The Use of Machine Learning in Real-World Data: A Systematic Review of Disease Prediction and Management

Norah Hamad Alhumaidi, Doni Dermawan, Hanin Farhana Kamaruzaman, Nasser Alotaiq

Submitted to: Journal of Participatory Medicine
on: November 17, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on it's website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressively prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Preprint
JMIR Publications

The Use of Machine Learning in Real-World Data: A Systematic Review of Disease Prediction and Management

Norah Hamad Alhumaidi[1*] MD; Doni Dermawan[2*] MSc; Hanin Farhana Kamaruzaman[3, 4] MSc; Nasser Alotaiq[5] PhD

[1]Qassim University College of Medicine Buraidah SA
[2]Warsaw University of Technology Applied Biotechnology, Faculty of Chemistry Warsaw PL
[3]Malaysian Health Technology Assessment Section (MaHTAS) Medical Development Division, Ministry of Health Malaysia Wilayah Persekutuan Putrajaya MY
[4]University of Glasgow Health Economics and Health Technology Assessment (HEHTA) School of Health and Wellbeing Glasgow GB
[5]Imam Mohammad ibn Saud Islamic University Health Sciences Research Center (HSRC) Riyadh SA
[*]these authors contributed equally

Corresponding Author:
Nasser Alotaiq PhD
Imam Mohammad ibn Saud Islamic University
Health Sciences Research Center (HSRC)
Riyadh
Riyadh
SA

Abstract

Background: Machine learning (ML) and big data analytics are revolutionizing healthcare, particularly in disease prediction, management, and personalized care. With vast amounts of real-world data (RWD) from sources like electronic health records (EHRs), patient registries, and wearable devices, ML offers significant potential to improve clinical outcomes. However, data quality, transparency, and clinical integration challenges remain.

Objective: This study aims to systematically review the use of ML in real-world data for disease prediction and management, identifying the most common ML methods, disease types, study designs, and sources of real-world evidence (RWE).

Methods: A systematic review followed the PRISMA guidelines to identify studies that utilized machine learning methods for analyzing real-world data in disease prediction and management. The review focused on extracting data related to the machine learning algorithms used, disease categories, types of studies, and sources of RWE, such as electronic health records (EHRs), patient registries, and wearable devices.

Results: The systematic review revealed that the most frequently employed machine learning methods were Random Forest (RF), Logistic Regression (LR), and Support Vector Machine (SVM). These methods were applied across various disease categories, with cardiovascular diseases, cancers, and neurological disorders being the most common. Real-world evidence primarily originated from EHRs, patient registries, and wearable devices, with a predominant focus on predictive modeling to improve clinical outcomes.

Conclusions: ML and big data hold significant promise for enhancing healthcare through better disease prediction and management. However, data quality, model interpretability, and generalizability must be addressed to integrate ML models fully into clinical practice.

(JMIR Preprints 17/11/2024:68898)
DOI: https://doi.org/10.2196/preprints.68898

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✔ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

- ✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
- Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
- Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Preprint
JMIR Publications

Original Manuscript

Review

The Use of Machine Learning in Real-World Data: A Systematic Review of Disease Prediction and Management

Norah Hamad Alhumaidi, MD[1]; Doni Dermawan[2], MSc. Eng; Hanin Farhana Kamaruzaman[3,4], M.Sc; Nasser Alotaiq[5], PhD

[1] College of Medicine, Qassim University, Qassim, Saudi Arabia
[2] Applied Biotechnology, Faculty of Chemistry, Warsaw University of Technology, 00-661 Warsaw, Poland
[3] Malaysian Health Technology Assessment Section (MaHTAS), Medical Development Division, Ministry of Health Malaysia
[4] Health Economics and Health Technology Assessment (HEHTA), School of Health and Wellbeing, University of Glasgow
[5] Health Sciences Research Center (HSRC), Imam Mohammad Ibn Saud Islamic University (IMSIU), 13317 Riyadh, Saudi Arabia

Corresponding Author:
Nasser Alotaiq, PhD
Health Sciences Research Center (HSRC)
Imam Mohammad Ibn Saud Islamic University (IMSIU)
P.O. Box 7544, Riyadh 13317
Saudi Arabia
Phone: +966 11 258 3011
Email: NAALOTAIQ@IMAMU.EDU.SA

preprint JMIR Publications

Abstract

Background: Machine learning (ML) and big data analytics are revolutionizing healthcare, particularly in disease prediction, management, and personalized care. With vast amounts of real-world data (RWD) from sources like electronic health records (EHRs), patient registries, and wearable devices, ML offers significant potential to improve clinical outcomes. However, data quality, transparency, and clinical integration challenges remain.
Objective: This study aims to systematically review the use of ML in real-world data for disease prediction and management, identifying the most common ML methods, disease types, study designs, and sources of real-world evidence (RWE).
Methods: A systematic review followed the PRISMA guidelines to identify studies that utilized machine learning methods for analyzing real-world data in disease prediction and management. The review focused on extracting data related to the machine learning algorithms used, disease categories, types of studies, and sources of RWE, such as electronic health records (EHRs), patient registries, and wearable devices.
Results: The systematic review revealed that the most frequently employed machine learning methods were Random Forest (RF), Logistic Regression (LR), and Support Vector Machine (SVM). These methods were applied across various disease categories, with cardiovascular diseases, cancers, and neurological disorders being the most common. Real-world evidence primarily originated from EHRs, patient registries, and wearable devices, with a predominant focus on predictive modeling to improve clinical outcomes.

Conclusions: ML and big data hold significant promise for enhancing healthcare through better disease prediction and management. However, data quality, model interpretability, and generalizability must be addressed to integrate ML models fully into clinical practice.

