future internet

Review

Edge AI for Early Detection of Chronic Diseases and the Spread of Infectious Diseases: Opportunities, Challenges, and Future Directions

Elarbi Badidi

Department of Computer Science and Software Engineering, College of Information Technology, UAE University, Al-Ain P.O. Box 15551, United Arab Emirates; ebadidi@uaeu.ac.ae

Abstract: Edge AI, an interdisciplinary technology that enables distributed intelligence with edge devices, is quickly becoming a critical component in early health prediction. Edge AI encompasses data analytics and artificial intelligence (AI) using machine learning, deep learning, and federated learning models deployed and executed at the edge of the network, far from centralized data centers. AI enables the careful analysis of large datasets derived from multiple sources, including electronic health records, wearable devices, and demographic information, making it possible to identify intricate patterns and predict a person's future health. Federated learning, a novel approach in AI, further enhances this prediction by enabling collaborative training of AI models on distributed edge devices while maintaining privacy. Using edge computing, data can be processed and analyzed locally, reducing latency and enabling instant decision making. This article reviews the role of Edge AI in early health prediction and highlights its potential to improve public health. Topics covered include the use of AI algorithms for early detection of chronic diseases such as diabetes and cancer and the use of edge computing in wearable devices to detect the spread of infectious diseases. In addition to discussing the challenges and limitations of Edge AI in early health prediction, this article emphasizes future research directions to address these concerns and the integration with existing healthcare systems and explore the full potential of these technologies in improving public health.

Keywords: artificial intelligence; edge computing; early health prediction; federated learning; wearable devices; chronic diseases; data privacy; public health; healthcare informatics; data analysis

check for updates

Citation: Badidi, E. Edge AI for Early Detection of Chronic Diseases and the Spread of Infectious Diseases: Opportunities, Challenges, and Future Directions. *Future Internet* **2023**, *15*, 370. https://doi.org/10.3390/fi15110370

Academic Editor: Hamid Mcheick

Received: 18 October 2023
Revised: 7 November 2023
Accepted: 10 November 2023
Published: 18 November 2023

1. Introduction

Digital technologies have led to an exponential increase in the amount of data generated by people and devices. In healthcare, this has led to new perspectives on the timely detection and interception of diseases, improving health outcomes and reducing overall healthcare spending. The use of AI and edge computing in healthcare is one of the most promising avenues of advancement, as it can revolutionize the way healthcare providers investigate and predict disease outbreaks. Healthcare data come from various sources, such as electronic health records, wearable devices, and various sensors. AI algorithms can be used to analyze these data to identify patterns and trends that may indicate disease onset [1,2]. By combining these data with other relevant information, such as demographic and environmental factors, AI can be used to make accurate predictions about future health outcomes and develop targeted interventions to prevent disease outbreaks [3,4]. Conversely, the emerging paradigm of edge computing seeks to solve the problems associated with transmitting and processing vast amounts of data as they are generated. Edge computing can significantly reduce latency and increase processing speeds by bringing computing power close to the data source. It is an optimal solution for systems that require real-time data processing and decision making [5,6]. In healthcare, edge computing is particularly useful in applications that require real-time monitoring and decision making, such as wearable devices for patient monitoring and remote diagnosis and treatment [7,8].

By combining the strengths of AI and edge computing, healthcare providers can create innovative solutions that improve public health outcomes, lower healthcare costs, and improve the overall patient experience [9,10].

Early health prediction plays a critical role in improving public health by enabling healthcare providers to identify and address potential health problems before they become more serious [11,12]. Using a variety of data sources, including electronic health records, demographic data, and wearable devices, the field of early health prediction manifests as a multifaceted approach that provides a holistic perspective on a person's well-being and proactively anticipates impending health prognoses. This information can then be used to develop personalized interventions, treatments, and preventive measures to improve an individual's health and prevent the onset of chronic diseases [13]. One of the most important benefits of early health prediction is the early detection of chronic diseases such as diabetes, cardiovascular disease, and cancer [14–16]. By detecting these diseases in their early stages, physicians can take therapeutic measures and recommend lifestyle changes that can improve an individual's health while preventing the disease's relentless progression. It is worth emphasizing that early health prediction promises to reduce healthcare costs by circumventing the need for later exhaustive and financially burdensome treatments. Early health prediction has great potential for infectious disease prevention and control. Careful examination of data obtained from wearable devices and various sources makes it possible to identify variations in a person's activity patterns and heart rate that serve as potential precursors to the onset of an infectious disease. With this knowledge, proactive measures can be taken to limit the transmission of the disease and strategic interventions can be made to mitigate its effects.

This article explores the role of Edge AI in early health prediction and highlights its benefits, limitations, and real-world applications in healthcare. It discusses the current state of the field and the challenges to be overcome. It also explores the future potential of Edge AI for improving public health. Within the context of Edge AI's flourishing significance in healthcare, it is essential to establish a comprehensive understanding of its applications and implications. Several prior review papers have explored facets of artificial intelligence (AI) in healthcare, with a predominant focus on specific diseases, imaging modalities, or AI techniques. Herein, we delineate how our work distinguishes itself from these review papers by taking a broader and more inclusive perspective. A paper by Xu et al. [17] offers an extensive survey of deep reinforcement learning (DRL) in the context of medical imaging and radiation therapy. It covers DRL's basic concepts and algorithms and discusses its applications in lesion localization, classification, registration, segmentation, and treatment planning. Montagnon et al. [18] provide a primer on the steps involved in developing and deploying deep learning models primarily tailored to radiology applications. It covers data collection and preprocessing, model design and training, model evaluation and validation, model deployment and integration, and model monitoring and maintenance. Furthermore, Cao et al. [19] reviewed deep learning principles and their applications in various biomedical domains, such as genomics, proteomics, metabolomics, microbiomics, medical imaging, drug discovery, and precision medicine. They also discussed the challenges and future directions of deep learning in biomedicine. Federated machine learning and its use in disease prediction was covered by Moshawrab et al. [20], who offered a comprehensive review of federated machine learning (FML) within the healthcare context. The paper also surveys the recent applications of FML in various disease prediction tasks, such as cancer, cardiovascular disease, and diabetes prediction. In another work, Moshawrab et al. [21] conduct a systematic literature review of the state-of-the-art smart wearables that can detect and monitor cardiovascular diseases (CVDs), the leading cause of death worldwide. They focus on the types, features, and applications of smart wearables, such as smartwatches, smart bands, smart rings, and smart glasses, that can measure various physiological signals, such as electrocardiogram (ECG), photoplethysmogram (PPG), blood pressure (BP), and heart rate (HR) signals.

In summary, our work stands out through its exploration of Edge AI applications in healthcare. It accentuates the potential impact of Edge AI on public health, early prediction of chronic diseases, and its seamless integration within established healthcare systems. Although the aforementioned review papers delve into specific niches within healthcare and AI, our work serves to provide a holistic understanding of the implications of Edge AI technologies in healthcare, emphasizing the necessity of a comprehensive grasp of Edge AI's role in healthcare settings. Our contributions can be summarized as follows:

1. We describe the synergy between AI and edge computing in healthcare to create opportunities for healthcare providers. These opportunities include cutting-edge tools that provide instant data and deep insights. These tools help accelerate disease detection and create tailored treatment plans. The plans consider the unique characteristics and needs of individual patients.
2. We review the various applications of Edge AI in the early prediction of health issues and the detection of chronic and infectious diseases.
3. We scrutinize machine learning and deep learning models instrumental in early health prediction and detection of chronic and infectious diseases.
4. We review the various applications of federated learning for disease and mortality prediction and highlight the critical dimension of privacy in early health prediction.
5. We identify and elaborate on the challenges faced by Edge AI in the healthcare domain, such as privacy, data accuracy, model bias, interoperability, and integration with existing health systems. Furthermore, we outline for each challenge the future research directions, emphasizing the transformative potential of Edge AI in the context of early health prediction.

The following sections of this review article are organized as follows: Section 2 provides background information on Edge AI. Section 3 describes the synergy between AI and edge computing in healthcare. Section 4 describes the research methodology used in this review. Section 5 describes the stakeholders of an Edge AI-based system for early healthcare prediction, reviews the various applications of Edge AI in the early prediction of health issues and the detection of chronic and infectious diseases, and examines the machine learning and deep learning models instrumental in early health prediction and detection of chronic and infectious diseases. Section 6 describes the potential of federated learning for early health prediction and the training process of the federated learning model with edge devices. It also reviews the various applications of federated learning for disease and mortality prediction. Section 7 discusses the challenges and limitations of Edge AI in early health prediction and the future research directions to address each challenge. Finally, Section 8 concludes the article.

2. Edge AI Overview

The emergence of Edge AI, an interdisciplinary technology that enables distributed intelligence with edge devices, is attributed to data analytics, machine learning, and deep learning taking place at the edge of the network, far from centralized data centers. Several initiatives are underway to develop Edge AI-based solutions in various application areas, such as the digital industry, which uses the technology to prevent early failures and perform predictive maintenance. Moreover, the use of Edge AI goes beyond the boundaries of specific industries and finds practical applications in various areas that shape our daily lives. For example, Edge AI is helping to advance smart buildings and efficient smart grids in the energy sector. In the automotive and transportation industries, as well as in the development of smart cities, Edge AI also plays a crucial role in optimizing operations and increasing overall efficiency. In addition, Edge AI is finding application in areas such as health and wellness, precision agriculture, and numerous other fields, where its transformative potential is being harnessed to achieve remarkable results.

Artificial intelligence has become a fascinating discipline within computer science, concerned with the careful development of algorithms and systems designed to mimic the cognitive abilities of human intelligence. Its overarching goal is to develop machines that are capable of learning, reasoning, and making data-driven decisions, similar to the cognitive abilities of humans. AI is playing a central role in healthcare, unleashing its potential across a broad spectrum that includes disease diagnosis, treatment modalities, personalized medicine, and continuous monitoring of patients' well-being. On the other hand, edge computing is a distributed computing architecture that brings computing power closer to the data source [5,6,22]. In today's dynamic landscape, edge computing has emerged as a timely and common-sense solution to meet the increasing need for fast data processing and rapid decision making in a variety of domains. Edge computing is relevant in various fields, such as the IoT, industrial automation, and the ever-evolving healthcare industry. Unlike the traditional cloud computing model, which relies predominantly on centralized data centers, edge computing involves the strategic deployment of compact and efficient computing devices at the periphery of the network, right where the data originate. There are notable benefits to be gained from this approach, such as increased processing speed and reduced latency, as data no longer need to be transferred to a central location. In healthcare, edge computing is useful in wearable devices for patient monitoring, remote diagnosis, and treatment. In addition, this approach is a valuable privacy and security enhancement because it allows sensitive information to be processed locally. By reducing reliance on the transmission of such data to a central location, the risks associated with long-distance data transmission are mitigated, and the protection and confidentiality of critical information are strengthened. Edge computing is an example of an innovative paradigm that optimizes the handling of data in real time and empowers various sectors, especially healthcare, with its transformative potential. Figure 1 shows the architecture of edge computing for healthcare.

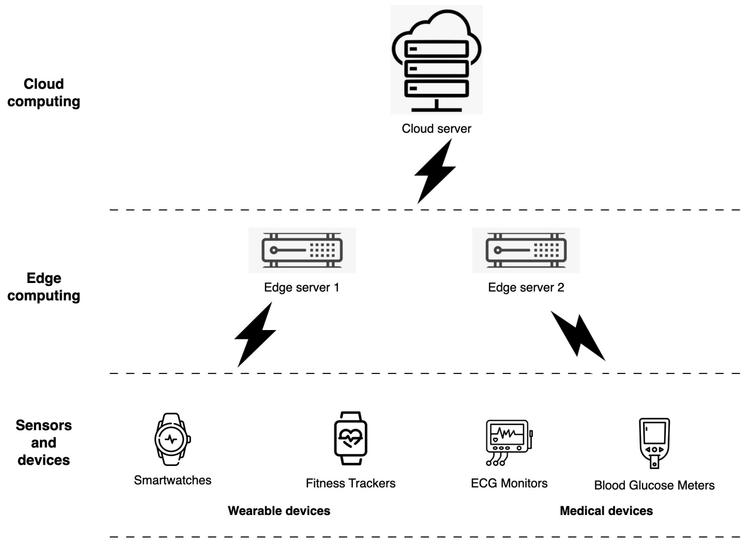

Figure 1. Edge computing architecture for healthcare.

A research report from Grand View Research [23] sheds light on the remarkable expansion of the global edge computing market. The report offers a compelling insight into the market dynamics, highlighting that the edge computing market size is estimated to reach a remarkable USD 11.24 billion by 2022, representing a significant economic scale. Moreover, the projected growth trajectory drives the market to an estimated value of USD 16.45 billion by 2023. Even further, the market is projected to grow at a remarkable compound annual growth rate (CAGR) of 37.9% from 2023 to 2030, resulting in a substantial market size of around USD 155.90 billion. The remarkable rise in the market is primarily due to the ever-increasing demand for real-time data processing and informed decision making across a wide range of industries. This insatiable appetite for instant data-driven insights spans multiple sectors, including but not limited to manufacturing, healthcare, and the flourishing landscape of smart cities. The research findings underscore the dominance of the Industrial Internet of Things segment in the edge computing industry, which accounts for over 29%

of the revenue share in 2022. Edge computing has played a significant role in enabling manufacturers to achieve the goal of digitizing their assets. A significant share of edge devices is installed in the manufacturing segment. The energy and utilities segment accounted for more than 14% of the revenue share in 2022. After the energy and utilities sector, the healthcare sector was a significant contributor to the revenue share in 2022, indicating its strong presence and influence. Moreover, this sector is projected to witness a CAGR of around 38% during the forecast period. In healthcare, edge computing holds immense transformative potential that will revolutionize early health prediction. By leveraging this cutting-edge technology, healthcare providers gain the ability to perform real-time analytics on large datasets from multiple channels seamlessly. An excellent example of this application is the effective use of wearable devices equipped with sophisticated sensors that enable the seamless collection of vital signs, physical activity measurements, and other key health indicators. These data points can then be seamlessly processed and analyzed, facilitating the timely detection of early signs of disease. This proactive approach enables healthcare providers to intervene quickly, prevent disease progression, improve health outcomes, and ultimately reduce overall healthcare spending.

Growing privacy concerns have led to the emergence of federated learning, an innovative method of machine learning that enables the training of models with distributed data across numerous devices without the need to aggregate the data in a central repository. This technique holds great promise, especially in healthcare, where it has significant potential for predicting and forecasting disease. By leveraging federated learning, physicians can harness the power of decentralized data while protecting individual privacy, ultimately advancing the field of predictive analytics in healthcare.

