthcare costs. This is particularly challenging in countries with limited financial resources, where the growing NCD burden threatens to overwhelm already stretched healthcare systems. Governments and policymakers must prioritize NCDs in health agendas, investing in prevention and early detection programs, and integrating AI and data science approaches to streamline healthcare delivery and improve outcomes (Haque *et al*., 2020) [12].

3. AI and Data Science Approaches in Health Data Analytics

Artificial intelligence (AI) and data science have revolutionized health data analytics, offering powerful tools for predicting, diagnosing, and managing non-communicable diseases (NCDs). These approaches enable the analysis of vast amounts of health data, including medical records, genomic information, and real-time patient data, to uncover trends, patterns, and insights that can improve patient outcomes. AI and data science techniques such as machine learning (ML), natural language processing (NLP), big data technologies, and personalized medicine are critical in the fight against NCDs by enhancing predictive accuracy, enabling early detection, and optimizing treatments.

3.1 Machine Learning (ML) and Predictive Analytics

Machine learning (ML) is a subset of AI that involves the development of algorithms capable of learning from data without being explicitly programmed. ML models have become a cornerstone of predictive analytics in healthcare, especially for NCDs. These models analyze large datasets to identify patterns and correlations between various health factors—such as genetics, lifestyle choices, and environmental exposures—and the likelihood of developing specific NCDs (Abdalrada, Abawajy, Al-Quraishi, & Islam, 2022) [1].

One of the most significant applications of ML in combating NCDs is predicting the onset of diseases such as diabetes, cardiovascular conditions, and cancer. For example, ML algorithms can be trained on datasets comprising patient demographics, medical histories, and lifestyle information to predict the risk of developing type 2 diabetes. By identifying at-risk individuals early, healthcare providers can implement preventative strategies, such as lifestyle interventions or medication, to delay or prevent disease onset (Wang & Wang, 2020) [45].

ML is also used to predict the progression of NCDs. For cardiovascular diseases, for instance, predictive models can forecast disease progression by analyzing cholesterol levels, blood pressure, and lifestyle habits. Such models are invaluable for informing clinical decisions and determining the appropriate interventions to slow or stop disease progression (Abdalrada *et al*., 2022) [1].

In addition, ML models help predict patient outcomes, including the likelihood of treatment success or the risk of complications. In oncology, machine learning algorithms can analyzetumor characteristics and patient profiles to predict responses to specific treatments, such as chemotherapy or immunotherapy. This helps clinicians make informed treatment decisions and reduces the risk of adverse effects by avoiding ineffective treatments. ML and predictive analytics offer a proactive approach to NCD management by facilitating early detection, improving prognosis, and supporting personalized care (Cuocolo, Caruso, Perillo, Ugga, & Petretta, 2020) [7].

3.2 Natural Language Processing (NLP)

Natural language processing (NLP) is another AI technique that is transforming health data analytics, particularly in the analysis of unstructured data. Much of the valuable health information is stored in unstructured formats, such as electronic health records (EHRs), clinical notes, discharge summaries, and research papers. This unstructured data contains critical insights that, when effectively analyzed, can provide a more comprehensive understanding of patient health and disease trends(Li *et al*., 2022) [18].

NLP enables the extraction of meaningful information from these unstructured text sources, converting them into structured data that can be used for further analysis. For example, NLP algorithms can be applied to clinical notes to identify symptoms, diagnoses, and treatment outcomes that might not be explicitly coded in structured medical records. This is particularly useful in detecting early signs of NCDs that may not be captured in standard diagnostic procedures. In oncology, NLP tools can analyze pathology reports to identify tumor characteristics and predict cancer progression (Kaswan, Gaur, Dhatterwal, & Kumar, 2021) [16].

NLP is also widely used in medical research to analyze vast scientific literature. With the constant influx of new research on NCDs, clinicians and researchers cannotkeep up with the latest findings manually. NLP-based tools can automatically scan and summarize relevant studies, enabling researchers to stay updated on emerging trends and treatment options. This accelerates the translation of research findings into clinical practice, ultimately benefiting patients (Nazir *et al*., 2020) [23].

By unlocking the value of unstructured data, NLP enhances the ability to make data-driven decisions in NCD management. It enables the synthesis of information from various sources and improves the accuracy of predictive models by incorporating a broader range of data, including patient-reported outcomes and real-time health updates.