KEYWORDS
big data; machine learning; real-world data; real-world evidence

Introduction

Background

Advances in big data analytics and the growing availability of real-world data (RWD) are transforming healthcare by enabling new applications of machine learning (ML) to improve health outcomes [1]. Real-world evidence (RWE) generated from diverse data sources like electronic health records (EHRs), patient registries, and wearable devices has become central to informed decision-making in clinical practice [2, 3]. When combined with ML, RWD presents a promising avenue to enhance disease prediction, personalize patient management, and optimize therapeutic effectiveness. By providing a comprehensive view of patient histories and real-world health outcomes, ML applications in healthcare can drive actionable insights across various domains, including disease diagnosis, treatment planning, and chronic disease management [4, 5]. Real-world data captures information about patients in naturalistic settings, revealing how healthcare is delivered and its outcomes. Unlike clinical trials that operate within controlled conditions, RWD offers a more representative view of patient experiences, treatment responses, and health outcomes [6]. The rise of big data technology and data management systems has facilitated the integration of vast, heterogeneous data types, allowing ML algorithms to identify complex patterns within high-dimensional datasets [7, 8]. These capabilities allow healthcare providers to predict health outcomes, identify at-risk populations, and tailor interventions based on individual patient factors, thus making strides toward precision medicine [9].

Challenges in Machine Learning with Big Data and Real-world Health Data

Despite its potential, ML applications in RWD and big data contexts face several challenges. Data quality remains a primary concern, as real-world data often features inconsistencies, missing values, and lack of standardization [10]. Unlike the structured data from controlled clinical trials, RWD demands extensive preprocessing, including advanced natural language processing (NLP) methods and imputation techniques to address data gaps. Such efforts are critical to enhancing ML model reliability and ensuring accurate, meaningful outcomes [11, 12]. Biases present another key issue. ML models trained on RWD may inherit biases from the data, often stemming from demographic imbalances or regional healthcare differences. Left unaddressed, these biases can lead to healthcare disparities, as ML-driven decisions might inaccurately represent minority populations or certain patient groups [13]. Incorporating fairness-aware ML algorithms and cross-validating models across multiple datasets can mitigate this challenge, though developing equitable ML models remains a high priority [14]. Another significant hurdle is the interpretability of ML models, especially deep neural networks, which are known for their "black box" nature. While complex models deliver high accuracy, their opaque decision-making process limits the ability to verify or explain predictions. Model transparency is crucial given the high stakes in healthcare, where ML-based recommendations can impact lives. Advances in interpretability tools like SHAP values and LIME have helped enhance model transparency, yet balancing interpretability with performance remains an area of active investigation [15, 16].

Ethical and Regulatory Considerations

The integration of ML with RWD poses ethical and regulatory challenges, especially regarding patient privacy, data security, and informed consent. Regulations like the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union impose strict standards for data protection. However, adapting these laws to the context of ML in healthcare is complex due to the scale and diversity of data involved [17, 18]. Solutions such as de-identification, secure data-sharing protocols, and clear data management strategies have become crucial to ensuring patient confidentiality while maximizing data utility [19]. Ensuring equitable treatment outcomes is another ethical imperative. Given that ML models trained on data from predominantly certain demographics may perform poorly on underrepresented groups, addressing these disparities is critical. By incorporating fairness-aware ML models and building representative datasets, healthcare practitioners can ensure that ML applications benefit all patient groups, regardless of demographics [20]. Regulatory bodies are beginning to develop specific guidelines for the use of ML and RWD in healthcare. The Food and Drug Administration (FDA), for example, has issued draft guidance on using RWE for regulatory decisions, and the European Medicines Agency (EMA) has also recognized the importance of RWE in evaluating drug safety and efficacy [21]. As ML applications in healthcare continue to grow, a solid regulatory framework will be necessary to safeguard patient health while supporting technological innovation.

Objectives

The systematic review aimed to explore and critically analyze the applications, challenges, and future directions of machine learning in processing real-world health data and big data across various disease domains. Specifically, it sought to identify prominent disease areas where machine learning with real-world data had demonstrated clinical utility, examine the types of machine learning algorithms and methodologies applied to big data in healthcare, and analyze the challenges associated with real-world data and big data, including issues of data quality, bias, and interpretability. Additionally, the review discussed ethical and regulatory frameworks relevant to the use of machine learning in healthcare, with an emphasis on patient privacy and equitable treatment, and outlined future research needs and opportunities to innovate using machine learning, real-world data, and big data for precision medicine and public health. This review drew on key databases, including PubMed, Scopus, Web of Science, and the Cochrane Library, along with regulatory websites from organizations such as the FDA and EMA.

Methods

Eligibility Criteria

For this systematic review, we focused exclusively on clinical trials and cohort studies that utilized ML techniques to analyze RWD for disease prediction and management. Studies were included if they met the following criteria: (1) they were randomized controlled trials (RCTs), pragmatic clinical trials, observational clinical trials, or cohort studies; (2) they involved the application of ML methods (e.g., supervised learning, unsupervised learning, deep learning) to RWD for clinical decision-making, disease prediction, or management of common diseases such as cardiovascular diseases, diabetes, cancer, and chronic conditions; and (3) they used real-world health data sources such as EHRs, patient registries, or wearable health devices. Exclusion criteria included trials that did not apply ML techniques or used only data from controlled clinical trials rather than real-world settings.

Information Sources

The following information sources were utilized to capture the most relevant clinical trial and cohort studies: PubMed, Scopus, and the Cochrane Library. PubMed was specifically targeted for clinical trials and biomedical research, particularly studies published in leading clinical journals. Scopus and the Cochrane Library were also searched to gather clinical trial reports within the healthcare and ML domains. To ensure comprehensive coverage, Google Scholar was included to identify grey literature, such as theses and reports not indexed in traditional databases. These sources were selected to provide a broad overview of clinical trial data and their relevance to ML applications in disease management. Additionally, regulatory bodies such as the U.S. FDA and the EMA were consulted to gain insights into clinical trial guidelines and regulatory standards regarding the use of RWD in healthcare.