3. AI and Edge Computing Synergy in Healthcare

The advent of artificial intelligence has opened up numerous opportunities in healthcare, giving healthcare providers a set of cutting-edge tools capable of delivering instant data and deep insights. This technological integration acts as a catalyst, enabling healthcare professionals to accelerate the process of disease detection and reap the benefits of early intervention. In addition, the use of AI enables physicians to carefully craft tailored treatment plans that take into account the unique characteristics and needs of individual patients. The multiple impacts of AI extend beyond diagnostics and treatment, leading to tangible improvements in patient outcomes and the overall quality of healthcare.

AI in healthcare typically uses electronic health records to identify patients at risk for chronic diseases such as heart disease and diabetes [24,25]. Using algorithms, various factors such as age, genetics, lifestyle habits, and medical history are taken into account to create comprehensive risk profiles. This enables targeted interventions and early disease surveillance. Another example of healthcare enrichment is the application of AI in drug discovery and development. Using sophisticated AI algorithms, large amounts of data from multiple channels such as the scientific literature and clinical trials are subjected to careful analysis [26–28]. This comprehensive investigation not only facilitates the identification of novel drug targets but also enables the prediction of drug efficacy. Integrating AI in this context accelerates the drug discovery process while increasing the precision and effectiveness of drug discovery efforts. Moreover, the use of AI algorithms is proving to be valuable as it enhances the capabilities of healthcare providers in making informed and accurate diagnoses [29]. These algorithms help formulate personalized medical plans that seamlessly integrate important aspects such as the patient's genetic makeup, medical background, and lifestyle habits. By complexly combining this diverse information, medical professionals can create treatment plans tailored to each patient's unique circumstances, ultimately ensuring optimal efficacy and therapeutic outcomes.

Deep learning and reinforcement learning are two powerful AI techniques for medical data analysis [17,30]. Through rigorous exploration of large and complicated datasets, deep learning effectively leverages artificial neural network capabilities to acquire knowledge and make predictions. This versatile approach can be applied to various medical data analysis tasks, including image analysis, natural language processing, and time series analysis. It has been used to analyze medical images such as X-rays and MRI scans, revealing complex patterns that significantly affect diagnostic accuracy. In addition, the keen eye of deep learning algorithms sifts through electronic health records, unearthing valuable insights that pave the way for predicting patient outcomes [31,32]. Machine learning involves the training technique known as reinforcement learning, based on the fundamental principle of rewarding and punishing favorable behavior. At its core, a reinforcement learning agent can understand and interpret its environment, enabling informed action and knowledge acquisition through the iterative process of trial and error. Analysis of medical data can fine-tune drug dosages for individual patients based on their characteristics and past responses [33]. The convergence of deep learning and reinforcement learning holds the potential to improve the analysis of medical data, ultimately leading to better treatment outcomes. However, the implementation of these techniques depends on large volumes of carefully curated, high-quality data that require careful validation and rigorous evaluation to ensure their safety and efficacy. It is equally important to consider ethical issues such as privacy and mitigating bias when implementing these innovative methods in healthcare. There is immense potential in integrating AI into healthcare, but there are still challenges to overcome. Privacy and security issues, representative training data, accurate algorithms, and prospective clinical trials are critical aspects to consider. Robust solutions are needed to ensure data privacy, obtain diverse and high-quality training data, develop reliable algorithms, and conduct rigorous clinical evaluations to integrate medical AI into current workflows seamlessly.

Healthcare could be significantly transformed by edge computing, which enables the timely examination of significant amounts of data derived from wearable devices, medical instruments, and other related sources in real time.

1. *Wearable devices*: Wearable technology, which includes devices such as fitness trackers and smartwatches, is evolving into a comprehensive data repository that contains various details about a person's physical movements, heart behavior, and sleep dynamics during nighttime hours (see Figure 2). The use of edge computing facilitates the instant processing of data originating from wearable devices. This accelerates the detection of changes in a person's health status that may serve as precursors to the onset of chronic disease. In a recent study by the authors in [21], several research papers were analyzed to examine the use of wearable devices in detecting and predicting cardiovascular disease. The study results suggest that wearable devices can effectively detect, predict, and treat cardiovascular disease. However, more research is needed to improve their use. In addition, in [34], the authors proposed a secure edge-computing-based framework for smart health systems. This framework focuses on real-time health monitoring and ensures data security and confidentiality through clustering approaches for anomaly detection and attribute-based encryption (ABE) for secure access to biosignal data. Experimental results of the proposed framework show an improved performance, an accuracy up to 98.5%, and data security.
2. *Medical Devices*: IoT sensors integrated into medical devices provide valuable data for early health prediction and continuous monitoring of patient's health conditions. Edge computing is a viable option for processing data obtained from medical devices such as blood glucose meters in real time. Processing data is an indispensable component that helps detect changes in a person's health that could indicate the development of chronic diseases such as diabetes [35]. Currently, there are several medical devices that use edge computing technology for their operations:
    - *Smart insulin pens*: Smart insulin pens, used to control blood glucose levels in diabetic patients, are advanced medical devices that integrate sensors and connectivity to track insulin doses and blood glucose levels and provide real-time feedback to patients [36]. These devices can process data in real time and provide patients with personalized insights through edge computing.

- *Wearable electrocardiogram (ECG) monitors*: Wearable ECG monitors are medical devices used to monitor the patient's heart's electrical activity continuously. These monitors are attached to various areas of the body, such as the wrist, chest, or torso, and are aimed at individuals struggling with heart problems [37]. By integrating edge computing, these devices are able to analyze ECG data instantly. As a result, they are able to quickly detect irregular heart rhythms and other critical abnormalities, warranting immediate medical intervention.
- *Smart inhalers*: Smart inhalers are devices designed to treat asthma or chronic obstructive pulmonary disease (COPD). They are equipped with sensors and connectivity features that allow them to track medication use and provide feedback to patients [38,39]. The use of edge computing in smart inhalers has further enhanced their functionality by enabling real-time data processing and personalized patient insights.
- *Vital sign monitors*: Vital sign monitors are medical devices that monitor various physiological parameters such as heart rate, blood pressure, and oxygen saturation. These devices can be used in hospitals, clinics, or at home and allow real-time monitoring of the patient's health status. Using edge computing, these devices can analyze vital signs in real time and alert healthcare providers to anomalies [40,41].

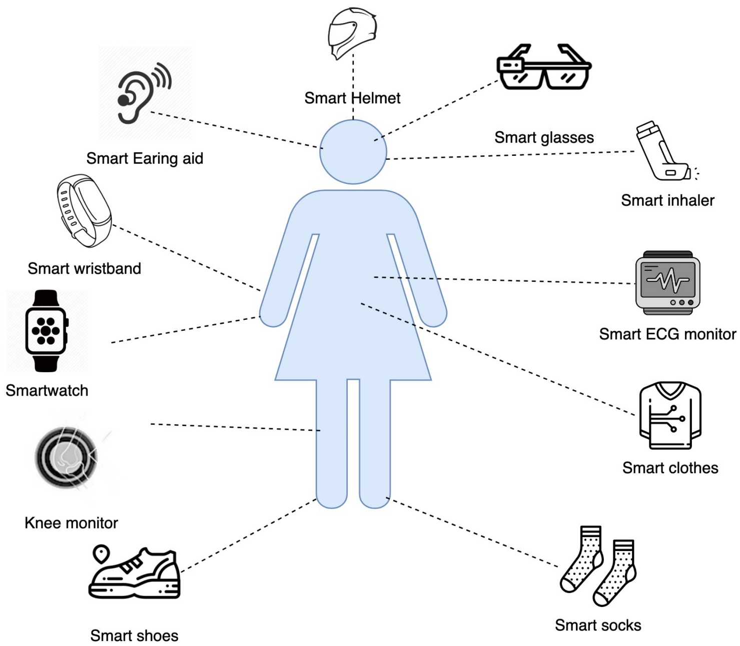

Figure 2. Common wearable devices to monitor several health parameters.

Edge computing facilitates remote patient monitoring and enables healthcare providers to monitor health statuses in real time and detect changes that signal the possible onset of a chronic disease [42,43]. It has the potential to significantly improve healthcare outcomes by enabling healthcare providers to detect chronic diseases at the onset when intervention is most effective. By processing data instantly and protecting sensitive information, edge computing can amplify public health outcomes by equipping healthcare providers with the insights needed to make informed judgments about an individual's health status. Edge computing has been used to predict the spread of infectious diseases through real-time analysis of data from multiple sources, including wearables, electronic health records, and social media. This method enables healthcare professionals to detect infectious disease outbreaks at an early stage, monitor the transmission of such diseases, and take proactive measures to contain their spread [43]. The framework proposed in [44] uses heart rate and sleep data from wearable devices to predict the progression of COVID-19 outbreaks in different countries and cities. The model claims that early detection and warning of epidemics can be facilitated by such predictive models. Edge computing offers a variety of advantages that make it particularly suitable for predicting health conditions in their early stages:

- *Real-time data analysis*: Edge computing offers the ability to analyze large datasets in a timely manner, thereby detecting changes in a person's health status in virtually real time. Timely detection of potential health problems is critical for early health prognosis, enabling health professionals to identify potential health problems before they worsen and become more difficult to treat.
- *Shortened latency*: By processing data at the periphery of the network and closer to its source, edge computing effectively reduces the latency that typically occurs when data are transmitted to a central data center for analysis. Incorporating edge computing technology can provide significant benefits to healthcare providers by enabling rapid data analysis and informed decision making regarding a patient's health status.
- *Improved data security*: Edge computing is a notable advancement that aims to protect sensitive data from security breaches through localized data processing at the network perimeter. Limiting the need for extensive data transmission over the internet provides a critical solution for early health prognosis. This attribute plays a central role in maintaining the confidentiality of an individual's private health information, ensuring that their information is used only to promote their well-being.
- *Cost efficiency*: The increased cost efficiency of edge computing makes it particularly beneficial for healthcare providers, as they have the ability to leverage artificial intelligence and other cutting-edge technologies to improve healthcare outcomes.
- *Scalability*: Edge computing is a highly scalable framework, making it a suitable choice for health predictive initiatives that operate at scale. By performing data processing at the periphery of the network, edge computing can handle large amounts of data, enabling real-time monitoring of the health status of multiple individuals simultaneously.

4. Methodology

This review article uses a systematic approach to gathering the relevant literature related to the article's focus on "Edge AI for Early Detection of Chronic Diseases and Infectious Disease Spread". The search for relevant references used the following keywords: "healthcare and medicine", "artificial intelligence", "machine learning", "federated learning", "edge computing", "IoT and wearable devices", "Edge AI", "predictive models", "security and privacy", "infectious diseases", and "keywords specific to various diseases". Without aiming at complete coverage, the methodology encompasses the systematic collection, screening, and analysis of academic references from popular databases.

4.1. Search Criteria Formulation

The search criteria used were:

- C1: ("Edge computing" OR "IoT" OR "wearables" OR "wearable sensors") AND ("chronic" OR "infectious") AND ("disease");
- C2: ("AI" OR "edge intelligence" OR "machine learning" OR "deep learning" OR "federated learning") AND ("chronic" OR "infectious") AND ("disease" OR "health prediction");
- C3: ("predictive models") AND ("chronic" OR "infectious") AND ("disease");
- C4: "Privacy" AND "health prediction".

The purpose of this review paper is to answer the following research questions.

- **RQ-1**: What are the myriad applications of Edge AI in early health prediction? This research question seeks to uncover research efforts and breakthroughs in the use of Edge AI for the early detection of chronic and infectious diseases.

- **RQ-2**: What machine learning and deep learning models are used in early health prediction?
- **RQ-3**: What techniques and methods are used to preserve privacy in early health prediction and detection of the onset of chronic and infectious diseases?
- **RQ-4**: What are the potential open research issues and future directions of Edge AI for early health prediction and detection of chronic and infectious diseases? This question seeks to define the unanswered inquiries and unexplored paths that hold the key to unlocking the full potential of Edge AI in early health prediction. By unraveling the challenges that can hinder their widespread adoption and delving into research directions, this query drives researchers to understand the current landscape of Edge AI and early disease detection, unraveling novel insights and paving the way for transformative advancements in this domain.

4.2. Source Selection and Approach

An extensive exploration was undertaken utilizing various popular databases and search engines to gather pertinent research material for this review. Three popular databases (Scopus, Google Scholar, and PubMed) renowned for their comprehensive coverage were used to search for scholarly works on the subject. The search strategy was based on the above search criteria. A time constraint was applied, restricting the search to include articles published between 2019 and 2023.

Most of the papers reviewed are journal articles or conference papers. They were selected on the basis of the quality of the journal and relevance to the topic and filtered by date of publication. The selection of articles is based on titles relevant to the topic of this review. The initial search for the above search criteria (C1–C4) found 556 references from Scopus, 565 references from Google Scholar, and 198 references from PubMed. However, the total number of references, 1319, was reduced to 280 after eliminating duplicates and incomplete references. Further screening of the title, abstract, and full text allowed the elimination of 182 references that addressed issues far from the main topic of this review article. The final number of references eligible for this study was 98. These references do not include the references we cited in the background sections.

5. Edge AI for Early Health Prediction

This section addresses **RQ-1**. Here, we discuss the myriad applications of Edge AI in early health prediction, which is the central focus. It also delves into the machine learning and deep learning models used for early health prediction, answering **RQ-2**.

5.1. Stakeholders and Architecture

The synergy of AI and edge computing promises to reshape healthcare by empowering healthcare providers with instant data and insights to drive early detection and prevention of diseases. Nevertheless, the full realization of these technologies depends on sustained investment in research and development, careful management of ethical and regulatory issues, and strategic management of challenges and constraints inherent in these areas. These obstacles are discussed in Section 7.

An Edge AI-based early health prediction system, such as the one depicted in Figure 3, involves several key stakeholders, including patients, healthcare providers, technology companies, governments and regulators, and research institutions. Each of these stakeholders plays an indispensable role in developing and implementing Edge AI-enabled systems tailored to early health prediction. Working together can help ensure the systems' efficiency, safety, and ethical behavior, with significant benefits for patients and healthcare providers. *Patients* are the primary users of Edge AI systems for early health prediction and are responsible for collecting and transmitting data to healthcare providers. *Healthcare providers* are tasked with reviewing the data collected from their patients and using the information to make informed judgments about patient care. This includes the processes of diagnosing diseases, developing appropriate treatment plans, and monitoring patient well-being over time. The task of developing and implementing Edge AI systems for healthcare is being undertaken by *technology enterprises*. This includes the development of sensors, devices, and AI algorithms in conjunction with the execution of edge computing and cloud infrastructure. *Governmental and regulatory entities* assume a central role in formulating benchmarks and guidelines for the use of Edge AI in healthcare. This includes ensuring compliance with privacy and security laws and protecting patients from potential harm. Finally, *Research Institutions* play a key role in advancing the field of Edge AI and developing new technologies for early health prediction. This includes participating in research and development projects, conducting clinical trials, and disseminating research findings through scientific publications.