3.3 Big Data and Real-Time Monitoring

Big data technologies are integral to the modern healthcare landscape, particularly for NCD management, where large and diverse datasets need to be analyzed in real time. Big data encompasses vast amounts of information from various sources, including wearable devices, mobile health apps, EHRs, genomic data, and population health databases. The ability to process and analyze such massive datasets in real-time provides a more comprehensive view of patient health and allows for continuous monitoring of disease progression (Okoduwa *et al*., 2024; Udegbe, Ebulue, Ebulue, & Ekesiobi, 2024b) [27, 43].

One of the most prominent applications of big data in NCDs is the use of wearable devices, such as fitness trackers and smartwatches, which continuously collect data on physical activity, heart rate, sleep patterns, and other health metrics. These devices enable real-time monitoring of patients with chronic conditions, allowing healthcare providers to track disease progression and intervene when necessary. For example, in patients with cardiovascular disease, wearable devices can detect irregular heart rhythms or abnormal blood pressure readings, prompting early medical intervention and potentially preventing life-threatening events like heart attacks (Pelluru, 2020) [32].

Electronic health records (EHRs) also contribute to the big data landscape by providing a wealth of clinical information on patient histories, treatments, and outcomes. When integrated with AI algorithms, EHR data can be used to predict disease trajectories, identify high-risk patients, and personalize treatment plans. Furthermore, big data analytics allows for population-level analyses, identifying trends in disease prevalence, treatment efficacy, and healthcare access disparities, which can inform public health strategies for NCD prevention and management (Reza, Prieto, & Julien, 2020) [36]. Real-time data monitoring, facilitated by big data technologies, is transforming NCD management by enabling continuous, personalized care. It shifts the focus from episodic, reactive healthcare to a proactive approach where patients and clinicians can access real-time insights, improving individual and population health outcomes (Odilibe *et al*., 2024) [25].

3.4 Personalized Medicine and Treatment Optimization

Personalized medicine, also known as precision medicine, is an approach that tailors medical treatment to the individual characteristics of each patient, taking into account their genetic makeup, environmental exposures, and lifestyle factors. AI and data science are crucial in enabling personalized medicine, particularly in the treatment of NCDs, where patients often respond differently to the same treatment regimens.

AI algorithms can analyze genomic data to identify genetic markers associated with specific NCDs, such as cancer or cardiovascular disease. This information allows for the development of personalized treatment plans that target the underlying biological mechanisms driving the disease. In oncology, for example, genetic profiling of tumors enables the identification of mutations that specific drugs, such as targeted therapies or immunotherapies can target. This approach not only improves treatment efficacy but also reduces the risk of side effects by avoiding treatments that are unlikely to be effective (Strianese *et al*., 2020) [40].

Moreover, AI helps optimize treatment regimens by continuously analyzing patient data to adjust treatment based on the patient's response. For patients with diabetes, AI-powered systems can analyze blood sugar levels, dietary habits, and exercise routines to provide personalized recommendations on insulin dosages and lifestyle modifications. This real-time, data-driven approach ensures that treatment is always optimized for the individual's unique needs, improving disease management and quality of life (Ellahham, 2020) [8]. Personalized medicine supported by AI and data science represents a shift from a one-size-fits-all approach to a more individualized care model. It holds the potential to significantly improve outcomes for patients with NCDs by offering more effective, targeted treatments and minimizing the risks associated with trial-and-error approaches to care (Udegbe, Ebulue, Ebulue, & Ekesiobi, 2024a) [42].

4. Challenges and Ethical Considerations

4.1 Data Privacy and Security

One of the most pressing ethical concerns in the use of AI for health data analytics is the issue of data privacy and security. Health data is highly sensitive, containing personal information about an individual's medical history, genetic makeup, and lifestyle. The use of large-scale datasets in AI-driven healthcare systems raises questions about how this data is collected, stored, and used. Ensuring that patient data is kept confidential and secure is paramount to maintaining trust in these technologies and the healthcare systems that employ them.

Data privacy regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in Europe, provide frameworks for protecting patient data. These regulations mandate that organizations handling health data must obtain informed consent from individuals before using their data, ensure that data is anonymized where possible, and implement robust security measures to prevent unauthorized access. However, in practice, ensuring compliance with these regulations can be challenging, especially when dealing with vast amounts of data collected from multiple sources, such as electronic health records (EHRs), wearable devices, and mobile health apps (Bakare, Adeniyi, Akpuokwe, & Eneh, 2024) [4].