Search Strategy

A targeted search strategy was developed to identify clinical trials and cohort studies focused on ML applications in RWD. The search query incorporated key terms related to ML (e.g., "machine learning," "deep learning," "artificial intelligence") and clinical trials (e.g., "clinical trial," "randomized controlled trial," "pragmatic clinical trial") along with terms related to disease management (e.g., "disease prediction," "healthcare outcomes"). For example, the search used: ("machine learning" OR "deep learning") AND ("clinical trial" OR "randomized controlled trial" OR "pragmatic trial" OR “cohort study”) AND ("real-world data" OR "electronic health records" OR "patient registries"). Boolean operators (AND, OR), truncation, and Medical Subject Headings (MeSH) terms were employed to refine the search and ensure comprehensive coverage. The search was limited to studies published in the last 10 years (2014-2024) to reflect recent developments. Only English-language publications were included. Additionally, relevant studies were identified through manual searches of reference lists from key articles and by reviewing clinical trial registries such as ClinicalTrials.gov to ensure comprehensive coverage of the clinical trials relevant to ML in disease management.

Study Selection

The study selection process was conducted in two stages: an initial screening of titles and abstracts, followed by a full-text review. In the first stage, the titles and abstracts of all identified articles were evaluated for relevance based on the predefined inclusion criteria. Studies that did not meet these criteria, such as those not involving clinical trials or cohort studies and not applying ML methods to RWD, were excluded. Studies that met the inclusion criteria proceeded to the second stage, where the full texts were retrieved for further evaluation. Each study was carefully assessed for adherence to the eligibility criteria, focusing on the clinical trial design, use of RWD, and application of ML algorithms for disease prediction or management. Any discrepancies in study selection were resolved through discussion. To ensure transparency and clarity in the selection process, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram was used to document the number of studies at each review stage, including identification, screening, eligibility, and final inclusion [22, 23]. This approach allowed for a systematic and comprehensive selection of studies, ensuring that only relevant clinical trials were included in the final analysis.

Data Extraction

Data extraction was performed independently by two reviewers using a standardized form. Key data points extracted from each clinical trial and cohort studies included study characteristics (e.g., author(s), year of publication, trial design), the specific machine learning methods employed (e.g., supervised learning, reinforcement learning, deep learning), disease areas targeted (e.g., cardiovascular diseases, diabetes, cancer), and the types of real-world data sources used (e.g., EHRs, patient registries, wearable devices). Additionally, we extracted the performance metrics of the ML models used, such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC), to evaluate their effectiveness in disease prediction and management. Information on the challenges and limitations of applying ML to RWD in clinical trials, such as data quality issues, biases, or model interpretability, was also collected. Any disagreements in data extraction were resolved through discussion. The extracted data was organized systematically to synthesize findings across studies.