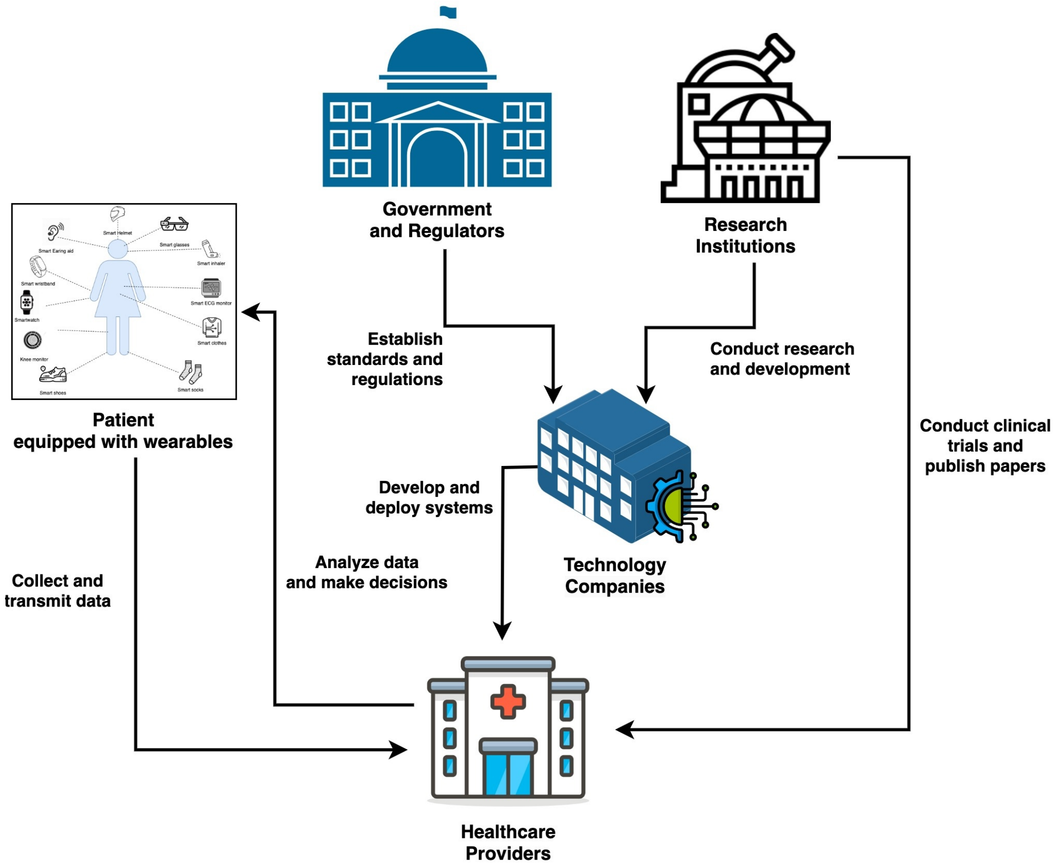

Figure 3. Main stakeholders of an Edge AI-based system for early health prediction.

As shown in Figure 4, an Edge AI-based architecture for early health prediction typically consists of several key components, including:

1. *Sensors and Devices*: The foundation of AI-driven edge systems for early health prediction depends on the use of sensors and devices that collect and transmit patient data. These devices typically include wearable devices such as fitness trackers and smartwatches and medical devices such as ECG monitors and glucometers.
2. *Edge Computing Devices*: Edge computing devices are responsible for processing and analyzing data at the point of origin. They include a variety of components, such as edge gateways, edge servers, and edge routers.
3. *AI Algorithms*: AI algorithms play a central role in carefully examining data collected by sensors and devices and facilitating the detection of latent health anomalies. By leveraging historical patient data, these algorithms can be trained to improve their accuracy in detecting patterns and trends indicative of an imminent health problem.

4. *Cloud Platforms*: Cloud platforms serve as repositories for data acquired by Edge AI systems. The information thus obtained can be used for deeper analysis and modeling. Subsequently, these data are even used to refine AI algorithms through training.
5. *User Interfaces*: User interfaces allow patients and healthcare providers to access the data and insights generated by Edge AI systems. Accessible through a variety of media, including desktop computers, mobile devices, or other internet-connected platforms, these interfaces play a critical communicative role.
6. *Data Management and Security*: Edge AI-based early health prediction systems require robust data management and security systems to protect patient data. This includes increased data storage, data encryption, and the implementation of strict access control protocols.

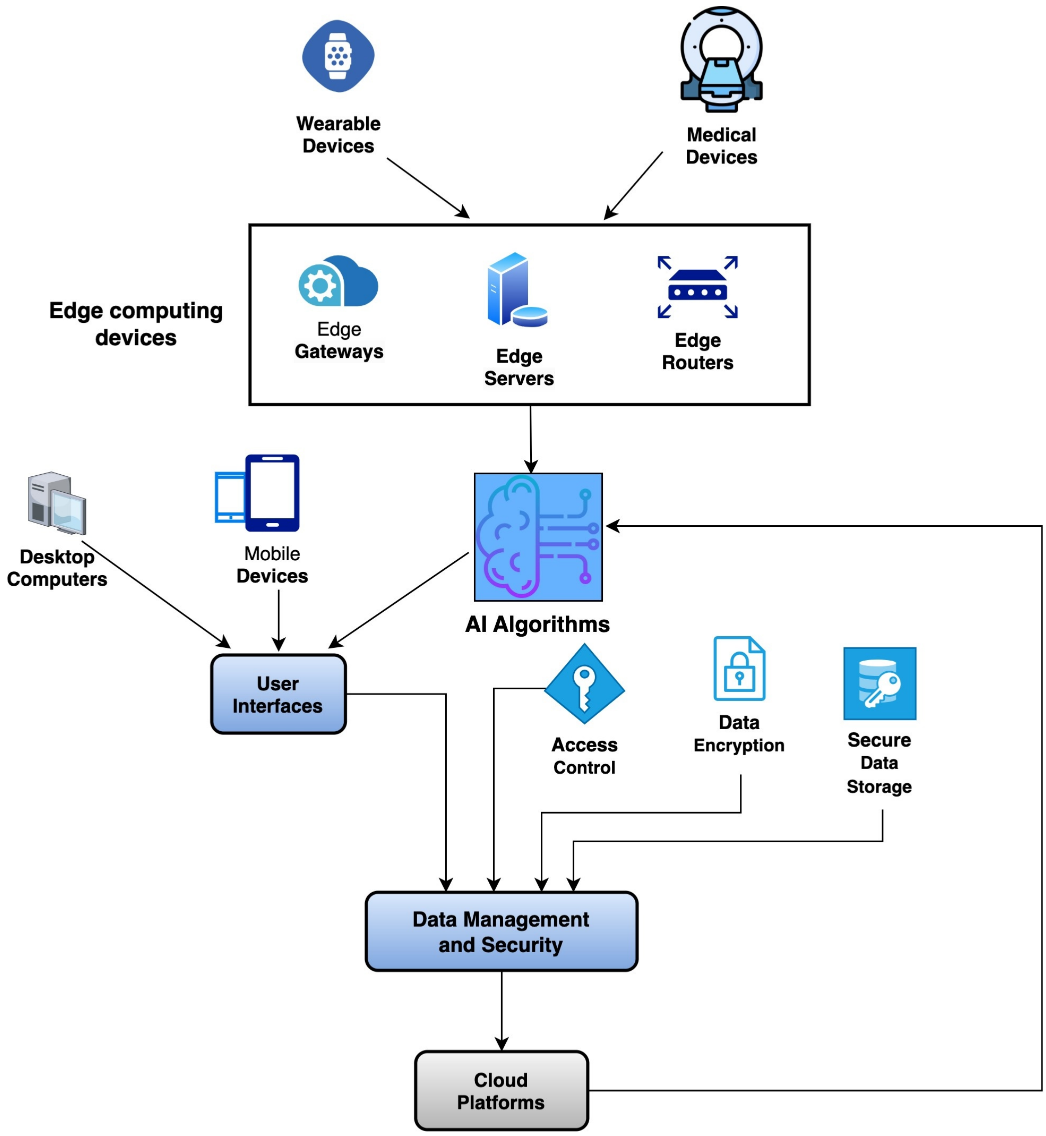

Figure 4. Edge AI-based architecture for early health prediction.

5.2. Early Detection of Chronic and Infectious Diseases

Several studies have investigated the use of AI for the early detection of chronic diseases such as cardiovascular disease, diabetes, cancer, and infectious diseases [14,15,45–52]. These research efforts highlight the ability of AI to significantly improve public health outcomes by allowing health professionals to detect chronic diseases in their early stages when interventions are most effective. By analyzing data from multiple sources, AI algorithms comprehensively summarize a person's health status and predict how health conditions will evolve. This allows healthcare providers to develop tailored interventions and therapies to improve an individual's health and slow disease progression.

Cardiovascular Disease: AI algorithms process data derived from wearables, electronic health records, and other sources to reveal deviations in a person's heart rate and activity metrics [53]. These discrepancies indicate the onset of cardiovascular disease. In [50], the authors describe recent developments in digital health applications tailored to cardiovascular disease. They focus on methods for detecting, diagnosing, and predicting cardiovascular disease using AI models based on data from wearables. They summarized the literature on wearables and AI in cardiovascular disease diagnosis, followed by a detailed description of the prevailing AI approaches for modeling and prediction using data collected from sensors such as wearables. They asserted that machine learning and AI-driven models outperform traditional statistical methods in predicting cardiovascular disease. Further confirmation comes from a separate study, as stated in [54], claiming that machine learning models are the first choice for prediction and classification measures in heart disease.

Diabetes: AI algorithms are able to decode data from electronic health records and wearable devices. These algorithms can evaluate changes in a person's blood glucose levels and activity parameters that may indicate the onset of diabetes. This idea is supported by the study in [55], in which the authors meticulously explain machine learning and AI-based methods capable of detecting and self-treating diabetes mellitus.

Cancer: AI algorithms skillfully analyze medical imaging and electronic health records to detect health changes that indicate the presence of cancer. This is illustrated in [48], in which the authors explore the role of AI in the digital pathology of breast cancer, outlining both the current state and the challenges ahead.

COVID-19: AI holds great promise in several areas of COVID-19 data reviews. Foremost is its role in developing predictive models for COVID-19 diagnosis and prognosis. Using large datasets of COVID-19 patient data, AI algorithms have been refined to identify predictive patterns and attributes associated with disease severity, mortality, and important outcomes. The application of AI also extends to accelerating medical image analysis, including chest X-rays and CT scans, which critically aids COVID-19 patient diagnosis and monitoring. The authors in [56] discuss the development of a federated learning model called EXAM, which predicts the oxygen requirements of COVID-19 patients using data from 20 institutions worldwide. EXAM achieved a mean AUC (Area Under The Curve) of 0.92 in predicting outcomes within 24 and 72 h, demonstrating an improved average AUC and a broader applicability in contrast to models trained at only one site. FL facilitated rapid collaboration without data sharing and enabled a model that could be generalized across heterogeneous datasets, demonstrating the potential of FL in healthcare.

5.3. Prediction of Future Health Outcomes

Early prediction of health status has attracted considerable research attention in recent years because of its potential to improve patient care and reduce healthcare costs. Figure 5 shows the different approaches to early health prediction. AI algorithms can analyze data from multiple sources to predict patients' future health status. They use machine learning and statistical methods to find patterns in the data and identify risk factors. AI algorithms can efficiently analyze data and uncover hidden patterns and correlations that indicate potential diseases [24]. Table 1 summarizes these research efforts, including using AI to predict future patient healthcare. By examining data from these various sources, AI algorithms provide clinicians with a comprehensive picture of an individual's well-being, allowing them to predict future health outcomes. Given the previously described findings, this opens up the possibility of developing personalized interventions, therapies, and preventive strategies, all aimed at improving a person's holistic well-being and preventing the onset of chronic diseases. It is imperative to acknowledge the shortcomings of AI algorithms. Therefore, their predictions must be integrated in the context of variables such as individual lifestyle, family lineage, and current health status. Nonetheless, the prudent use of AI algorithms promises to improve public health significantly. By enabling physicians to identify and address latent health problems in their early stages, these algorithms are an excellent tool to prevent potential problems from escalating.

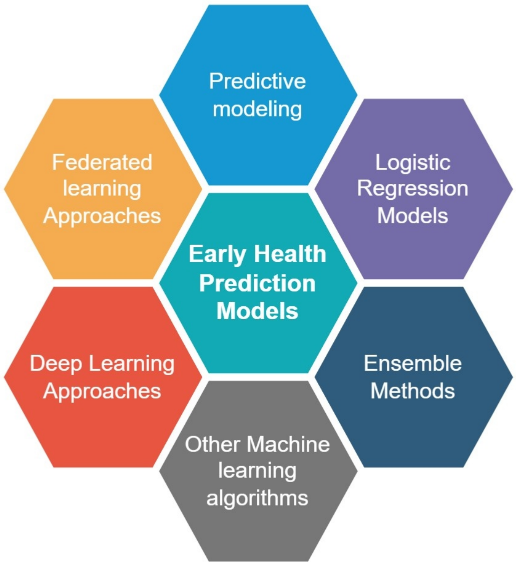

Figure 5. Early health prediction approaches.

5.3.1. Predictive Modeling

Predictive modeling uses data and statistical techniques to create models that can forecast future outcomes or behaviors. In healthcare, predictive modeling can be used for various purposes, such as identifying patients at a high risk of developing chronic diseases or complications [57–59], optimizing resource allocation and scheduling, improving quality of care and patient satisfaction, reducing costs and waste, and enhancing clinical decision making and diagnosis. Predictive modeling in healthcare requires a multidisciplinary approach that involves data collection, preprocessing, analysis, validation, and deployment. It encompasses various models designed to analyze historical data, identify patterns, and predict future trends. Among the commonly used models are classification, clustering, and time series models. Classification models are frequently employed in healthcare settings to make decisions related to improving patient health, optimizing healthcare services, and detecting health insurance fraud. Clustering models allow for profiling individuals based on characteristics such as age, inpatient admissions, and the risk of emergency hospital admission in the next 12 months. Time series models enable the plotting of observations made over time, such as monthly emergency department admissions or annual healthcare expenditures. Predictive modeling in healthcare faces many challenges, such as data quality, privacy, ethics, and interpretability. Therefore, it is important to ensure that the models are accurate, reliable, transparent, and fair.