The use of AI systems in healthcare also presents new challenges for data security. AI algorithms require large volumes of data to train predictive models, and this data is often shared across different platforms and organizations, increasing the risk of data breaches. In recent years, healthcare institutions have become prime targets for cyberattacks, as they hold large repositories of sensitive data. A breach in a health data system could expose patients to identity theft, discrimination, and other harms. Therefore, ensuring robust cybersecurity measures is essential when implementing AI systems in healthcare (Martinez-Martin *et al*., 2021) [20].

In addition to security concerns, there are ethical questions surrounding informed consent in the context of AI. Patients may not fully understand how their data is being used, especially in cases where AI models analyze data for secondary purposes beyond their immediate care, such as research or public health monitoring. Ensuring that patients are adequately informed about how their data will be used, and that they have the option to opt out without compromising their care is critical in addressing these ethical concerns (Naik *et al*., 2022) [22].

4.2 Bias and Fairness in AI Models

Another major ethical concern in using AI for NCD management is the potential for bias in predictive models. AI systems are trained on historical data, and if that data contains biases—such as underrepresentation of certain populations or unequal treatment outcomes—the AI models may perpetuate or even amplify these biases. This can lead to unequal access to care or disparities in treatment recommendations, particularly for underrepresented or marginalized groups (Ferrara, 2023) [9]. For example, studies have shown that AI models trained on predominantly white patient populations may be less accurate in predicting outcomes for individuals from minority racial or ethnic groups. Similarly, gender bias can emerge in predictive models if the training data does not adequately represent women, leading to less accurate diagnoses or treatment plans for female patients. These biases can have serious consequences, particularly in the management of NCDs, where early detection and tailored interventions are critical to preventing disease progression (Panesar & Panesar, 2021) [30].

Addressing bias in AI models requires a concerted effort to ensure that the data used to train these systems is representative of diverse populations. This involves collecting data from a wide range of demographic groups, including individuals of different ages, genders, ethnicities, and socioeconomic backgrounds. In addition, algorithms need to be continuously monitored and tested to identify and mitigate potential biases. Researchers and developers must also be transparent about the limitations of AI models and avoid overreliance on these tools without proper clinical oversight (Schwartz *et al*., 2022) [39].

Fairness in AI models goes beyond just addressing bias in training data. It also involves ensuring that the benefits of AI technologies are distributed equitably across all populations. This includes providing equal access to AI-driven healthcare tools and ensuring that underserved communities are not left behind in the digital health revolution. Public health systems must work to bridge the digital divide and ensure that all individuals, regardless of their background, have access to the benefits of AI in healthcare (Venkatasubbu & Krishnamoorthy, 2022) [44].

4.3 Technical Barriers

While AI and data science hold great promise for healthcare, several technical barriers limit their widespread implementation. One of the key challenges is data quality. AI models rely on high-quality, accurate data to make reliable predictions, but in practice, health data is often incomplete, inconsistent, or fragmented. For example, EHRs may contain missing information, duplicate records, or errors in data entry. This can reduce the accuracy of predictive models and lead to incorrect diagnoses or treatment recommendations. Data quality is essential to ensure that AI systems can deliver accurate and reliable insights (Ajegbile, Olaboye, Maha, Igwama, & Abdul, 2024; Ogugua, Okongwu, Akomolafe, Anyanwu, & Daraojimba, 2024; Olorunyomi, Sanyaolu, Adeleke, & Okeke, 2024) [3, 26, 28].

Another technical challenge is interoperability. Health data is often siloed in different systems, making it difficult to integrate data from multiple sources for comprehensive analysis. EHR systems, wearable devices, and mobile health apps often use different data formats and standards, creating data sharing and analysis barriers. For AI systems to be effective, they must be able to access and analyze data from a wide range of sources, but the lack of interoperability between healthcare systems remains a significant obstacle. Initiatives to standardize health data formats and improve data-sharing protocols are crucial for overcoming this challenge (Petch, Di, & Nelson, 2022) [33].