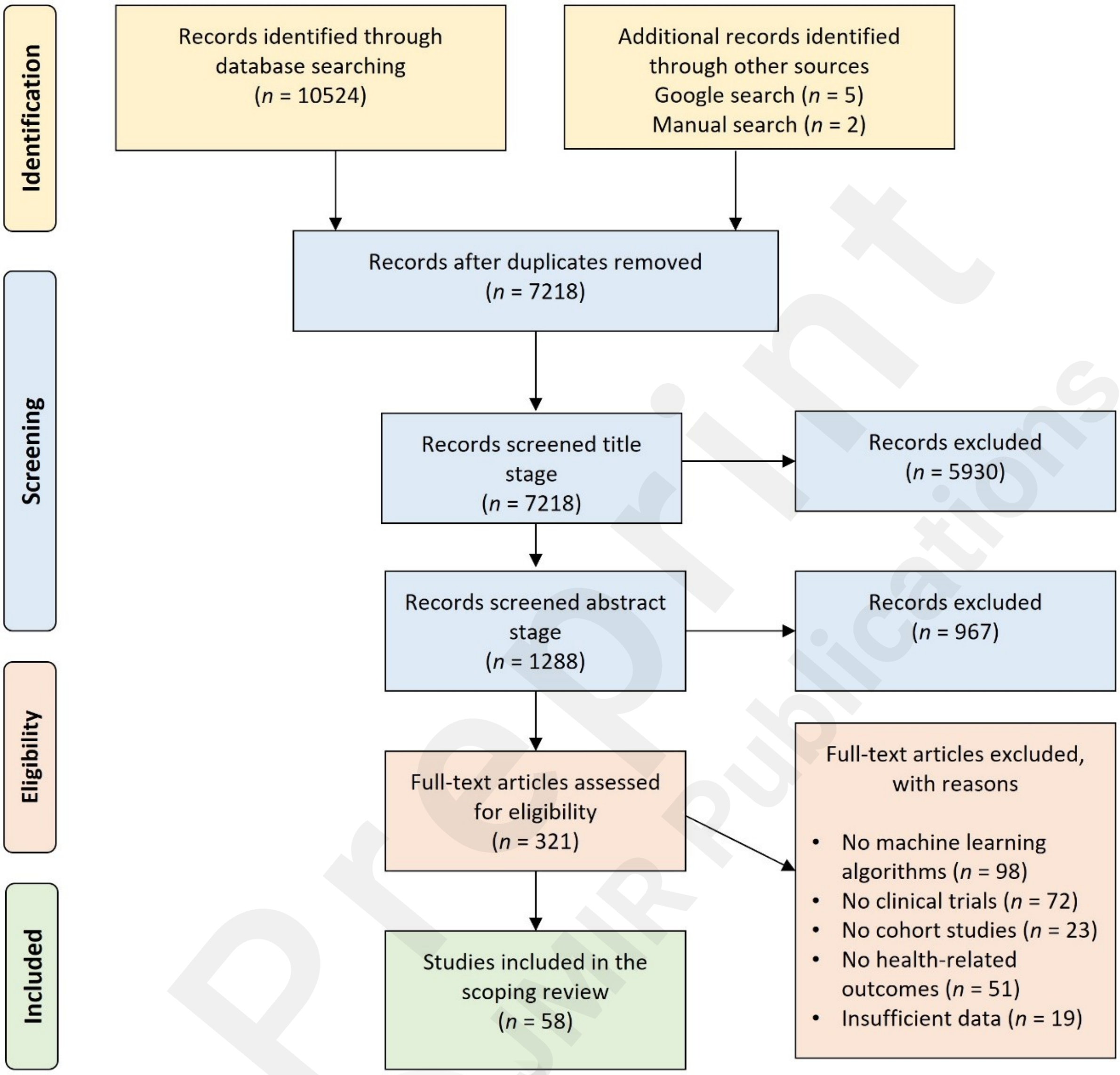
Results

Systematic Literature Search and Study Selection Workflow

The systematic literature search was conducted to identify studies applying ML techniques to RWD in clinical trials and cohort studies, focusing on disease prediction and management. The search covered multiple databases, including PubMed, Scopus, Web of Science, and the Cochrane Library, intending to capture a broad range of studies from biomedical, clinical, and healthcare research fields. This search yielded a total of 10,524 records, as shown in the PRISMA flow diagram (**Figure 1**). To ensure comprehensive coverage, an additional eight records were identified through external sources, such as Google searches and manual hand-searching of journals not indexed by these main databases, including grey literature and nontraditional academic sources. Following the removal of duplicates, 7,218 unique studies remained for further screening. The screening phase commenced with title-based screening of the 7,218 records. During this phase, titles were reviewed to assess their relevance based on predefined inclusion criteria. Studies with titles that did not suggest the application of ML to RWD in clinical or disease management contexts were excluded. This stage resulted in the exclusion of 5,930 records deemed irrelevant based on their titles, leaving 1,288 studies to move on to abstract screening.

Abstract screening followed, with a detailed examination of each abstract for inclusion criteria, such as the use of ML techniques, RWD sources (e.g., electronic health records, patient registries, wearable device data), and relevance to disease prediction or management. This phase resulted in the exclusion of 967 studies for failing to meet these criteria. Common reasons for exclusion included a lack of ML application (e.g., studies using traditional statistical models only) or studies not addressing disease management. After this filtering process, 321 studies remained for full-text review. The eligibility assessment involved thoroughly reviewing each full text to confirm that all inclusion criteria were met. A total of 263 articles were excluded at this stage for specific reasons: 98 studies lacked ML algorithms (instead using conventional statistical methods), 72 were unrelated to clinical trial methodologies, 23 did not involve study cohorts, 51 were unrelated to healthcare outcomes, and 19 provided insufficient data on ML model performance or data sources. Each exclusion ensured that only studies fully relevant to ML-based disease prediction or management using RWD were retained. Following this rigorous multi-stage selection process, 58 studies met all eligibility criteria and were included in the final scoping review. These selected studies represent a diverse range of clinical applications, disease areas, ML methodologies, and RWD sources, offering insights into the current role of ML in clinical trials and cohort studies.

Figure 1. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram depicts the study selection process from initial identification to final inclusion, detailing the number of records screened, excluded, and ultimately included in the scoping review.



The results of this review demonstrate that, despite the initial large number of records identified through comprehensive database searches, only a small fraction of the studies ultimately met the stringent inclusion criteria necessary for assessing the application of ML techniques in RWD for disease prediction and management. Out of the 10,524 records initially identified, a significant portion was excluded during each phase of the screening process, primarily due to the lack of alignment with the predefined criteria. In the title-based screening phase, the vast majority of studies (5,930 out of 7,218) were excluded for not being relevant to the focus of this review, particularly those that did not explore the use of ML or RWD in clinical decision-making. This initial filtering reduced the pool of records considerably. A substantial number of studies were further excluded during abstract screening (967 records), as many failed to involve the application of ML algorithms, were not related to clinical trials or cohort studies, or did not address disease prediction and management. At the full-text assessment stage, the exclusion of 263 studies highlighted the challenge of identifying studies that precisely met the criteria for this review. The primary reasons for exclusion included a lack of machine learning techniques, as many studies used traditional statistical methods, such as regression models or expert systems, which are not classified as machine learning. Additionally, many studies did not pertain to clinical trials or cohort studies. In contrast, others lacked specific data on health outcomes or failed to provide sufficient details on the ML models used. This rigorous multi-step selection process underscores researchers' challenges when searching for high-quality studies that align with niche research areas, such as ML applications in real-world clinical settings. **Table 1** provides an overview of the studies included in the scoping review, with the key findings from each study presented in **Multimedia Appendix 1**.

Table 1. A summary of the studies included in the scoping review, outlining the study characteristics, diseases/medical conditions, type of study, source of real-world evidence (RWE), and machine learning (ML) methods used.

Study; Year	Database	Diseases/ Medical Conditions (Category)	Type of Study	Type of Real-world Evidence (RWE)	Machine Learning (ML) Methods
Wissel, BD. et al [24], 2023	PubMed	Epilepsy (Neurological diseases)	Evaluation of healthcare outcomes	Electronic health record (EHR)	Natural language processing (NLP)
Ayers, B. et al [25], 2021	PubMed	Orthotopic heart transplantation (Cardiovascular diseases)	Survival prediction	Patient registries	Deep neural network (DNN), random forest (RF), and adaptive boosting (AdaBoost)
Nadarajah, R. et al [26], 2023	PubMed	Atrial fibrillation (Cardiovascular diseases)	Disease prediction	EHR	FIND-AF machine learning algorithm
Yadgir, SR. et al [27], 2021	PubMed	Cognitive impairment (Neurological diseases)	Disease prediction	EHR	Extreme gradient boosting (XGBoost)
Liu, Y. et al [28], 2023	PubMed	Peripheral artery disease (Cardiovascular diseases)	Survival prediction	EHR	Logistic regression (LR), gradient boosting machine (GBM), RF, decision tree (DT), XGBoost, neural network, Cox regression, and random survival forest (RSF)
Hill, NR. et al [29], 2022	PubMed	Atrial fibrillation (Cardiovascular diseases)	Disease prediction and cost-effectiveness	EHR	Prediction of Undiagnosed atriaL fibrillation using a machinE learning AlgorIthm (PULsE-AI)
Sheth, SA. et al [30], 2019	PubMed	Acute ischemic stroke (Cardiovascular diseases)	Disease prediction	EHR	Convolutional neural network (CNN)
Barton, C. et al [31], 2019	PubMed	Sepsis (Infectious diseases)	Disease prediction	EHR	XGBoost
Kao, YT. et al [32], 2023	PubMed	Atrial fibrillation (Cardiovascular diseases)	Disease prediction	EHR	DT, support vector machine (SVM), LR, and RF
Kim, M. et al [33], 2022	PubMed	Atrial high-rate episodes (AHREs) (Cardiovascular diseases)	Disease prediction	Wearable devices	RF, SVM, and XGBoost
Park, HB. et al [34], 2023	PubMed	Coronary artery disease (Cardiovascular diseases)	Disease prediction	Patient registries	Bayesian quantile regression (BQR)
Hilbert, A. et al [35], 2019	PubMed	Acute ischemic stroke (Cardiovascular diseases)	Healthcare outcomes and decision-	Wearable devices	Residual Neural Network (ResNet)