5.3.2. Machine Learning Algorithms

Machine learning algorithms are able to analyze information and detect changes that may indicate a medical problem [60–62]. Early-stage health prediction has been approached using various machine learning techniques. Notable models and methods in this area include logistic regression models, ensemble methods, and decision trees [45,61,63–65]. Through rigorous training on large datasets, these algorithms reveal hidden patterns and relationships in the data structure. Consequently, these detected patterns can be used to predict individuals' future health trajectories by gaining insights from their medical history and a constellation of relevant factors. Logistic regression models have been widely used in health status prediction because of their simplicity and interoperability [45,66–68]. These models are suitable for binary classification tasks that involve predicting the presence or absence of a particular health state based on patient data. Ensemble methods, such as random forest and Gradient Boosting, have gained popularity due to their ability to combine multiple base models to improve prediction accuracy [46,47,69,70]. They have been applied in the context of early health prediction to capture complex interactions in health data.

5.3.3. Deep Learning Approaches

Deep learning is a subset of machine learning that relies on neural networks as a fundamental framework. It has gained prominence due to the convergence of two factors: the availability of Big Data and the accessibility of cost-effective parallel computing hardware like Graphical Processing Units (GPUs) and computer clusters. Deep learning uses computational intelligence to acquire knowledge, learn from experience, and develop complex concepts from simple ones. Unlike traditional machine learning, deep learning can autonomously learn features from input data. A deep neural network develops its classification capabilities by analyzing thousands of labeled images during the training process. When presented with an input image or video, the neural network layers respond to complex shapes and structures, comparing them to the training data to identify the image or extract relevant features. The commonly used models are:

- *Convolutional Neural Networks (CNNs)*. CNNs are pivotal in edge analytics, excelling in object detection and recognition. Their capacity to handle substantial quantities of visual data and discern complicated patterns is of utmost importance in the interpretation of meaningful observations from visual information.
- *Recurrent Neural Networks (RNNs)*. RNNs serve a pivotal function in video analysis by effectively managing tasks such as tracking, segmentation, and action recognition. With their proficiency in processing sequential data, RNNs are well suited for analyzing videos, which are essentially temporal image sequences.

Deep learning excels at recognizing and classifying complex objects. Deep learning techniques, especially neural networks, have shown remarkable performance in early disease detection [66,67,71]. CNNs and RNNs have been used to analyze medical images and time series data for early disease detection [48].

5.3.4. Federated Learning (FL)

By implementing federated learning, healthcare providers can effectively train models on patient data that come from many different sources. This approach allows for a more comprehensive and diverse set of data from which the model can learn, leading to greater accuracy in predicting and detecting diseases in their early stages. In preliminary health diagnosis, the implementation of federated learning is an effective mechanism to develop predictive models that can predict the likelihood of diseases such as cancer, diabetes, and heart disease [72–74]. By thoroughly analyzing data collected from different patients, the model can efficiently identify patterns and risk factors that would be difficult to detect in a single patient. In addition, federated learning can help address privacy issues in healthcare. The risk of data breaches is significantly reduced because patients' data are stored on their devices rather than sent to a central server, which increases patient privacy [72,75].

5.3.5. Evaluation Metrics of Edge AI Algorithms

Various metrics are commonly employed to assess the efficacy of Edge AI algorithms. The key metrics of utmost significance include the following:

- *Accuracy*. Accuracy refers to the measure of how well the algorithm performs in correctly identifying or classifying data in the edge devices or edge computing systems. It quantifies the level of agreement between the algorithm's predictions and the ground truth or desired outcomes.

- *Precision and Recall*. Precision and recall are two important metrics used to evaluate the performance of classification models. They provide insights into how well the algorithm is identifying and classifying data at the edge devices or edge computing systems. Precision measures the accuracy of positive predictions made by the algorithm. It is the ratio of true positives (correctly predicted positive instances) to the sum of true positives and false positives (incorrectly predicted positive instances). In other words, precision indicates the proportion of correctly identified positive instances out of all the instances predicted as positive. A higher precision value indicates a lower rate of false positives. Recall, also known as sensitivity or the true positive rate, measures the algorithm's ability to correctly identify positive instances. It is the ratio of true positives to the sum of true positives and false negatives (positive instances incorrectly classified as negative). Recall represents the proportion of correctly identified positive instances out of all the actual positive instances. A higher recall value indicates a lower rate of false negatives.
- *Intersection over Union (IoU)*. The IoU metric is particularly useful in evaluating the performance of object detection algorithms. It helps assess how well the algorithm accurately localizes and identifies objects within an image or video frame. By comparing the predicted and ground truth regions, IoU provides insights into the algorithm's ability to detect and segment objects accurately. It is used in Edge AI algorithms dealing with visual data to measure the overlap between the predicted bounding box or region and the ground truth bounding box or region in object detection and segmentation tasks. It provides a quantitative measure of the accuracy of the algorithm's predictions. IoU is calculated by dividing the area of intersection between the predicted and ground truth regions by the area of their union. The resulting value indicates the extent of overlap between the two regions. A higher IoU value indicates a better match between the predicted and ground truth regions. IoU is often used as a threshold to determine whether a predicted bounding box or region is considered a true positive or a false positive. If the IoU value exceeds a certain threshold (commonly 0.5 or 0.7), the prediction is considered a true positive. Otherwise, it is classified as a false positive.
- *Latency*. Latency in Edge AI algorithms refers to the time it takes for an algorithm to process input data and produce an output on edge devices or edge computing platforms. It is a measure of the delay or response time experienced during the execution of the algorithm. In the context of Edge AI, latency is a critical factor as it directly impacts the real-time performance and responsiveness of the system. Edge devices, such as smartphones, IoT devices, or edge servers, often have limited computational resources compared to cloud-based servers. This constraint can lead to a higher latency in executing AI algorithms on the edge. The latency in Edge AI algorithms depends on various factors, including the complexity of the algorithm, the computational power of the edge device, data processing requirements, and network connectivity. The goal is to minimize latency to ensure efficient and timely decision making at the edge.

Table 1. Edge AI for early detection of the onset of a disease and mortality prediction.

Ref.	Year	Application	Methods	Results
[45]	2022	Identification and prediction of chronic diseases using machine learning.	Data collection from various sources. Training with CNN and KNN algorithms	CNN and KNN models outperformed Naive Bayes, decision tree, and logistic regression algorithms. CNN and KNN achieved a higher precision, recall, and F1-score.
[75]	2022	A survey on Federated Learning for Privacy Preservation in Smart Healthcare Systems	FL for privacy preservation in IoMT. Advanced FL architectures incorporating DRL, digital twins, and GANs	Description of some advanced FL architectures incorporating deep reinforcement learning (DRL), digital twins, and generative adversarial networks (GANs) for detecting privacy threats.
[72]	2022	Review of Federated Learning for Healthcare	Systematic literature review methodology. Definition of research questions	Systematic literature review on FL in healthcare. Proposed architecture for FL applied to healthcare data
[63]	2022	Machine learning models used for early diabetes prediction.	SVM-ANN ensemble. SVM, KNN, NB, C4.5 DT, Adaboost DT with Bagging, Bagged DT, K-means clustering and RF, KNN and AB, Fusion ML Decision, LR ensemble, RF, Multilayer Perceptron, SVM with feed backward feature elimination	Random forest (RF) model achieved the highest accuracy of 82.26% in predicting diabetes. The Naive Bayes (NB) model performed the worst, with an accuracy rate of 70.56%.
[59]	2019	Health risk prediction models incorporating personality data.	Informal rule for cross-validation error. Anti-conservative approach to protect against overfitting	Four-year incidence rate of possible MCI is roughly 20%. Model C with personality data shows a significant improvement in overall performance.
[66]	2015	Comparison of predictive models for early hospital readmissions.	Logistic regression with maximum likelihood estimator. Logistic regression with multi-step heuristic approach	Random forest and penalized logistic regressions are the best methods for predicting early readmissions. Deep learning methods outperform regression methods in the healthcare literature.
[67]	2021	Predicting mortality among patients with liver cirrhosis.	Deep neural network (DNN), random forest (RF), and logistic regression (LR) algorithms. Multiple imputation for missing values in variables.	Models with all variables outperformed those with four MELD-NA variables. The DNN model achieved a higher AUC than the LR and RF models.
[76]	2022	Clinical prediction models to estimate disease probability and health outcomes.	Split-sample method for model development and internal validation. Resampling methods, especially the bootstrap method, for stable estimates	Overview of developing and validating clinical prediction models by applying traditional regression models or machine learning models.
[49]	2023	AI models for predicting and early diagnosis of pancreatic cancer.	Scoping review conducted following PRISMA-ScR guidelines. Two reviewers independently performed study selection and data extraction	Initially identified 18,285 articles from various databases. After screening and exclusion, 30 articles were included.
[64]	2022	Edge computing-based heart disease prediction.	Decision-tree-based classifier for analyzing health data. Pre-trained machine learning processing module for analysis	The decision tree classifier shows 99% accuracy for classifying the subject's position. The decision tree classifier shows 98% accuracy for heart disease prediction.

Table 1. *Cont.*

Ref.	Year	Application	Methods	Results
[77]	2023	Clinical language models for a wide range of clinical and operational predictive tasks.	Pretraining datasets: NYU Notes, NYU Notes-Manhattan, NYU Notes-Brooklyn. Fine-tuning datasets: NYU Readmission	NYUTron has an AUC of 78.7–94.9%. NYUTron improves the AUC by 5.36–14.7% compared to traditional models.
[78]	2023	Prediction of mental health problems after military deployment.	Neural network models used for prediction. Pre-deployment registry data combined with post-deployment questionnaire data	Approximately 95% of participants were male. The percentage of individuals deployed in a combat unit was highest for the first deployment (32.7%), compared to the second (25.6%) and the third (20.8%) deployment.
[65]	2022	Early-stage Alzheimer's prediction.	Decision tree, random forest, Support Vector Machine, Gradient Boosting, and voting classifiers. Machine learning techniques applied to Alzheimer's disease datasets.	Evaluation metrics: precision, recall, accuracy. Men are more likely to have dementia than women. Achieved 83% accuracy on test data
[57]	2014	Risk predictive modeling for diabetes and cardiovascular disease	Collection of data during a prospective study. Estimation of regression coefficients for predictor–outcome association	Emphasizes the importance of validating existing CVD and diabetes prediction models to improve their adoption in routine practice.
[68]	2019	Identifications of patients at risk of uncontrolled hypertension.	Logistic regression and recurrent neural networks	Best model achieved an AUROC of 0.719. Linear models performed better than recurrent neural networks
[79]	2021	Predicting critical state after COVID-19 diagnosis	Prognostic model trained on US electronic health records. Feature reduction process based on SHAP values	ROC AUC: 0.861 [0.838, 0.883]. Precision–recall AUC: 0.434 [0.414, 0.485]
[61]	2022	Machine learning for healthcare wearable devices.	Homomorphic Encryption (HE). Secure Multiparty Computation (SMPC)	Review of different areas of machine learning research for wearable healthcare devices.
[58]	2018	Clinical prediction models (CPMs) with statistical updating models.	Regression coefficient updating. Meta-model updating	Original ES overestimated mortality (calibration intercept—1.06, slope—0.97). All updating strategies improved calibration performance.
[80]	2019	Continuous risk predictions for acute kidney injury.	Binary variable prediction for AKI occurrence. Eight future time horizons for predictions.	Model achieves a higher ROC AUC in shorter time windows. The model achieves a lower PR AUC in shorter time windows.
[71]	2019	Deep learning models for early prediction of acute adverse events.	Developing deep learning continuous risk models. Integrating domain knowledge into the technical specification	Continuous risk models identified 55.8% of AKI cases up to 48 h early with a false positive rate of 2:1. The model correctly predicted 90.2% of AKI cases requiring dialysis within 90 days.
[48]	2020	AI in digital pathology for breast cancer diagnosis.	Image analysis with deep learning (DL). Non-CNN algorithms for segmentation and detection	Review of the basics of digital pathology and AI and the challenges in the field.

6. Federated Learning for Early Health Prediction

This section addresses federated learning's role in health prediction. It aligns with **RQ-1**. It also examines the issue of preserving privacy in early health prediction, which relates to **RQ-3**.

Federated learning holds tremendous potential for early health prediction in scenarios where healthcare providers equip patients with wearables and various medical devices. This approach to training AI models offers a variety of benefits that can usher in transformative change in healthcare [81,82]. Let us take a closer look at the reasons for the promising outlook:

- *Preserving Data Privacy:* In federated learning, patient data are securely stored on edge devices such as medical devices and wearables without being transferred to a central server for storage. This decentralized approach ensures strict data privacy and security and mitigates the risk of data breaches or violations.
- *Access to Multiple Data Sources:* Federated learning enables healthcare providers to access diverse patient data collected by wearables and medical devices. This data diversity promotes the development of more robust and accurate predictive models by encompassing a wide range of health-related information from multiple sources.
- *Leveraging Large-Scale Data:* With federated learning, healthcare providers can leverage a vast amount of data from a distributed patient network. By aggregating and combining these disparate datasets, healthcare professionals can build comprehensive models capable of capturing nuanced patterns and early health indicators, resulting in highly accurate predictions and timely interventions.
- *Continuous Learning and Adaptation:* As patients continuously use wearables and medical devices, new data become available over time. Federated learning facilitates continuous training of the model and adaptation to changing health conditions. The global model can be periodically updated with the latest aggregated parameters to incorporate new knowledge and improve the system's predictive capabilities.
- *Personalized and Adaptive Models:* Federated learning facilitates the creation of personalized models tailored to individual patients. By training local models based on their specific data, the system can tailor predictions and interventions to individual health profiles, leading to highly personalized healthcare and early detection of health problems.
- *Reduced Data Transfer and Computational Overhead:* Federated learning minimizes the need to transfer huge amounts of raw patient data to a central server. Instead, the focus is on transferring aggregated model parameters, significantly reducing bandwidth requirements and computational overhead while carefully protecting the privacy of patient data.
- *Collaborative Research and Knowledge Sharing:* Federated learning promotes collaborative research and knowledge sharing among healthcare providers by pooling anonymized, aggregated parameters, insights, and discoveries can be shared without compromising patient privacy. This collective intelligence accelerates medical progress and improves early health prediction capabilities.

By leveraging the power of federated learning in scenarios where patients use wearables and various medical devices, healthcare providers can improve early prediction of health statuses, facilitate personalized interventions, and improve overall healthcare outcomes while carefully protecting patient privacy and data security.