Scalability is another issue that healthcare systems must address when implementing AI solutions. AI systems that work well in research settings or pilot projects may face difficulties when deployed at scale in real-world healthcare environments. This can be due to a variety of factors, including the need for significant computational resources, the complexity of integrating AI into existing workflows, and the challenges of maintaining system performance as the volume of data grows. Developing AI systems that are scalable, cost-effective, and easy to implement in diverse healthcare settings is essential for realizing the full potential of AI in healthcare (Oyeniran, Adewusi, Adeleke, Akwawa, & Azubuko, 2022) [29].

Finally, there is a need for greater transparency and explainability in AI systems. Many AI models, particularly those based on machine learning, function as "black boxes" that generate predictions without clearly explaining how those predictions were made. This lack of transparency can make it difficult for clinicians to trust AI-driven recommendations, particularly in critical care settings where decisions have life-or-death consequences. Ensuring that AI systems are explainable and that clinicians understand how the algorithms work is essential for fostering trust in these technologies and ensuring that they are used effectively in healthcare (Hassija *et al*., 2024) [13].

5. Conclusion and Future Directions

5.1 Conclusion

One of the key contributions of AI in NCD management is its ability to analyze large datasets to predict the onset and progression of diseases. Machine learning (ML) models, for example, are highly effective in identifying at-risk individuals based on their health data and predicting disease outcomes, which enables timely interventions and improves patient care. AI is also crucial for optimizing treatments through personalized medicine, where individual patient profiles guide clinical decisions, enhancing treatment efficacy while minimizing risks.

Another significant application of AI is in natural language processing (NLP), which helps convert unstructured data from clinical notes and medical *records* into structured information that can be analyzed to provide deeper insights into patient health. Furthermore, big data technologies allow for real-time monitoring of patients through continuous data collection from wearable devices, enabling proactive disease management. Together, these AI and data science approaches contribute to a more effective and efficient healthcare system, reducing the burden of NCDs on individuals and society.

Despite these advancements, challenges persist, particularly around data privacy and security, bias in AI models, and the technical difficulties associated with data interoperability and system scalability. These barriers must be addressed to ensure that AI can deliver on its promise of improving NCD management across all populations.

5.2 Future Directions

Several recommendations can be made to enhance the application of AI and data science in NCD research and management. First, healthcare providers, technologists, and policymakers must work together to establish comprehensive data privacy frameworks that protect patient information while enabling the sharing of health data for research and treatment purposes. This requires the development of stronger regulations that promote transparency in data usage and ensure patients' informed consent.

Additionally, efforts must be made to mitigate bias in AI models by ensuring that datasets used for training are representative of diverse populations. This can be achieved through targeted data collection efforts that include underrepresented groups and by incorporating fairness auditing practices into developing and deploying AI systems.

Moreover, improving technical infrastructure is critical to scaling AI systems in healthcare. Enhancing data quality and ensuring interoperability between different health information systems will facilitate the integration of AI into clinical workflows. Investment in these areas will enable healthcare systems to leverage AI at a broader scale, leading to better disease management and patient outcomes.

Lastly, fostering collaboration between healthcare providers, technologists, and regulators is essential for shaping the future of AI in healthcare. Policymakers must create environments that encourage innovation while ensuring that the ethical use of AI remains a priority. Through continued research, innovation, and interdisciplinary cooperation, AI and data science can become integral tools in the global fight against NCDs, ultimately reducing their prevalence and impact on public health.

6. References

1. Abdalrada AS, Abawajy J, Al-Quraishi T, Islam SMS. Machine learning models for prediction of co-occurrence of diabetes and cardiovascular diseases: A retrospective cohort study. Journal of Diabetes & Metabolic Disorders. 2022; 21(1):251-261.
2. Åberg F, Byrne CD, Pirola CJ, Männistö V, Sookoian S. Alcohol consumption and metabolic syndrome: Clinical and epidemiological impact on liver disease. Journal of hepatology. 2023; 78(1):191-206.
3. Ajegbile MD, Olaboye JA, Maha CC, Igwama GT, Abdul S. Health Informatics and Public Health: Leveraging Technology for Improved Health Outcomes. Health. 2024; 20(8):229-235.
4. Bakare SS, Adeniyi AO, Akpuokwe CU, Eneh NE. Data privacy laws and compliance: A comparative review of the EU GDPR and USA regulations. Computer Science & IT Research Journal. 2024; 5(3):528-543.
5. Bawack RE, Fosso Wamba S, Carillo KDA. A framework for understanding artificial intelligence research: Insights from practice. Journal of Enterprise Information Management. 2021; 34(2):645-678.
6. Coates MM, Kintu A, Gupta N, Wroe EB, Adler AJ, Kwan GF, *et al*. Burden of non-communicable diseases from infectious causes in 2017: A modelling study. The Lancet Global Health. 2020; 8(12):e1489-e1498.
7. Cuocolo R, Caruso M, Perillo T, Ugga L, Petretta M. Machine learning in oncology: A clinical appraisal. Cancer letters. 2020; 481:55-62.
8. Ellahham S. Artificial intelligence: The future for diabetes care. The American Journal of Medicine. 2020; 133(8):895-900.
9. Ferrara E. Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation

strategies. Sci. 2023; 6(1):3.
10. Ganju A, Goulart AC, Ray A, Majumdar A, Jeffers BW, Llamosa G, *et al*. Systemic solutions for addressing non-communicable diseases in low-and middle-income countries. Journal of Multidisciplinary Healthcare, 2020, 693-707.
11. Gould GS, Hurst JR, Trofor A, Alison JA, Fox G, Kulkarni MM, *et al*. Recognising the importance of chronic lung disease: A consensus statement from the Global Alliance for Chronic Diseases (Lung Diseases group). Respiratory Research. 2023; 24(1):15.
12. Haque M, Islam T, Rahman NAA, McKimm J, Abdullah A, Dhingra S. Strengthening primary health-care services to help prevent and control long-term (chronic) non-communicable diseases in low-and middle-income countries. Risk Management and Healthcare Policy, 2020, 409-426.
13. Hassija V, Chamola V, Mahapatra A, Singal A, Goel D, Huang K, *et al*. Interpreting black-box models: A review on explainable artificial intelligence. Cognitive Computation. 2024; 16(1):45-74.
14. He H, Pan Z, Wu J, Hu C, Bai L, Lyu J. Health effects of tobacco at the global, regional, and national levels: Results from the 2019 global burden of disease study. Nicotine and Tobacco Research. 2022; 24(6):864-870.
15. Jakovljevic M, Westerman R, Sharma T, Lamnisos D. Aging and global health. Handbook of global health, 2021, 73-102.
16. Kaswan KS, Gaur L, Dhatterwal JS, Kumar R. AI-based natural language processing for the generation of meaningful information electronic health record (EHR) data. In Advanced AI techniques and applications in bioinformatics (pp. 41-86): CRC Press, 2021.
17. Le Foll B, Piper ME, Fowler CD, Tonstad S, Bierut L, Lu L, *et al*. Tobacco and nicotine use. Nature Reviews Disease Primers. 2022; 8(1):19.
18. Li I, Pan J, Goldwasser J, Verma N, Wong WP, Nuzumlalı MY, *et al*. Neural natural language processing for unstructured data in electronic health records: A review. Computer Science Review. 2022; 46:100511.
19. Luna F, Luyckx VA. Why have non-communicable diseases been left behind? Asian bioethics review. 2020; 12(1):5-25.
20. Martinez-Martin N, Luo Z, Kaushal A, Adeli E, Haque A, Kelly SS, *et al*. Ethical issues in using ambient intelligence in health-care settings. The Lancet Digital Health. 2021; 3(2):e115-e123.
21. McGarvey N, Gitlin M, Fadli E, Chung KC. Increased healthcare costs by later stage cancer diagnosis. BMC Health Services Research. 2022; 22(1):1155.
22. Naik N, Hameed BZ, Shetty DK, Swain D, Shah M, Paul R, *et al*. Legal and ethical consideration in artificial intelligence in healthcare: Who takes responsibility? Frontiers in surgery. 2022; 9:862322.
23. Nazir S, Khan S, Khan HU, Ali S, Garcia-Magarino I, Atan RB, *et al*. A comprehensive analysis of healthcare big data management, analytics and scientific programming. IEEE Access. 2020; 8:95714-95733.
24. Ngoma TA, Ngoma M. Global Burden of Cancer: Prevalence, Pattern, and Trends. In Handbook of global health (pp. 459-494): Springer, 2021.
25. Odilibe IP, Akomolafe O, Arowoogun JO, Anyanwu EC, Onwumere C, Ogugua JO. Mental health policies: A comparative review between the USA and African nations. International Medical Science Research Journal. 2024; 4(2):141-157.
26. Ogugua JO, Okongwu CC, Akomolafe OO, Anyanwu EC, Daraojimba OD. Mental health and digital technology: A public health review of current trends and responses. International Medical Science Research Journal. 2024; 4(2):108-125.