Study; Year	Database	Diseases/ Medical Conditions (Category)	Type of Study	Type of Real-world Evidence (RWE)	Machine Learning (ML) Methods
			making		
Chen, W. et al [36], 2021	PubMed	Ewing sarcoma (Tumors)	Survival prediction	Patient registries	Boosted DT, SVM, nonparametric RF, and neural network
Koutsouleris, N. et al [37], 2017	PubMed	Schizophrenia (Neurological diseases)	Healthcare outcomes and decision-making	Patient registries	Non-linear SVM
Strömblad, CT. et al [38], 2021	PubMed	Colorectal and gynecology cancer (Cancers)	Healthcare outcomes	EHR	GBM and LR
Wang, SV. et al [39], 2019	PubMed	Atrial fibrillation (Cardiovascular diseases)	Decision-making	EHR	DT, RF, and LR
Tan, TH. et al [40], 2021	PubMed	Influenza (Infectious diseases, respiratory diseases)	Healthcare outcomes	EHR	RF, XGBoost, and LR
Goerigk, S. et al [41], 2020	PubMed	Depression (Neurological diseases)	Decision-making	Patient registries	LR, SVM, RF, tree-based stochastic gradient boosting, and XGBoost
Kijpaisalratana, N. et al [42], 2024	PubMed	Sepsis (Infectious disease)	Decision-making	EHR	RF, XGBoost, LR, and SVM
Sharma, A. et al [43], 2019	PubMed	Acute coronary syndrome (Cardiovascular diseases)	Survival prediction	Patient registries	Cox regression
Singhal, L. et al [44], 2021	PubMed	Acute Respiratory Distress Syndrome (ARDS) (Respiratory diseases)	Disease prediction	EHR	Neural Networks, SVM, RF, LR, and XGBoost
Kanchanatawan, B. et al [45], 2018	PubMed	Schizophrenia (Neurological diseases)	Disease prediction	Patient registries	SVM and RF
Huang, J. et al [46], 2022	PubMed	Ischemic stroke (Cardiovascular diseases)	Survival prediction	EHR	Naive Bayes (NB), XGBoost, and LR
She, H. et al [47], 2023	PubMed	Sepsis (Infectious disease)	Disease prediction	Patient registries	SVM and RF
Sundar, R. et al [48], 2022	PubMed	Gastric cancer (Cancers)	Survival prediction	Patient registries	RF
Alaa, AM. et al [49], 2019	PubMed	Cardiovascular disease risk (Cardiovascular diseases)	Disease prediction	Patient registries	Linear SVM, RF, neural networks, AdaBoost, and XGBoost
Azimi, P. et al [50], 2017	PubMed	Lumbar spinal canal stenosis (LSCS) (Spinal diseases)	Decision-making	EHR	Artificial neural network (ANN) and LR
Baxter, SL. et al [51], 2019	PubMed	Glaucoma (Ocular diseases)	Decision-making	EHR	Multivariable logistic regression (MLR), RF, and ANN
Anderson, JP. et al [52], 2015	PubMed	Type 2 Diabetes (Metabolic diseases)	Disease prediction	EHR	RF and SVM
Bannister, CA. et al [53], 2018	PubMed	Stroke and myocardial infarction (Cardiovascular diseases)	Survival prediction	Patient registries	Cox regression
Scheer, JK. et al [54], 2017	Web of Science	Spinal deformity surgery (Spinal diseases)	Decision-making	Patient registries	DT and ANN
Rau, HH. et al	Web of Science	Liver cancer (Cancers)	Disease prediction	EHR	ANN and LR