6.1. Federated Learning Methodology

Figure 6 shows the training process of a federated learning model for early health prediction. The methodology of federated learning with edge devices involves the following steps:

1. **Data Distribution**: The training data are distributed across multiple edge devices. Each device holds its own local data, which may be collected from different sources or users.
2. **Local Model Training**: A local model is trained using its own local data on each edge device. The training process can be performed using various machine learning algorithms and techniques.
3. **Model Aggregation**: After the local model training, the updated models from each edge device are sent to a central server or aggregator. The aggregator collects the models and performs model aggregation techniques, such as averaging or weighted averaging, to create a global model.
4. **Model Update**: The global model is then sent back to the edge devices, where it replaces the local models. This updated global model incorporates the knowledge learned from all the edge devices' local models.
5. **Iterative Process**: The above steps are repeated iteratively, allowing the edge devices to continuously improve the global model by training on their local data. This iterative process helps capture the data diversity across different edge devices and improve the overall model performance.

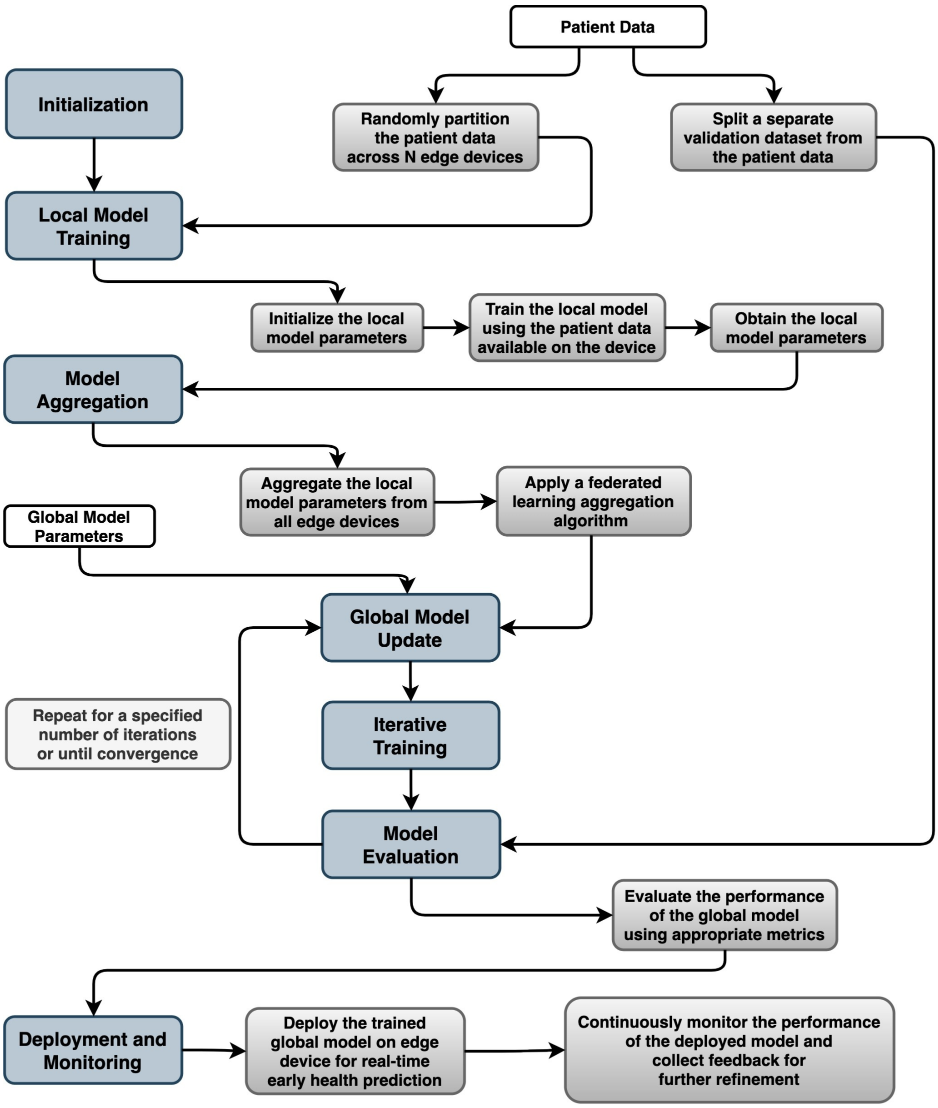

Figure 6. Model training in a federated learning scenario for early health prediction.

Collaborative training on edge devices is a core aspect of federated learning in this context. Researchers develop federated learning techniques specifically tailored for edge devices to enable collaborative model training. Secure aggregation methods are employed to combine model updates contributed by edge devices. Federated learning aggregation methods play a critical role in combining local model updates from different edge devices, resulting in a global model that encompasses the collective knowledge. Let N be the number of edge devices (1 to N) participating in the federated learning process. Each edge device i has a local model w_i that it trains on its own local dataset. After training, the local model w_i is updated and must be merged with the models of the other edge devices to form the global model W. The selection of an aggregation approach depends on certain assumptions, constraints, and properties of the federated learning context. Different techniques may be more suitable for different implementations, and new aggregation methods are continuously being sought to increase the effectiveness and efficiency of federated learning systems.

Averaging is the most widely used aggregation method in federated learning. In this approach, consensus is reached on the global model by averaging the local model updates. Each update is treated equally in this approach without considering specific weighting factors. The averaging approach can be represented mathematically as follows:

$$W = \frac{1}{N}\sum_{i=1}^{N} w_i$$

Here:

- $\sum$ indicates the summation of local models across all edge devices.
- $\frac{1}{N}$ is the normalization factor to ensure that the aggregation results in an averaged global model.

In this approach, local models are averaged to create a global model representing the collective learning of all edge devices. The initial phase of the training procedure involves using the global model as the basic framework, which is then iterated until a state of convergence is reached. Algorithm 1 below depicts how averaging can be used in federated learning with edge devices.

Weighted averaging extends the basic averaging method by assigning different weights to each local model update. These weights can be determined based on factors such as the reliability of the edge device, the amount of local data, or the performance of the local model. Consequently, devices with better performance or more representative data significantly affect the global model. The weighted averaging approach can be represented mathematically as follows:

$$W = \sum_{i=1}^{N} \alpha_i \cdot w_i$$

Here:

- α_i represents the weight assigned to the edge device i. Factors such as the amount of locally available data and the computational capacity of the edge device are considered in determining these weights.
- $\sum$ represents the sum of the weighted local models across all edge devices.

In this method, the local models are consolidated through the use of weighted averaging, where each edge device's individual input to the global model is adjusted in proportion to its corresponding weight [83]. In this way, edge devices with more relevant or accurate data can impact the global model more. Algorithm 2 below shows how the weighted averaging method can be used in federated learning with edge devices.

In Algorithm 1, federated learning with averaging, the gradient updates ($\nabla F(w_i)$) received from each edge device i are equally weighted during aggregation (computation of the aggregate gradient update as $\Delta w_i = \frac{1}{N}\nabla F(w_i)$. This means that each local model's contribution to the global model update is treated identically.

Algorithm 1 Federated Learning with Averaging

Require: Global model W, Edge devices $\{1, 2, \ldots, N\}$, Learning rate η, Rounds T

Ensure: Updated global model W

```
Initialize global model W
for t = 1 to T do
    Send global model W to all edge devices
    for each edge device i in {1, 2, ..., N} do
        Receive local model w_i from edge device i
        Compute gradient ∇F(w_i) on local model w_i
    end for
    Initialize aggregated gradient ΔW = 0
    for each edge device i in {1, 2, ..., N} do
        Compute gradient update Δw_i = (1/N) ∇F(w_i)
        Aggregate gradient updates: ΔW = ΔW + Δw_i
    end for
    Update global model: W = W − η · ΔW
end for
Return Updated global model W
```

Algorithm 2 Federated Learning with Weighted Averaging

Require: Global model W, Edge devices $\{1, 2, \ldots, N\}$, Learning rate η, Rounds T

Ensure: Updated global model W

```
Initialize global model W
for t = 1 to T do
    Send global model W to all edge devices
    for each edge device i in {1, 2, ..., N} do
        Receive local model w_i from edge device i
        Compute gradient ∇F(w_i) on local model w_i
        Compute weight α_i based on device-specific factors
    end for
    Initialize aggregated gradient ΔW = 0
    for each edge device i in {1, 2, ..., N} do
        Compute weighted gradient Δw_i = α_i · ∇F(w_i)
        Aggregate weighted gradient: ΔW = ΔW + Δw_i
    end for
    Update global model: W = W − η · ΔW
end for
Return Updated global model W
```

In Algorithm 2, federated learning with weighted averaging, we introduce the concept of device-specific factors or weights, denoted as α_i, for each edge device i. These factors are based on device-specific characteristics or criteria. The local models' gradient updates are multiplied by their respective weights before aggregation (weighted gradient $\Delta w_i = \alpha_i \cdot \nabla F(w_i)$), and the aggregate weighted gradient is computed as $\Delta W = \Delta W + \Delta w_i$. This allows for individualized contributions to the global model, accounting for device-specific considerations. Therefore, Algorithm 2 incorporates a level of customization not present in Algorithm 1. By introducing device-specific factors, it accommodates the potential variations among edge devices. This customization can be based on factors such as device capabilities, data quality, or performance history. Doing so provides a mechanism to incorporate this diversity into the federated learning process.

Secure aggregation methods emphasize privacy and security during the aggregation process. Techniques such as secure multi-party computation (SMPC) or homomorphic encryption ensure that local model updates remain confidential. These methods aggregate updates without disclosing sensitive information to the central server or other parties, thus maintaining privacy. SMPC allows edge devices to perform computations on their local models while maintaining privacy. The aggregation process can be represented mathematically as follows:

$$W = \text{SMPCAggregation}(w_1, w_2, \ldots, w_N)$$

Here:

- SMPCAggregation is the secure multi-party computation aggregation function.

In SMPC, each edge device i contributes its local model w_i without revealing the parameters of the model or the data used for training. Secure protocols such as homomorphic encryption or secret sharing facilitate the computation of aggregate statistics, including gradients, on encrypted data. These encrypted statistics are then combined to update the global model W. SMPC ensures that individual edge device data remain secret throughout the aggregation process, making it suitable for scenarios with strict privacy requirements [84].

6.2. Federated Learning for Disease and Mortality Prediction

A growing body of research is addressing the application of federated learning in the area of disease prediction and mortality forecasting. The results of these research efforts are clearly summarized in Table 2.

Table 2. Federated learning for early detection of the onset of a disease and mortality prediction.

Ref.	Year	Application	Methods	Results	Limitations
[85]	2022	Prediction of ICU mortality risk	Deep federated learning, centralized machine learning, and local machine learning.	Federated learning performs equally well as the centralized approach. Federated learning is substantially better than the local approach.	- The generalizability of the approach beyond the MIMIC-III dataset needs to be tested. - Does not address the performance of the FL workflow in predicting ICU mortality at an early stage using other datasets. - Does not discuss potential challenges or limitations of implementing FL in a healthcare setting.

Table 2. *Cont.*

Ref.	Year	Application	Methods	Results	Limitations
[20]	2023	Review of federated learning being used for early health prediction in diseases such as cardiovascular disease, diabetes, and cancer.	Federated learning and FL aggregation algorithms.	Federated learning helps solve privacy concerns in machine learning; 32% of companies plan to implement federated learning.	- FL in disease prediction is still in the early stages. - Limited real-world examples and use of smart wearables - The paper acknowledges that there is no unified classification of FL challenges in the literature
[86]	2023	Early prediction of cardiovascular disease.	Modified artificial bee colony optimization with Support Vector Machine (MABC-SVM). Federated matched averaging for the HSP server.	The proposed hybrid technique for federated learning improves the prediction accuracy by 1.5% and achieves a 1.6% fewer classification errors	- The paper does not mention any limitations of the proposed hybrid classifier-based FL framework. - The paper does discuss the limitations or constraints of the dataset used for testing and evaluation - No mention of ethical or legal considerations associated with the collection, sharing, and use of biomedical data in the proposed framework.
[87]	2022	Federated learning to predict heart disease by training a shared model while keeping patient data distributed across multiple locations.	Logistic regression and Support Vector Machine (SVM).	Accuracy of around 89% on the UCI benchmark dataset.	- The paper does not provide details on the size or diversity of the dataset used for training the shared model. - The paper does not provide a comparison of the FL approach with other existing machine learning methods for heart disease prediction.
[88]	2022	Federated learning to predict chronic kidney diseases	Federated learning and image processing technique to identify the affected area.	Higher accuracy, efficiency, specificity, and sensitivity. Increased accuracy through training with a single image.	- No mention of the size or diversity of the decentralized data used for training the algorithm. - Lack of information on the dataset used for training and testing the model - Insufficient explanation of the image processing techniques employed for identifying the affected area of the kidney.
[89]	2022	Combining the IoT and blockchain in healthcare. Focus on individualized health monitoring and early disease detection using wearable gadgets.	Combining the IoT and blockchain technologies and applying blockchain-based federated learning.	Blockchain-based federated learning offers benefits such as smarter simulations, lower latency, lower power consumption, and privacy.	- Limited battery life of IoT devices is a major obstacle to integrating blockchain with the IoT. - High processing and bandwidth requirements for blockchain integration

Table 2. *Cont.*

Ref.	Year	Application	Methods	Results	Limitations
[90]	2023	Dynamic federated meta-learning approach to improve rare disease prediction using federated learning.	Dynamic federated meta-learning (DFML) and inaccuracy-focused meta-learning (IFML) approach.	The proposed model improves the prediction accuracy by 13.28% and outperforms the original federated meta-learning algorithm in accuracy and speed.	- Lack of information on the specific rare diseases studied and the datasets used for evaluation. - Does not provide details on the computational resources required for implementing the DFML approach. - Does not compare the proposed DFML approach with other state-of-the-art methods for rare disease prediction.
[91,92]	2021	Federated learning to predict mortality in hospitalized patients with COVID-19 within 7 days using electronic health record data.	Logistic regression with L1 regularization/least absolute shrinkage and selection operator (LASSO) and multilayer perceptron (MLP)	The LASSO federated model outperformed the LASSO local model at three hospitals. The MLP federated model performed better than the MLP local model at all five hospitals.	- The study was limited to data collected from hospitals within the Mount Sinai Health System in NYC. - The study only included clinical data in the models. - The study only implemented two widely used classifiers within the framework.

6.2.1. Cardiovascular Diseases

The authors in [86] introduced an innovative approach to leveraging federated learning for the prediction of cardiovascular diseases in healthcare settings. Their study proposed a hybrid classifier-based framework that combines local models on individual health service providers' data with a centralized classifier, thereby striking a balance between data privacy and model accuracy. By harnessing the power of federated learning, this research aimed to enhance the accuracy of cardiovascular disease predictions while preserving sensitive patient data, making it a promising solution for healthcare institutions seeking to improve patient care and outcomes in a privacy-conscious manner.