27. Okoduwa IO, Ashiwaju BI, Ogugua JO, Arowoogun JO, Awonuga KF, Anyanwu EC. Reviewing the progress of cancer research in the USA. World Journal of Biology Pharmacy and Health Sciences. 2024; 17(2):068-079.
28. Olorunyomi TD, Sanyaolu TO, Adeleke AG, Okeke IC. Integrating FinOps in healthcare for optimized financial efficiency and enhanced care, 2024.
29. Oyeniran C, Adewusi AO, Adeleke AG, Akwawa LA, Azubuko CF. Ethical AI: Addressing bias in machine learning models and software applications. Computer Science & IT Research Journal. 2022; 3(3):115-126.
30. Panesar A, Panesar H. Artificial intelligence and machine learning in global healthcare. In Handbook of global health (pp. 1775-1813): Springer, 2021.
31. Paramesha M, Rane NL, Rane J. Big data analytics, artificial intelligence, machine learning, internet of things, and blockchain for enhanced business intelligence. Partners Universal Multidisciplinary Research Journal. 2024; 1(2):110-133.
32. Pelluru K. Prospects and Challenges of Big Data Analytics in Medical Science. Journal of Innovative Technologies. 2020; 3(1):1-18-11-18.
33. Petch J, Di S, Nelson W. Opening the black box: The promise and limitations of explainable machine learning in cardiology. Canadian Journal of Cardiology. 2022; 38(2):204-213.
34. Ramesh S, Kosalram K. The burden of non-communicable diseases: A scoping review focus on the context of India. Journal of education and health promotion. 2023; 12(1):41.
35. Rao GH. Management of diabetes epidemic: Global perspective. Current Trends in Diabetes. 2020; 25.
36. Reza F, Prieto JT, Julien SP. Electronic health records: Origination, adoption, and progression. Public Health Informatics and Information Systems, 2020, 183-201.
37. Roth GA, Mensah GA, Fuster V. The global burden of cardiovascular diseases and risks: A compass for global action. In (Vol. 76, pp. 2980-2981): American College of Cardiology Foundation Washington DC, 2020.
38. Scherrenberg M, Marinus N, Giallauria F, Falter M, Kemps H, Wilhelm M, *et al*. The need for long-term personalized management of frail CVD patients by rehabilitation and telemonitoring: A framework. Trends in cardiovascular medicine. 2023; 33(5):283-297.
39. Schwartz R, Schwartz R, Vassilev A, Greene K, Perine L, Burt A, *et al*. Towards a standard for identifying and managing bias in artificial intelligence (Vol. 3): US Department of Commerce, National Institute of Standards and Technology, 2022.
40. Strianese O, Rizzo F, Ciccarelli M, Galasso G, D'Agostino Y, Salvati A, *et al*. Precision and personalized medicine: How genomic approach improves the management of cardiovascular and neurodegenerative disease. Genes. 2020; 11(7):747.
41. Tahir F, Khan M. Big Data: The Fuel for Machine

Learning and AI Advancement (2516-2314), 2023. Retrieved from:
42. Udegbe FC, Ebulue OR, Ebulue CC, Ekesiobi CS. AI's impact on personalized medicine: Tailoring treatments for improved health outcomes. Engineering Science & Technology Journal. 2024a; 5(4):1386-1394.
43. Udegbe FC, Ebulue OR, Ebulue CC, Ekesiobi CS. The role of artificial intelligence in healthcare: A systematic review of applications and challenges. International Medical Science Research Journal. 2024b; 4(4):500-508.
44. Venkatasubbu S, Krishnamoorthy G. Ethical Considerations in AI Addressing Bias and Fairness in Machine Learning Models. Journal of Knowledge Learning and Science Technology. 2022; 1(1):130-138. ISSN: 2959-6386 (online).
45. Wang Y, Wang J. Modelling and prediction of global non-communicable diseases. BMC public health. 2020; 20:1-13.