Preprint JMIR Publications

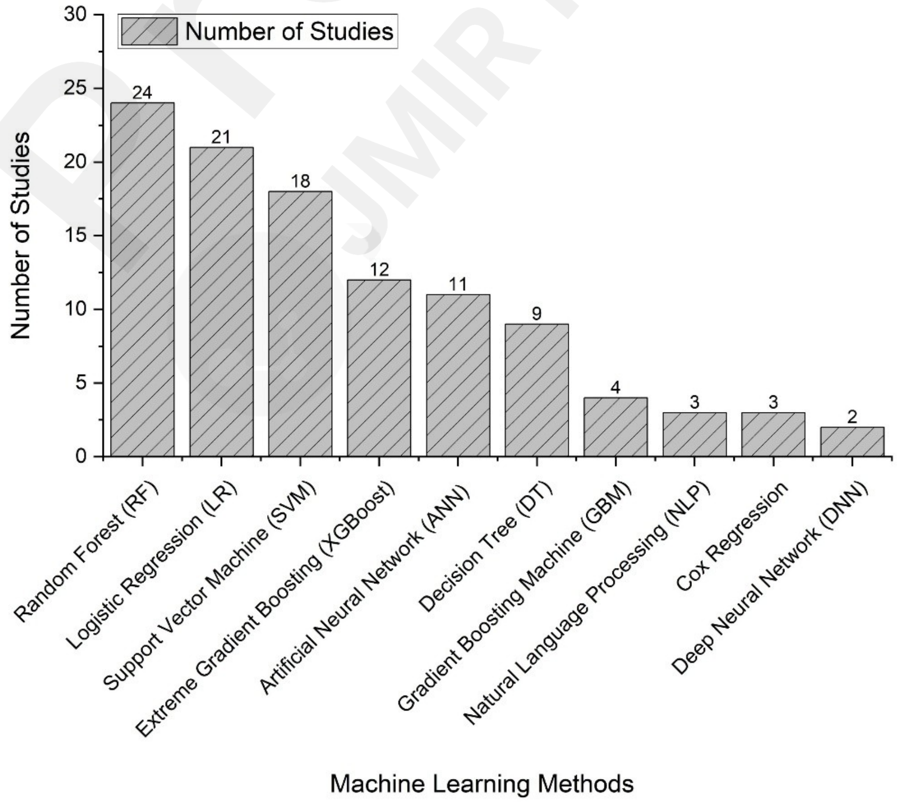
Study; Year	Database	Diseases/ Medical Conditions (Category)	Type of Study	Type of Real-world Evidence (RWE)	Machine Learning (ML) Methods
[55], 2016					
Ramezankhani, A. et al [56], 2016	Web of Science	Type 2 Diabetes (Metabolic diseases)	Disease prediction	Patient registries	DT
Pei, D. et al [57], 2019	Web of Science	Type 2 Diabetes (Metabolic diseases)	Disease prediction	EHR	DT
Oviedo, S. et al [58], 2019	Web of Science	Postprandial hypoglycemia (Metabolic diseases)	Decision-making	Wearable devices	SVM
Mubeen, AM. et al [59], 2017	Web of Science	Alzheimer's disease (Neurological diseases)	Disease prediction	EHR	RF
Lopez-de-Andres, A. et al [60], 2016	Web of Science	Type 2 Diabetes (Metabolic diseases)	Survival prediction	Patient registries	ANN
Kwon, JM. et al [61], 2019	Web of Science	Cardiac arrest (Cardiovascular diseases)	Survival prediction	Patient registries	DNN, LR, SVM, and RF
Kim, I. et al [62], 2019	Web of Science	Breast cancer (Cancers)	Decision-making	EHR	Two-class Decision Jungle and Two-class Neural Network
Khanji, C. et al [63], 2019	Web of Science	Hypertension and dyslipidemia (Cardiovascular diseases)	Healthcare outcomes	EHR	LR
Karhade, AV. et al [64], 2019	Web of Science	Lumbar disc herniation (Spinal diseases)	Decision-making	EHR	LR, RF, XGBoost, ANN, and SVM
Jovanovic, P. et al [65], 2014	Web of Science	Choledocholithiasis (Gastrointestinal diseases)	Decision-making	EHR	ANN
Kang, AR. et al [66], 2020	Web of Science	Postinduction hypotension (Anaesthesia-related complications)	Healthcare outcomes	EHR	NB, LR, RF, and ANN
Isma'eel, HA. et al [67], 2018	Web of Science	Coronary artery disease (Cardiovascular diseases)	Healthcare outcomes	EHR	ANN
Hill, NR. et al [68], 2019	Web of Science	Atrial fibrillation (Cardiovascular diseases)	Disease prediction	EHR	RF, SVM, and Cox regression
Dong, Y. et al [69], 2019	Web of Science	Chinese Crohn's disease (Gastrointestinal diseases)	Decision-making	EHR	RF, LR, SVM, DT, and ANN
Bowman, A. et al [70], 2018	Web of Science	Carpal tunnel syndrome (Musculoskeletal diseases)	Healthcare outcomes	EHR	LR and ANN
Bertsimas, D. et al [71], 2018	Web of Science	Breast, lung, ovarian cancers (Cancers)	Survival prediction	EHR	DT
Manz, CR. et al [72], 2020	Cochrane Library	Cancer-related serious illness (Cancers)	Decision-making	EHR	RF and SVM
Tian, D. et al [73], 2023	Cochrane Library	Lung transplantation (Respiratory diseases)	Survival prediction	EHR	Random survival forests (RSF)
Li, E. et al [74], 2022	Cochrane Library	Latent profile analysis (Cancers)	Decision-making	EHR	GBM
Tedeschi, SK. et al [75], 2021	Cochrane Library	Pseudogout (Rheumatic diseases)	Disease prediction	EHR	NLP
Wissel, BD. et al [24], 2023	Cochrane Library	Epilepsy surgery (Neurological diseases)	Decision-making	EHR	NLP
Ambwani, G. et al [76], 2019	Cochrane Library	Cancer biomarkers (Cancers)	Healthcare outcomes	EHR	LR
Jorge, A. et al [77], 2019	Cochrane Library	Lupus (Autoimmune disease)	Disease prediction	EHR	LR
Shimabukuro, DW. et al [78],	Cochrane Library	Sepsis (Infectious diseases)	Healthcare outcomes	EHR	LR

Study; Year	Database	Diseases/ Medical Conditions (Category)	Type of Study	Type of Real-world Evidence (RWE)	Machine Learning (ML) Methods
2017					
Sarraju, A. et al [79], 2021	Cochrane Library	Atherosclerosis (Cardiovascular diseases)	Healthcare outcomes	EHR	RF, GBM, XGBoost, and LR
Ye, B. et al [80], 2019	Cochrane Library	Myopia (Ocular diseases)	Healthcare outcomes	Wearable devices	SVM

Implementation of Machine Learning (ML) in Real-World Data (RWD) for Disease Prediction and Management

Machine learning (ML) methods have become integral tools in analyzing real-world data (RWD) for disease prediction and management. **Figure 2** showcases the top 10 most frequently employed ML methods across the studies reviewed, each contributing unique strengths to clinical decision-making and health outcome predictions. These methods analyze complex medical data, helping clinicians make informed decisions for better patient care. Random Forest (RF) is one of the most widely used ML methods, appearing in 24 studies. It is an ensemble learning technique that builds multiple decision trees and combines their outputs to enhance prediction accuracy [81]. In real-world data, RF is particularly effective for disease prediction because it can manage large datasets with numerous variables, including electronic health records (EHRs), which are common medical data sources. Its robustness against overfitting and ability to handle missing data make it ideal for clinical applications, where data quality can vary [82, 83]. RF excels in handling high-dimensional datasets, making it highly valuable for predictive modeling in healthcare, where diverse and complex datasets are prevalent [84]. It is often applied to predict disease outcomes, assess treatment responses, and identify patient risk factors.