6.2.2. Heart Diseases

Bharathi et al. in [87] presented a novel application of federated learning in the context of heart disease prediction. Their research introduced a federated learning framework that allows multiple healthcare institutions to collaboratively train a predictive model while keeping patient data decentralized and secure. By aggregating local model updates from different data sources, this approach achieves improved accuracy in heart disease prediction, facilitating early detection and intervention. The study highlights the potential of federated learning to harness the collective knowledge of distributed healthcare providers, enhancing the quality of predictive models without compromising patient data privacy, which is crucial for advancing cardiovascular disease management in a privacy-aware healthcare landscape.

6.2.3. Chronic Kidney Diseases

The authors in [88] introduced an innovative approach to predicting chronic kidney diseases using federated learning. Their investigation leverages the collaborative power of federated learning to train predictive models on data from various healthcare providers while preserving the privacy of patient information. By aggregating insights from multiple sources, their study aimed to enhance the accuracy of chronic kidney disease prediction, offering valuable early diagnosis capabilities. They underscored the significance of federated learning in healthcare, particularly in chronic disease management, where privacy concerns are paramount. This approach holds promise for improving patient outcomes and resource allocation while maintaining the highest standards of data privacy and security.

6.2.4. Rare Diseases

Chen et al. in [90] introduced an innovative approach to address the challenging task of predicting rare diseases using dynamic federated meta-learning. This research presents a unique framework that adapts to the inherently scarce data associated with rare diseases. By leveraging federated meta-learning techniques, the study combines knowledge from diverse sources while allowing models to adjust to each rare disease's specific characteristics dynamically. This approach significantly enhances the accuracy of rare disease prediction, which is often hampered by limited data availability. The paper underscores the potential of dynamic federated meta-learning to improve healthcare outcomes for individuals with rare diseases, showcasing the versatility of federated learning in addressing complex and under-studied medical conditions while safeguarding patient privacy.

6.2.5. ICU Mortality

Randl et al. in [85] presented an innovative application of deep federated learning in the context of predicting ICU mortality risk. This research introduced a novel framework that leverages deep learning models while preserving the privacy of sensitive patient data in intensive care unit (ICU) settings. Their approach aimed to improve the accuracy of ICU mortality risk predictions by enabling collaborative model training across multiple healthcare institutions. It highlights the critical role of federated learning in healthcare, particularly in ICU scenarios, where early risk assessments can lead to better patient outcomes and resource allocation. Furthermore, it demonstrates the potential of deep federated learning to enhance critical care by predicting mortality risk while upholding the highest data privacy and security standards, making it a significant advancement in the field of healthcare analytics.

6.2.6. COVID-19 Mortality

Vaid et al. in [91] introduced a pioneering application of federated learning to enhance mortality prediction in hospitalized COVID-19 patients. Their research employed machine learning techniques in a federated framework, enabling collaborative model training on electronic health records (EHRs) from various healthcare facilities while maintaining data privacy. By aggregating insights from different sources, this approach seeks to improve the accuracy of mortality prediction for COVID-19 patients, a critical aspect of healthcare management during the pandemic. The paper underscores the significance of federated learning in addressing urgent healthcare challenges, such as COVID-19, by leveraging distributed data to enhance patient outcomes without compromising data security, exemplifying the potential of federated learning in public health emergencies.

In another work, Vaid et al. in [92] presented a significant advancement in healthcare analytics by applying federated learning to enhance mortality prediction for COVID-19 patients. Their research leverages electronic health records (EHRs) from multiple healthcare facilities to collaboratively train predictive models while preserving patient data privacy. By aggregating insights from diverse sources, the study aims to improve the accuracy of mortality prediction, a crucial aspect of managing patients during the COVID-19 pandemic. The paper underscores the potential of federated learning in addressing pressing healthcare challenges, demonstrating how it can harness distributed data to enhance patient outcomes while ensuring data security and privacy in the context of a global health crisis.

7. Challenges of Edge AI in Early Health Prediction and Future Directions

This section primarily addresses **RQ-4**, highlighting the challenges in Edge AI for early health prediction and suggesting future directions to overcome them.

Notwithstanding the potential benefits of using AI and edge computing for early health prediction, it remains essential to address a number of challenges and limitations to ensure optimal application.

7.1. Privacy and Security

Protecting personal health data is a major barrier to using AI and edge computing to predict health outcomes in advance. With the proliferation of wearable devices and other connected technologies, there are growing concerns about the security, privacy, and protection of personal health data [93,94].

7.1.1. Privacy

Here are the key aspects to consider:

- **Data collection:** Edge AI relies on collecting vast amounts of personal health data, including medical records, biometric data, and lifestyle information. Safeguarding the privacy of this sensitive data is crucial.
- **Informed consent:** Obtaining informed consent from individuals for data collection and usage can be challenging. Clear communication about the purpose, risks, and benefits of using their data in early health prediction is essential.
- **Data anonymization:** Anonymizing health data is critical to protecting individuals' privacy. However, achieving complete anonymity while maintaining data utility for accurate predictions can be a complex task.

Future research directions to address the privacy challenges in Edge AI for early health prediction include:

- **Privacy-preserving algorithms:** Developing advanced algorithms that can perform accurate health prediction while preserving privacy is a promising research direction. Techniques like federated learning and differential privacy can help achieve this goal.
- **Consent mechanisms:** Exploring innovative consent mechanisms that empower individuals to have more control over their data and make informed decisions regarding data usage in early health prediction.
- **Privacy-enhancing technologies:** Investigating the use of privacy-enhancing technologies such as secure multi-party computation (SMPC), differential privacy, and homomorphic encryption to enable analysis of sensitive health data without compromising privacy.
- **Policy and regulation:** Establishing comprehensive policies and regulations that govern the collection, storage, and usage of health data in Edge AI systems. This includes ensuring compliance with privacy laws and implementing ethical frameworks.

7.1.2. Security

Here are some key aspects to consider:

Security challenges:

- **Data breaches:** Edge AI systems store and process sensitive health data, making them potential targets for cyberattacks. Robust security measures must be in place to prevent unauthorized access and data breaches.
- **Secure communication:** Edge AI systems often rely on transmitting data between devices and cloud servers. Ensuring secure communication channels and encryption protocols is vital to protect data during transmission.
- **Adversarial attacks:** Edge AI models can be vulnerable to adversarial attacks, where malicious actors manipulate input data to deceive the system. Developing robust defenses against such attacks is crucial to maintaining the integrity of early health prediction.

Future research directions to address the security challenges in Edge AI for early health prediction include:

- **Robust authentication:** Exploring advanced authentication mechanisms, such as multi-factor authentication and biometrics, to enhance the security of Edge AI systems and prevent unauthorized access.
- **Secure hardware and firmware:** Enhancing the security of edge devices, such as wearables and IoT devices, by implementing secure hardware components and regularly updating the firmware to mitigate potential vulnerabilities.
- **Intrusion detection and prevention:** Developing sophisticated intrusion detection and prevention systems specifically designed for Edge AI in early health prediction. This helps detect and mitigate potential security breaches.
- **Resilient AI models:** Designing AI models that are resilient to adversarial attacks and can identify and reject manipulated input data, ensuring the accuracy and trustworthiness of early health predictions.

7.2. Data Quality and Accuracy

Another obstacle AI algorithms face is the quality and accuracy of the data they analyze. Information coming from wearable devices and other sources can be inaccurate, inconsistent, or erroneous. All of this can lead to inaccurate predictions and false warnings [95,96]. Here are the key aspects to consider:

- **Labeling and Annotation:** High-quality labeled data are crucial for training accurate and reliable AI models. However, in early health prediction, obtaining ground truth labels can be challenging. This may require expert knowledge, clinical validation, or long-term follow-up to confirm the accuracy of predictions.
- **Data Variability:** Health data collected at the edge can be highly variable due to various factors such as device sensors, user behavior, environmental conditions, and data collection protocols [97]. This variability can impact the performance and generalizability of AI models, making it essential to address data variability challenges.
- **Data Imbalance:** In health prediction tasks, class imbalances are common, where certain health conditions or outcomes may be significantly less frequent than others [98]. This can lead to biased models that perform poorly on minority classes. Techniques such as data augmentation, oversampling, or ensemble methods need to be explored to address data imbalance challenges.
- **Data Quality Control:** Ensuring the quality and reliability of collected health data is critical. Sensor errors, noise, missing values, and data corruption can negatively impact the performance of AI models [99]. Quality control mechanisms, data preprocessing techniques, and outlier detection methods need to be developed to improve data quality.

Future research directions to address data quality challenges in Edge AI for early health prediction include:

- **Data Augmentation Techniques:** Exploring data augmentation methods specifically tailored for health data to increase the diversity and size of the training dataset, improving model robustness and generalizability.
- **Collaborative Data Sharing:** Encouraging collaboration and data sharing among healthcare institutions while ensuring privacy and security. Pooling diverse datasets can help overcome data variability and improve the representativeness of AI models.
- **Data Quality Assessment Frameworks:** Developing standardized frameworks to assess the quality and reliability of health data collected at the edge. This can involve metrics, guidelines, and best practices for data collection, labeling, and preprocessing.

7.3. Model Bias

AI algorithms may exhibit bias when trained on datasets with inherent biases. This can result in predictions that are also biased, leading to unfair results, especially in early-stage health prediction, where the detrimental effects of erroneous predictions can be quite significant. Here are the key aspects to consider:

- **Dataset Bias:** AI models are trained on datasets that may not be representative of the diverse population they aim to serve. Biases in the data, such as under-representation of certain demographics or health conditions, can lead to biased predictions and inequitable healthcare outcomes [100].
- **Algorithmic Bias:** Biases can also be introduced through the design and implementation of AI algorithms [101]. If the training data contain inherent biases or if the algorithm itself is biased, this can perpetuate and amplify existing disparities in healthcare.
- **Interpretability and Transparency:** Lack of interpretability and transparency in AI models can make it challenging to identify and address biases. Understanding how the model makes predictions and uncovering any underlying biases require explainable AI techniques [102].

Future research directions to address model bias challenges in Edge AI for early health prediction include:

- **Bias Detection and Mitigation:** Designing methods to detect and quantify bias in AI models and developing techniques to mitigate its impact. This can include techniques like debiasing algorithms, data augmentation, and fairness-aware feature selection.
- **Diverse and Representative Datasets:** Collecting and using diverse and representative datasets that encompass different demographics, health conditions, and socioeconomic backgrounds. This helps reduce dataset bias and improves the generalizability of AI models.
- **Ethical Guidelines and Regulations:** Establishing clear ethical guidelines and regulations for the development and deployment of AI in healthcare. This can help address biases, promote fairness, and ensure accountability and transparency in early health prediction.
- **Collaboration and Interdisciplinary Research:** Encouraging collaboration between AI researchers, healthcare professionals, ethicists, and policymakers to collectively address model bias challenges. Interdisciplinary research can provide a holistic perspective and help develop comprehensive solutions.

7.4. Interoperability

Ensuring that different devices, systems, and data sources work together is a challenge. Ensuring smooth data sharing and analysis across many platforms and devices is critical. This cohesion is critical to enabling effective prediction of health outcomes. Here are the key aspects to consider:

- **Heterogeneous Data Sources:** In early health prediction, data are collected from various sources such as wearable devices, electronic health records, and sensors. These sources often use different data formats, protocols, and standards, making it challenging to integrate and analyze data seamlessly.
- **Data Integration and Fusion:** Aggregating and fusing data from multiple sources are essential for building comprehensive AI models. However, the lack of interoperability can hinder this process, leading to difficulties in harmonizing and combining heterogeneous data effectively.
- **Connectivity and Communication:** Edge AI systems rely on efficient communication between edge devices, cloud infrastructure, and central servers. Interoperability issues can arise due to differences in communication protocols, network connectivity, and compatibility between devices and systems.
- **Privacy and Security Concerns:** Interoperability can raise privacy and security concerns, especially when sensitive health data are shared or exchanged between different systems. Ensuring secure data transmission, secure access controls, and compliance with privacy regulations becomes crucial.

Future research directions to address interoperability challenges in Edge AI for early health prediction include:

- **Standardization and Data Formats:** Developing standardized data formats, protocols, and interfaces for health data exchange. This enables seamless interoperability between different systems and facilitates data integration and analysis.
- **Ontologies and Semantic Interoperability:** Utilizing ontologies and semantic models to establish a common understanding of health data and data relationships. This promotes interoperability by enabling efficient data integration, data mapping, and knowledge sharing across different platforms.
- **Interoperability Frameworks and Middleware:** Designing interoperability frameworks and middleware that facilitate data exchange, communication, and integration among diverse edge devices, cloud systems, and healthcare infrastructure. These frameworks can provide standard APIs, data transformation capabilities, and connectivity support.
- **Secure Data Sharing and Privacy-Preserving Mechanisms:** Developing secure and privacy-preserving methods for data sharing and exchange. Techniques such as federated learning, differential privacy, and encryption can enable collaborative analysis while protecting sensitive health information.
- **Collaborative Ecosystem:** Encouraging collaboration between stakeholders, including researchers, healthcare providers, device manufacturers, and policymakers, to establish interoperability standards, guidelines, and best practices. Collaboration can drive the adoption of interoperable solutions and facilitate the seamless integration of Edge AI systems in healthcare.

7.5. Integration with Existing Health Systems

Integrating Edge AI technologies with existing health systems is imperative to realize their full potential. This effort will require collaboration among healthcare providers, technology companies, and researchers, all working together to ensure the seamless integration of AI and edge computing technologies into the existing healthcare infrastructure. Here are the key aspects to consider:

- **Compatibility with Legacy Systems:** Many healthcare organizations have established legacy systems and infrastructure that may not be designed to integrate with newer Edge AI technologies. These systems often have different data formats, protocols, and interfaces, making it difficult to seamlessly incorporate Edge AI solutions.
- **Data Synchronization and Exchange:** Integrating Edge AI for early health prediction requires smooth data synchronization and exchange between edge devices, cloud platforms, and existing healthcare systems. Ensuring data consistency, real-time updates, and bidirectional communication becomes crucial for effective integration.
- **Workflow and Process Alignment:** Integrating Edge AI into existing healthcare systems requires careful consideration of workflow and process alignment. The introduction of Edge AI should seamlessly fit into existing clinical workflows, ensuring minimal disruption and maximizing efficiency.

Future research directions to address the integration challenge with existing healthcare systems in Edge AI for early health prediction include:

- **Interoperability Standards:** Developing industry-wide interoperability standards and guidelines that facilitate the integration of Edge AI technologies with existing healthcare systems. These standards should address data formats, communication protocols, and interoperability interfaces to ensure seamless integration [103].
- **Application Programming Interfaces (APIs):** Creating standardized APIs that enable easy integration between Edge AI systems and existing healthcare systems. These APIs should provide clear specifications for data exchange, functionality access, and system integration, simplifying the integration process.
- **Middleware and Integration Platforms:** Designing middleware and integration platforms specifically tailored for integrating Edge AI into existing healthcare systems.

These platforms can provide tools, libraries, and frameworks that facilitate data integration, process alignment, and workflow integration.

- **Proof-of-Concept Projects:** Conducting pilot projects and proof-of-concept studies to demonstrate the feasibility and benefits of integrating Edge AI in early health prediction with existing healthcare systems. These projects can showcase successful integration strategies, identify challenges, and provide insights for future implementation.
- **Collaboration and Partnerships:** Encouraging collaboration and partnerships between Edge AI solution providers and healthcare organizations. Close collaboration can help identify integration requirements, co-design solutions, and establish a mutually beneficial integration process.

8. Conclusions

Integration of AI and edge computing in early health prediction has transformative potential. These technologies enable real-time data analysis, aiding health risk detection and prevention. This systematic review analyzes Edge AI's use in early health prediction and chronic and infectious disease detection, utilizing predictive modeling, machine learning, deep learning, and federated learning. As well as enhancing prediction and maintaining privacy, federated learning collaboratively trains AI models on distributed edge devices, allowing local processing for instant decision making and reduced latency. However, challenges like privacy, security, data quality, model bias, interoperability, and system integration must be addressed. Continued research, collaboration, and strategic investments are imperative for the improvement and widespread adoption of Edge AI in healthcare. Furthermore, staying informed about emerging trends and ethical concerns is essential as Edge AI evolves in early health prediction.

Funding: This work was supported by the United Arab Emirates (UAE) University UAEU Program for Advanced Research (UPAR) under Grant G00003443.

Conflicts of Interest: The author declares no conflict of interest.

Sample Availability: Samples of the compounds are available from the authors.

References

1. Yu, K.H.; Beam, A.L.; Kohane, I.S. Artificial intelligence in healthcare. *Nat. Biomed. Eng.* **2018**, *2*, 719–731. [CrossRef]
2. Amisha; Malik, P.; Pathania, M.; Rathaur, V.K. Overview of artificial intelligence in medicine. *J. Family Med. Prim. Care* **2019**, *8*, 2328–2331.
3. Bohr, A.; Memarzadeh, K. *Artificial Intelligence in Healthcare*; Academic Press: Cambridge, MA, USA, 2020.
4. Rong, G.; Mendez, A.; Bou Assi, E.; Zhao, B.; Sawan, M. Artificial Intelligence in Healthcare: Review and Prediction Case Studies. *Proc. Est. Acad. Sci. Eng.* **2020**, *6*, 291–301. [CrossRef]
5. Shi, W.; Cao, J.; Zhang, Q.; Li, Y.; Xu, L. Edge Computing: Vision and Challenges. *IEEE Internet Things J.* **2016**, *3*, 637–646. [CrossRef]
6. Satyanarayanan, M. The Emergence of Edge Computing. *Computer* **2017**, *50*, 30–39. [CrossRef]
7. Ray, P.P.; Dash, D.; De, D. Edge computing for Internet of Things: A survey, e-healthcare case study and future direction. *J. Netw. Comput. Appl.* **2019**, *140*, 1–22. [CrossRef]
8. Sodhro, A.H.; Luo, Z.; Sangaiah, A.K.; Baik, S.W. Mobile edge computing based QoS optimization in medical healthcare applications. *Int. J. Inf. Manag.* **2019**, *45*, 308–318. [CrossRef]
9. Kamruzzaman, M.M.; Alrashdi, I.; Alqazzaz, A. New Opportunities, Challenges, and Applications of Edge-AI for Connected Healthcare in Internet of Medical Things for Smart Cities. *J. Healthc. Eng.* **2022**, *2022*, 2950699. [CrossRef]
10. Rathi, V.K.; Rajput, N.K.; Mishra, S.; Grover, B.A.; Tiwari, P.; Jaiswal, A.K.; Hossain, M.S. An edge AI-enabled IoT healthcare monitoring system for smart cities. *Comput. Electr. Eng.* **2021**, *96*, 107524. [CrossRef]
11. Dong, J.; Feng, T.; Thapa-Chhetry, B.; Cho, B.G.; Shum, T.; Inwald, D.P.; Newth, C.J.L.; Vaidya, V.U. Machine learning model for early prediction of acute kidney injury (AKI) in pediatric critical care. *Crit. Care* **2021**, *25*, 288. [CrossRef]
12. Layeghian Javan, S.; Sepehri, M.M.; Layeghian Javan, M.; Khatibi, T. An intelligent warning model for early prediction of cardiac arrest in sepsis patients. *Comput. Methods Programs Biomed.* **2019**, *178*, 47–58. [CrossRef]
13. Hong, W.; Lee, W.G. Wearable sensors for continuous oral cavity and dietary monitoring toward personalized healthcare and digital medicine. *Analyst* **2021**, *145*, 7796–7808. [CrossRef] [PubMed]

14. Thipsawat, S. Early detection of diabetic nephropathy in patient with type 2 diabetes mellitus: A review of the literature. *Diab. Vasc. Dis. Res.* **2021**, *18*, 14791641211058856. [CrossRef] [PubMed]
15. Wander, G.S.; Bansal, M.; Kasliwal, R.R. Prediction and early detection of cardiovascular disease in South Asians with diabetes mellitus. *Diabetes Metab. Syndr.* **2020**, *14*, 385–393. [CrossRef] [PubMed]
16. Schiffman, J.D.; Fisher, P.G.; Gibbs, P. Early detection of cancer: Past, present, and future. *Am. Soc. Clin. Oncol. Educ. Book* **2015**, 57–65. [CrossRef] [PubMed]
17. Xu, L.; Zhu, S.; Wen, N. Deep reinforcement learning and its applications in medical imaging and radiation therapy: A survey. *Phys. Med. Biol.* **2022**, *67*, 22TR02. [CrossRef]
18. Montagnon, E.; Cerny, M.; Cadrin-Chênevert, A.; Hamilton, V.; Derennes, T.; Ilinca, A.; Vandenbroucke-Menu, F.; Turcotte, S.; Kadoury, S.; Tang, A. Deep learning workflow in radiology: A primer. *Insights Imaging* **2020**, *11*, 22. [CrossRef]
19. Cao, C.; Liu, F.; Tan, H.; Song, D.; Shu, W.; Li, W.; Zhou, Y.; Bo, X.; Xie, Z. Deep Learning and Its Applications in Biomedicine. *Genom. Proteom. Bioinform.* **2018**, *16*, 17–32. [CrossRef]
20. Moshawrab, M.; Adda, M.; Bouzouane, A.; Ibrahim, H.; Raad, A. Reviewing Federated Machine Learning and Its Use in Diseases Prediction. *Sensors* **2023**, *23*, 2112. [CrossRef]
21. Moshawrab, M.; Adda, M.; Bouzouane, A.; Ibrahim, H.; Raad, A. Smart Wearables for the Detection of Cardiovascular Diseases: A Systematic Literature Review. *Sensors* **2023**, *23*, 828. [CrossRef]
22. Garcia Lopez, P.; Montresor, A.; Epema, D.; Datta, A.; Higashino, T.; Iamnitchi, A.; Barcellos, M.; Felber, P.; Riviere, E. Edge-centric Computing: Vision and Challenges. *SIGCOMM Comput. Commun. Rev.* **2015**, *45*, 37–42. [CrossRef]
23. Grand View Research. Edge Computing Market Size, Share & Growth Report. 2023–2030. Available online: https://www.grandviewresearch.com/industry-analysis/edge-computing-market (accessed on 16 February 2023).
24. Mohsen, F.; Ali, H.; El Hajj, N.; Shah, Z. Artificial intelligence-based methods for fusion of electronic health records and imaging data. *Sci. Rep.* **2022**, *12*, 17981. [CrossRef]
25. Briganti, G.; Le Moine, O. Artificial Intelligence in Medicine: Today and Tomorrow. *Front. Med.* **2020**, *7*, 27. [CrossRef]
26. Lavecchia, A. Deep learning in drug discovery: Opportunities, challenges and future prospects. *Drug Discov. Today* **2019**, *24*, 2017–2032. [CrossRef] [PubMed]
27. Mak, K.K.; Pichika, M.R. Artificial intelligence in drug development: Present status and future prospects. *Drug Discov. Today* **2019**, *24*, 773–780. [CrossRef]
28. Jiménez-Luna, J.; Grisoni, F.; Weskamp, N.; Schneider, G. Artificial intelligence in drug discovery: Recent advances and future perspectives. *Expert Opin. Drug Discov.* **2021**, *16*, 949–959. [CrossRef] [PubMed]
29. Topol, E.J. High-performance medicine: The convergence of human and artificial intelligence. *Nat. Med.* **2019**, *25*, 44–56. [CrossRef]
30. Adegun, A.A.; Viriri, S.; Ogundokun, R.O. Deep Learning Approach for Medical Image Analysis. *Comput. Intell. Neurosci.* **2021**, *2021*, 6215281. [CrossRef]
31. Fraiwan, M.; Audat, Z.; Fraiwan, L.; Manasreh, T. Using deep transfer learning to detect scoliosis and spondylolisthesis from X-ray images. *PLoS ONE* **2022**, *17*, e0267851. [CrossRef]
32. Panayides, A.S.; Amini, A.; Filipovic, N.D.; Sharma, A.; Tsaftaris, S.A.; Young, A.; Foran, D.; Do, N.; Golemati, S.; Kurc, T.; et al. AI in Medical Imaging Informatics: Current Challenges and Future Directions. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 1837–1857. [CrossRef]
33. Coronato, A.; Naeem, M.; De Pietro, G.; Paragliola, G. Reinforcement learning for intelligent healthcare applications: A survey. *Artif. Intell. Med.* **2020**, *109*, 101964. [CrossRef] [PubMed]
34. Singh, A.; Chatterjee, K. Edge computing based secure health monitoring framework for electronic healthcare system. *Clust. Comput.* **2022**, *26*, 1205–1220. [CrossRef]
35. Klonoff, D.C. Fog Computing and Edge Computing Architectures for Processing Data From Diabetes Devices Connected to the Medical Internet of Things. *J. Diabetes Sci. Technol.* **2017**, *11*, 647–652. [CrossRef]
36. American Diabetes Association. What Is a Smart Insulin Pen? 2023. Available online: https://diabetes.org/about-diabetes/devices-technology/smart-insulin-pen (accessed on 18 February 2023).
37. Kamga, P.; Mostafa, R.; Zafar, S. The Use of Wearable ECG Devices in the Clinical Setting: A Review. *Curr. Emerg. Hosp. Med. Rep.* **2022**, *10*, 67–72. [CrossRef] [PubMed]
38. Findair. IoT with the Use of Smart Inhalers-Resources-Blog. 2021. Available online: http://findair.eu/resources/blog/iot-with-the-use-of-smart-inhalers/ (accessed on 18 February 2023).
39. Dey, A.; Haque, K.A.; Nayan, A.A.; Kibria, M.G. IoT Based Smart Inhaler For Context-Aware Service Provisioning. In Proceedings of the 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), Dhaka, Bangladesh, 28–29 November 2020; pp. 410–415.
40. Mohan, H.M.; Anitha, S.; Chai, R.; Ling, S.H. Edge Artificial Intelligence: Real-Time Noninvasive Technique for Vital Signs of Myocardial Infarction Recognition Using Jetson Nano. *Adv. Hum.-Comput. Interact.* **2021**, *2021*, 6483003. [CrossRef]
41. Islam, K.; Alam, F.; Zahid, A.I.; Khan, M.M.; InamAbbasi, M. Internet of Things- (IoT-) Based Real-Time Vital Physiological Parameter Monitoring System for Remote Asthma Patients. *Proc. Int. Wirel. Commun. Mob. Comput. Conf.* **2022**, *2022*, 1191434. [CrossRef]