Figure 2. Top 10 most employed machine learning (ML) methods in real-world data related to disease prediction and management.



Logistic Regression (LR) is a fundamental method used for binary classification tasks and was employed in 21 studies. LR estimates the probability of a particular class, which is essential for predicting binary outcomes such as disease presence or absence. It is a simple yet powerful tool that works well with smaller datasets and provides results that are easy to interpret, making it particularly useful in clinical settings where transparency is crucial [85, 86]. LR is commonly used for disease risk prediction, helping clinicians assess the likelihood of a patient developing a condition based on their medical history and other clinical factors. Its interpretability allows for clear communication of results to healthcare providers, enhancing decision-making [87, 88]. Support Vector Machine (SVM), employed in 18 studies, is known for its ability to handle high-dimensional data, making it suitable for complex medical datasets, including genomic and imaging data. SVM works by finding the optimal hyperplane that separates different classes in the feature space [89-91]. This method is beneficial in clinical settings where the relationship between variables is non-linear and can be adapted for classification and regression tasks. SVM is applied in disease prediction, particularly when the dataset has many features relative to the number of observations. It is also useful for classifying patients based on genetic or demographic factors, making it a powerful tool for precision medicine [92, 93]. Extreme Gradient Boosting (XGBoost) appears in 12 studies and is a highly effective method for improving predictive accuracy through boosting. XGBoost builds models sequentially to correct errors made by previous models and uses regularization to prevent overfitting. This method is precious in handling large datasets, common in clinical studies, and where computational efficiency is essential [94, 95]. XGBoost is often used for survival analysis and disease outcome prediction, where it can effectively manage the complexity of large datasets and missing data. Its flexibility allows it to be applied across various disease areas, from cancer prognosis to cardiovascular risk assessment [96, 97].

Artificial Neural Networks (ANN), utilized in 11 studies, are powerful tools for modeling complex, nonlinear relationships in data. With multiple layers of interconnected neurons, ANN can learn intricate patterns from large datasets. It is widely used in applications that involve unstructured data, such as medical imaging and genetic data, where traditional models might struggle [98, 99]. ANN is frequently applied to predict disease progression, response to treatments, and identify potential biomarkers. In real-world data, ANN helps identify subtle patterns in complex datasets that simpler models might not capture, such as predicting cancer progression from radiological images [100, 101]. Decision Trees (DT), featured in 9 studies, are straightforward and interpretable models that split data into subsets based on feature values. DT models are highly useful in real-world healthcare settings, where interpretability is essential for clinical decision-making. They are often applied in clinical decision support systems (CDSS) to guide treatment decisions based on patient data [102, 103]. In healthcare, DTs predict disease outcomes, stratify patients by risk, and recommend treatment plans. Their transparency allows clinicians to understand the decision-making process, which is critical for patient trust and informed consent [104]. Gradient Boosting Machine (GBM) is used in 4 studies and is a powerful ensemble method that focuses on correcting errors made by previous models. It is effective in producing highly accurate predictions, particularly in the presence of noisy or incomplete data. GBM is more computationally intensive than other methods but often outperforms simpler models in accuracy. GBM is particularly useful for predicting disease progression and evaluating treatment efficacy in longitudinal studies, where multiple factors influence outcomes over time [105, 106].

Natural Language Processing (NLP), employed in 3 studies, is a subfield of AI focused on analyzing unstructured textual data. In healthcare, NLP extracts relevant information from clinical notes, electronic health records, and medical literature. It enables clinicians and researchers to analyze vast amounts of text data to identify trends, predict disease outcomes, and assess treatment effectiveness [107, 108]. NLP is crucial in extracting insights from EHRs and other textual data sources. It can help in disease prediction by identifying patterns from patient narratives, diagnostic codes, and clinician notes that would otherwise remain hidden in unstructured formats [109]. Cox Regression, used in 3 studies, is designed explicitly for survival analysis. It is widely applied in clinical research to model the time of an event, such as the onset of a disease or patient survival. This method is precious for understanding how various predictors affect the risk of an event occurring over time. In real-world data, Cox Regression is often used in cancer studies and other chronic diseases to predict survival times and assess the impact of different treatment regimens, making it indispensable in clinical trials and outcome-based research [110, 111]. Deep Neural Networks (DNN), employed in 2 studies, are a more complex version of ANN with multiple hidden layers. DNNs identify intricate patterns and are increasingly used in healthcare applications involving large and complex data types such as medical imaging, genomics, and sensor data. DNN is particularly useful for analyzing high-dimensional data such as medical images (e.g., X-rays, MRIs) or genomic data, where the relationships between variables are complex and nonlinear. It helps identify disease markers and predict outcomes based on these complex datasets [112-114]. The diverse range of machine learning methods employed in real-world data for disease prediction and management demonstrates the adaptability of these techniques in clinical practice. From interpretable models like Logistic Regression and Decision Trees to more complex methods such as Deep Neural Networks and Extreme Gradient Boosting, each ML technique uniquely enhances predictive capabilities. These methods enable healthcare providers to make more accurate, data-driven decisions, ultimately improving patient outcomes and advancing personalized medicine.