42. Paul, A.; Pinjari, H.; Hong, W.H.; Seo, H.C.; Rho, S. Fog Computing-Based IoT for Health Monitoring System. *J. Sens.* **2018**, *2018*, 1386470. [CrossRef]
43. Jiang, Q.; Zhou, X.; Wang, R.; Ding, W.; Chu, Y.; Tang, S.; Jia, X.; Xu, X. Intelligent monitoring for infectious diseases with fuzzy systems and edge computing: A survey. *Appl. Soft Comput.* **2022**, *123*, 108835. [CrossRef]
44. Zhu, G.; Li, J.; Meng, Z.; Yu, Y.; Li, Y.; Tang, X.; Dong, Y.; Sun, G.; Zhou, R.; Wang, H.; et al. Learning from Large-Scale Wearable Device Data for Predicting the Epidemic Trend of COVID-19. *Discrete Dyn. Nat. Soc.* **2020**, *2020*, 6152041. [CrossRef]
45. Alanazi, R. Identification and Prediction of Chronic Diseases Using Machine Learning Approach. *J. Healthc. Eng.* **2022**, *2022*, 2826127. [CrossRef]
46. Manjulatha, B.; Pabboju, S. An Ensemble Model for Predicting Chronic Diseases Using Machine Learning Algorithms. In *Smart Computing Techniques and Applications*; Springer: Singapore, 2021; pp. 337–345.
47. Chhabra, D.; Juneja, M.; Chutani, G. An efficient ensemble based machine learning approach for predicting Chronic Kidney Disease. *Curr. Med. Imaging Rev.* **2023**, *Online ahead of print*.
48. Ibrahim, A.; Gamble, P.; Jaroensri, R.; Abdelsamea, M.M.; Mermel, C.H.; Chen, P.H.C.; Rakha, E.A. Artificial intelligence in digital breast pathology: Techniques and applications. *Breast* **2020**, *49*, 267–273. [CrossRef]
49. Jan, Z.; El Assadi, F.; Abd-Alrazaq, A.; Jithesh, P.V. Artificial Intelligence for the Prediction and Early Diagnosis of Pancreatic Cancer: Scoping Review. *J. Med. Internet Res.* **2023**, *25*, e44248. [CrossRef]
50. Huang, J.D.; Wang, J.; Ramsey, E.; Leavey, G.; Chico, T.J.A.; Condell, J. Applying Artificial Intelligence to Wearable Sensor Data to Diagnose and Predict Cardiovascular Disease: A Review. *Sensors* **2022**, *22*, 8002. [CrossRef]
51. Tran, N.K.; Kretsch, C.; LaValley, C.; Rashidi, H.H. Machine learning and artificial intelligence for the diagnosis of infectious diseases in immunocompromised patients. *Curr. Opin. Infect. Dis.* **2023**, *36*, 235. [CrossRef]
52. Parums, D.V. Editorial: Infectious Disease Surveillance Using Artificial Intelligence (AI) and its Role in Epidemic and Pandemic Preparedness. *Med. Sci. Monit.* **2023**, *29*, e941209-1–e941209-4. [CrossRef]
53. Shah, W.; Aleem, M.; Iqbal, M.A.; Islam, M.A.; Ahmed, U.; Srivastava, G.; Lin, J.C.W. A Machine-Learning-Based System for Prediction of Cardiovascular and Chronic Respiratory Diseases. *J. Healthc. Eng.* **2021**, *2021*, 2621655. [CrossRef]
54. Ramesh, T.R.; Lilhore, U.K.; Poongodi, M.; Simaiya, S.; Kaur, A.; Hamdi, M. Predictive analysis of heart diseases with machine learning approaches. *Malays. J. Comput. Sci.* **2022**, *2022*, 132–148.
55. Chaki, J.; Thillai Ganesh, S.; Cidham, S.K.; Ananda Theertan, S. Machine learning and artificial intelligence based Diabetes Mellitus detection and self-management: A systematic review. *J. King Saud Univ. Comput. Inf. Sci.* **2022**, *34*, 3204–3225. [CrossRef]
56. Dayan, I.; Roth, H.R.; Zhong, A.; Harouni, A.; Gentili, A.; Abidin, A.Z.; Liu, A.; Costa, A.B.; Wood, B.J.; Tsai, C.S.; et al. Federated learning for predicting clinical outcomes in patients with COVID-19. *Nat. Med.* **2021**, *27*, 1735–1743. [CrossRef]
57. Kengne, A.P.; Masconi, K.; Mbanya, V.N.; Lekoubou, A.; Echouffo-Tcheugui, J.B.; Matsha, T.E. Risk predictive modelling for diabetes and cardiovascular disease. *Crit. Rev. Clin. Lab. Sci.* **2014**, *51*, 1–12. [CrossRef]
58. Su, T.L.; Jaki, T.; Hickey, G.L.; Buchan, I.; Sperrin, M. A review of statistical updating methods for clinical prediction models. *Stat. Methods Med. Res.* **2018**, *27*, 185–197. [CrossRef] [PubMed]
59. Chapman, B.P.; Lin, F.; Roy, S.; Benedict, R.H.B.; Lyness, J.M. Health risk prediction models incorporating personality data: Motivation, challenges, and illustration. *Personal. Disord.* **2019**, *10*, 46–58. [CrossRef] [PubMed]
60. Beniczky, S.; Karoly, P.; Nurse, E.; Ryvlin, P.; Cook, M. Machine learning and wearable devices of the future. *Epilepsia* **2021**, *62* (Suppl. S2), S116–S124. [CrossRef] [PubMed]
61. Sabry, F.; Eltaras, T.; Labda, W.; Alzoubi, K.; Malluhi, Q. Machine Learning for Healthcare Wearable Devices: The Big Picture. *J. Healthc. Eng.* **2022**, *2022*, 4653923. [CrossRef]
62. Castelli Gattinara Di Zubiena, F.; Menna, G.; Mileti, I.; Zampogna, A.; Asci, F.; Paoloni, M.; Suppa, A.; Del Prete, Z.; Palermo, E. Machine Learning and Wearable Sensors for the Early Detection of Balance Disorders in Parkinson's Disease. *Sensors* **2022**, *22*, 9903. [CrossRef]
63. Chang, V.; Ganatra, M.A.; Hall, K.; Golightly, L.; Xu, Q.A. An assessment of machine learning models and algorithms for early prediction and diagnosis of diabetes using health indicators. *Healthc. Anal.* **2022**, *2*, 100118. [CrossRef]
64. Jenifer, A.; Jeba, G.; Paulraj, L.; Kumar, N.; Yuvaraj, T.; Alen, G.; Rozario, P.; Amoli, R. Edge-based heart disease prediction device using internet of things. In Proceedings of the 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 9–11 May 2022.
65. Kavitha, C.; Mani, V.; Srividhya, S.R.; Khalaf, O.I.; Tavera Romero, C.A. Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models. *Front. Public Health* **2022**, *10*, 853294. [CrossRef]
66. Futoma, J.; Morris, J.; Lucas, J. A comparison of models for predicting early hospital readmissions. *J. Biomed. Inform.* **2015**, *56*, 229–238. [CrossRef]
67. Guo, A.; Mazumder, N.R.; Ladner, D.P.; Foraker, R.E. Predicting mortality among patients with liver cirrhosis in electronic health records with machine learning. *PLoS ONE* **2021**, *16*, e0256428. [CrossRef]
68. Mohammadi, R.; Jain, S.; Agboola, S.; Palacholla, R.; Kamarthi, S.; Wallace, B.C. Learning to Identify Patients at Risk of Uncontrolled Hypertension Using Electronic Health Records Data. *AMIA Jt. Summits Transl. Sci. Proc.* **2019**, *2019*, 533–542.
69. Nikhila. Chronic Kidney Disease Prediction using Machine Learning Ensemble Algorithm. In Proceedings of the 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 19–20 February 2021; pp. 476–480.

70. Khalid, H.; Khan, A.; Zahid Khan, M.; Mehmood, G.; Shuaib Qureshi, M. Machine Learning Hybrid Model for the Prediction of Chronic Kidney Disease. *Comput. Intell. Neurosci.* **2023**, *2023*, 9266889. [CrossRef]
71. Tomašev, N.; Glorot, X.; Rae, J.W.; Zielinski, M.; Askham, H.; Saraiva, A.; Mottram, A.; Meyer, C.; Ravuri, S.; Protsyuk, I.; et al. Developing Deep Learning Continuous Risk Models for Early Adverse Event Prediction in Electronic Health Records: An AKI Case Study. *Protoc. Exch.* **2019**. [CrossRef]
72. Antunes, R.S.; André da Costa, C.; Küderle, A.; Yari, I.A.; Eskofier, B. Federated Learning for Healthcare: Systematic Review and Architecture Proposal. *ACM Trans. Intell. Syst. Technol.* **2022**, *13*, 1–23. [CrossRef]
73. Nazir, S.; Kaleem, M. Federated Learning for Medical Image Analysis with Deep Neural Networks. *Diagnostics* **2023**, *13*, 1532. [CrossRef] [PubMed]
74. Sohan, M.F.; Basalamah, A. A systematic review on federated learning in medical image analysis. *IEEE Access* **2023**, *11*, 28628–28644. [CrossRef]
75. Ali, M.; Naeem, F.; Tariq, M.; Kaddoum, G. Federated Learning for Privacy Preservation in Smart Healthcare Systems: A Comprehensive Survey. *IEEE J. Biomed. Health Inform.* **2022**, *27*, 778–789. [CrossRef] [PubMed]
76. Iwagami, M.; Matsui, H. Introduction to Clinical Prediction Models. *Ann. Clin. Epidemiol.* **2022**, *4*, 72–80. [CrossRef]
77. Jiang, L.Y.; Liu, X.C.; Nejatian, N.P.; Nasir-Moin, M.; Wang, D.; Abidin, A.; Eaton, K.; Riina, H.A.; Laufer, I.; Punjabi, P.; et al. Health system-scale language models are all-purpose prediction engines. *Nature* **2023**, *619*, 357–362. [CrossRef]
78. Karstoft, K.I.; Eskelund, K.; Gradus, J.L.; Andersen, S.B.; Nissen, L.R. Early prediction of mental health problems following military deployment: Integrating pre- and post-deployment factors in neural network models. *J. Psychiatr. Res.* **2023**, *163*, 109–117. [CrossRef]
79. Rinderknecht, M.D.; Klopfenstein, Y. Predicting critical state after COVID-19 diagnosis: Model development using a large US electronic health record dataset. *NPJ Digit. Med.* **2021**, *4*, 113. [CrossRef]
80. Tomašev, N.; Glorot, X.; Rae, J.W.; Zielinski, M.; Askham, H.; Saraiva, A.; Mottram, A.; Meyer, C.; Ravuri, S.; Protsyuk, I.; et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* **2019**, *572*, 116–119. [CrossRef] [PubMed]
81. Nguyen, D.C.; Pham, Q.V.; Pathirana, P.N.; Ding, M.; Seneviratne, A.; Lin, Z.; Dobre, O.; Hwang, W.J. Federated Learning for Smart Healthcare: A Survey. *ACM Comput. Surv.* **2022**, *55*, 1–37. [CrossRef]
82. Yang, Q.; Liu, Y.; Chen, T.; Tong, Y. Federated Machine Learning: Concept and Applications. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, 1–19. [CrossRef]
83. Beaussart, M.; Grimberg, F.; Hartley, M.A.; Jaggi, M. WAFFLE: Weighted averaging for personalized federated learning. *arXiv* **2021**, arXiv:2110.06978.
84. Kalapaaking, A.P.; Khalil, I.; Yi, X. Blockchain-based Federated Learning with SMPC Model Verification against Poisoning Attack for Healthcare Systems. *arXiv* **2023**, arXiv:2304.13360.
85. Randl, K.; Armengol, N.L.; Mondrejevski, L.; Miliou, I. Early prediction of the risk of ICU mortality with Deep Federated Learning. In Proceedings of the 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS), L'Aquila, Italy, 22–24 June 2023; pp. 706–711.
86. Yaqoob, M.M.; Nazir, M.; Khan, M.A.; Qureshi, S.; Al-Rasheed, A. Hybrid Classifier-Based Federated Learning in Health Service Providers for Cardiovascular Disease Prediction. *NATO Adv. Sci. Inst. Ser. E Appl. Sci.* **2023**, *13*, 1911. [CrossRef]
87. Bharathi, K.; Dhavamani, M.; Niranjan, K. A federated learning based approach for heart disease prediction. In Proceedings of the 2022 6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 29–31 March 2022.
88. Nandhini, J.M.; Joshi, S.; Anuratha, K. Federated learning based prediction of chronic kidney diseases. In Proceedings of the 2022 1st International Conference on Computational Science and Technology (ICCST), Chennai, India, 9–10 November 2022.
89. Farooq, K.; Syed, H.J.; Alqahtani, S.O.; Nagmeldin, W.; Ibrahim, A.O.; Gani, A. Blockchain Federated Learning for In-Home Health Monitoring. *Electronics* **2022**, *12*, 136. [CrossRef]
90. Chen, B.; Chen, T.; Zeng, X.; Zhang, W.; Lu, Q.; Hou, Z.; Zhou, J.; Helal, S. DFML: Dynamic Federated Meta-Learning for Rare Disease Prediction. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2023**, *Online ahead of print*.
91. Vaid, A.; Jaladanki, S.K.; Xu, J.; Teng, S.; Kumar, A.; Lee, S.; Somani, S.; Paranjpe, I.; De Freitas, J.K.; Wanyan, T.; et al. Federated Learning of Electronic Health Records to Improve Mortality Prediction in Hospitalized Patients With COVID-19: Machine Learning Approach. *JMIR Med. Inform.* **2021**, *9*, e24207. [CrossRef] [PubMed]
92. Vaid, A.; Jaladanki, S.K.; Xu, J.; Teng, S.; Kumar, A.; Lee, S.; Somani, S.; Paranjpe, I.; De Freitas, J.K.; Wanyan, T.; et al. Federated Learning of Electronic Health Records Improves Mortality Prediction in Patients Hospitalized with COVID-19. *medRxiv* **2020**, *Preprint*.
93. Jiang, D.; Shi, G. Research on Data Security and Privacy Protection of Wearable Equipment in Healthcare. *J. Healthc. Eng.* **2021**, *2021*, 6656204. [CrossRef]
94. Arias, O.; Wurm, J.; Hoang, K.; Jin, Y. Privacy and Security in Internet of Things and Wearable Devices. *IEEE Trans. Multi-Scale Comput. Syst.* **2015**, *1*, 99–109. [CrossRef]
95. Böttcher, S.; Vieluf, S.; Bruno, E.; Joseph, B.; Epitashvili, N.; Biondi, A.; Zabler, N.; Glasstetter, M.; Dümpelmann, M.; Van Laerhoven, K.; et al. Data quality evaluation in wearable monitoring. *Sci. Rep.* **2022**, *12*, 21412. [CrossRef]
96. Poli, A.; Cosoli, G.; Scalise, L.; Spinsante, S. Impact of Wearable Measurement Properties and Data Quality on ADLs Classification Accuracy. *IEEE Sens. J.* **2021**, *21*, 14221–14231. [CrossRef]

97. Souza, J.; Caballero, I.; Vasco Santos, J.; Lobo, M.; Pinto, A.; Viana, J.; Sáez, C.; Lopes, F.; Freitas, A. Multisource and temporal variability in Portuguese hospital administrative datasets: Data quality implications. *J. Biomed. Inform.* **2022**, *136*, 104242. [CrossRef]
98. Khushi, M.; Shaukat, K.; Alam, T.M.; Hameed, I.A.; Uddin, S.; Luo, S.; Yang, X.; Reyes, M.C. A Comparative Performance Analysis of Data Resampling Methods on Imbalance Medical Data. *IEEE Access* **2021**, *9*, 109960–109975. [CrossRef]
99. Isgut, M.; Gloster, L.; Choi, K.; Venugopalan, J.; Wang, M.D. Systematic Review of Advanced AI Methods for Improving Healthcare Data Quality in Post COVID-19 Era. *IEEE Rev. Biomed. Eng.* **2023**, *16*, 53–69. [CrossRef]
100. Vaughn, J.; Baral, A.; Vadari, M.; Boag, W. Dataset Bias in Diagnostic AI systems: Guidelines for Dataset Collection and Usage. In Proceedings of the ACM Conference on Health, Inference and Learning, Toronto, ON, Canada, 2–4 April 2020.
101. Agarwal, R.; Bjarnadottir, M.; Rhue, L.; Dugas, M.; Crowley, K.; Clark, J.; Gao, G. Addressing algorithmic bias and the perpetuation of health inequities: An AI bias aware framework. *Health Policy Technol.* **2023**, *12*, 100702. [CrossRef]
102. Ennab, M.; Mcheick, H. Designing an Interpretability-Based Model to Explain the Artificial Intelligence Algorithms in Healthcare. *Diagnostics* **2022**, *12*, 1557. [CrossRef]
103. Lee, Y.F. An Interoperability Solution for Legacy Healthcare Devices. *IT Prof.* **2015**, *17*, 51–57. [CrossRef]