Distribution of Diseases, Study Types, and Real-World Evidence (RWE) Sources in Machine Learning Applications

The distribution of disease types in studies utilizing ML for disease prediction and management reveals a strong emphasis on cardiovascular diseases, with 19 studies focusing on various conditions within this category (**Figure 3a**). This high representation can be attributed to the multifactorial and complex nature of cardiovascular diseases, which often involve a combination of genetic, environmental, and lifestyle factors. Conditions such as atrial fibrillation, heart transplantation, and peripheral artery disease are prominent in these studies, where advanced ML models are utilized to enhance predictive accuracy and improve patient management. For instance, studies on heart transplantation and atrial fibrillation highlight the potential of ML algorithms in survival prediction and early disease detection. The work by Ayers et al. demonstrated that ensemble models, combining random forest, deep neural networks, and adaptive boosting, significantly outperformed traditional logistic regression for predicting 1-year survival rates after orthotopic heart transplantation, with an AUROC of 0.764 [25]. Meanwhile, Nadarajah et al. explored using the FIND-AF ML algorithm to identify undiagnosed atrial fibrillation using data from EHR, aiming to improve early detection and intervention. In addition, studies on peripheral artery disease (PAD) and atrial fibrillation in older adults underscore the utility of ML models in survival prediction and risk assessment [26]. Liu et al. developed a predictive model for amputation-free survival post-revascularization, with the random survival forest model achieving the highest accuracy in predicting long-term outcomes [28]. Similarly, Kao et al. used various ML methods, including decision trees and random forests, to predict new-onset atrial fibrillation in older adults, achieving high specificity and performance, particularly with the random forest model [32]. Furthermore, the use of ML in acute ischemic stroke, including studies by Sheth et al. and Hilbert et al., illustrates the growing role of deep learning techniques, such as convolutional neural networks (CNN) and residual neural networks (ResNet), in improving diagnostic accuracy and predicting patient outcomes [35]. These advancements in ML can potentially revolutionize clinical decision-making and treatment selection, especially for conditions like stroke, where rapid and accurate assessment is critical.

A significant portion of studies also targeted cancers and tumors (9 studies), which are often characterized by their heterogeneity and the need for personalized treatment plans. Machine learning algorithms, such as RF and SVM, have enhanced early cancer detection, predicted disease recurrence, and assessed the effectiveness of different treatment protocols, offering great potential in oncology settings. One key area of focus is the prediction of disease outcomes. For instance, Chen et al. (2023) developed a series of ML models to predict the 5-year survival rate for patients with Ewing sarcoma, a rare type of cancer. By utilizing data from 2,332 patients, including various algorithms such as boosted decision trees, SVM, random forests, and neural networks, the study found that the random forest method performed best, with impressive sensitivity and specificity. This model is now available through a web-based application, providing a valuable tool for clinicians to assess survival probabilities for Ewing sarcoma patients [36]. Another study by Strömblad et al. (2023) employed a predictive ML model to improve surgical scheduling in cancer surgeries, specifically for colorectal and gynecologic cancers. This research utilized gradient boosting and linear regression techniques to predict surgical durations, reducing operational inefficiencies such as patient wait times and optimizing the use of surgical resources, demonstrating how ML can streamline healthcare operations while maintaining treatment quality [38]. Furthermore, in survival prediction, Sundar et al. (2023) utilized a random forest model to develop a gene signature that predicts the response of gastric cancer patients to paclitaxel treatment. Their model, which identified a 19-gene signature, enabled the classification of patients into those who would benefit from the treatment, providing a novel approach to personalized cancer therapy [48].

The studies focusing on neurological diseases, including conditions like epilepsy, cognitive impairment, and schizophrenia, highlight the significant impact of ML in improving diagnosis, treatment prediction, and healthcare outcomes. These studies underscore the potential of ML to personalize patient care and optimize clinical decision-making. For instance, the study by Wissel et al. (2024) investigated the application of NLP embedded in EHR to automate alerts for pediatric epilepsy patients. This ML-driven clinical decision support system successfully increased referrals for epilepsy surgery, with a marked improvement in presurgical evaluation rates and even higher rates of actual surgery, illustrating how NLP-based interventions can influence healthcare outcomes by improving referral efficiency and treatment access [24]. Similarly, Yadgir et al. (2024) focused on the use of XGBoost, a machine learning algorithm, to identify older emergency department patients at high risk for cognitive impairment. This predictive model, using EHR data, demonstrated high sensitivity and specificity, with the potential to reduce the need for in-person screenings and prioritize high-risk patients. By streamlining screening processes, this approach could enhance the detection of cognitive impairments in older adults, potentially leading to earlier interventions and better management of conditions like dementia [27]. In schizophrenia, Koutsouleris et al. (2024) developed a non-linear SVM model to predict treatment outcomes for first-episode psychosis patients. The model was trained on pre-treatment patient-reported data and successfully predicted poor versus good treatment outcomes, thus supporting clinical decision-making in terms of which treatments might be more effective for certain patients, and identifying those at risk for non-adherence or poor prognosis [37]. These studies collectively demonstrate how machine learning methods like NLP, XGBoost, and Random Forest are revolutionizing the management of neurological diseases. By enabling early detection, better prediction of disease outcomes, and more informed decision-making, these tools offer substantial improvements in both clinical and healthcare settings. Infectious diseases, metabolic diseases, spinal diseases, gastrointestinal diseases, ocular diseases, and respiratory diseases each had a smaller but notable presence in the studies (5 or fewer studies). These applications generally focus on disease prediction, early diagnosis, and treatment optimization. Machine learning models such as XGBoost and DNN have been employed to predict disease onset, assess risks, and improve patient management in these areas.

Figure 3. Distribution of disease types, study types, and real-world evidence (RWE) sources in the included studies. This figure illustrates the disease breakdown, the study designs employed, and the sources of real-world evidence (such as electronic health records, patient registries, etc.) used to inform disease prediction and management in the selected studies.

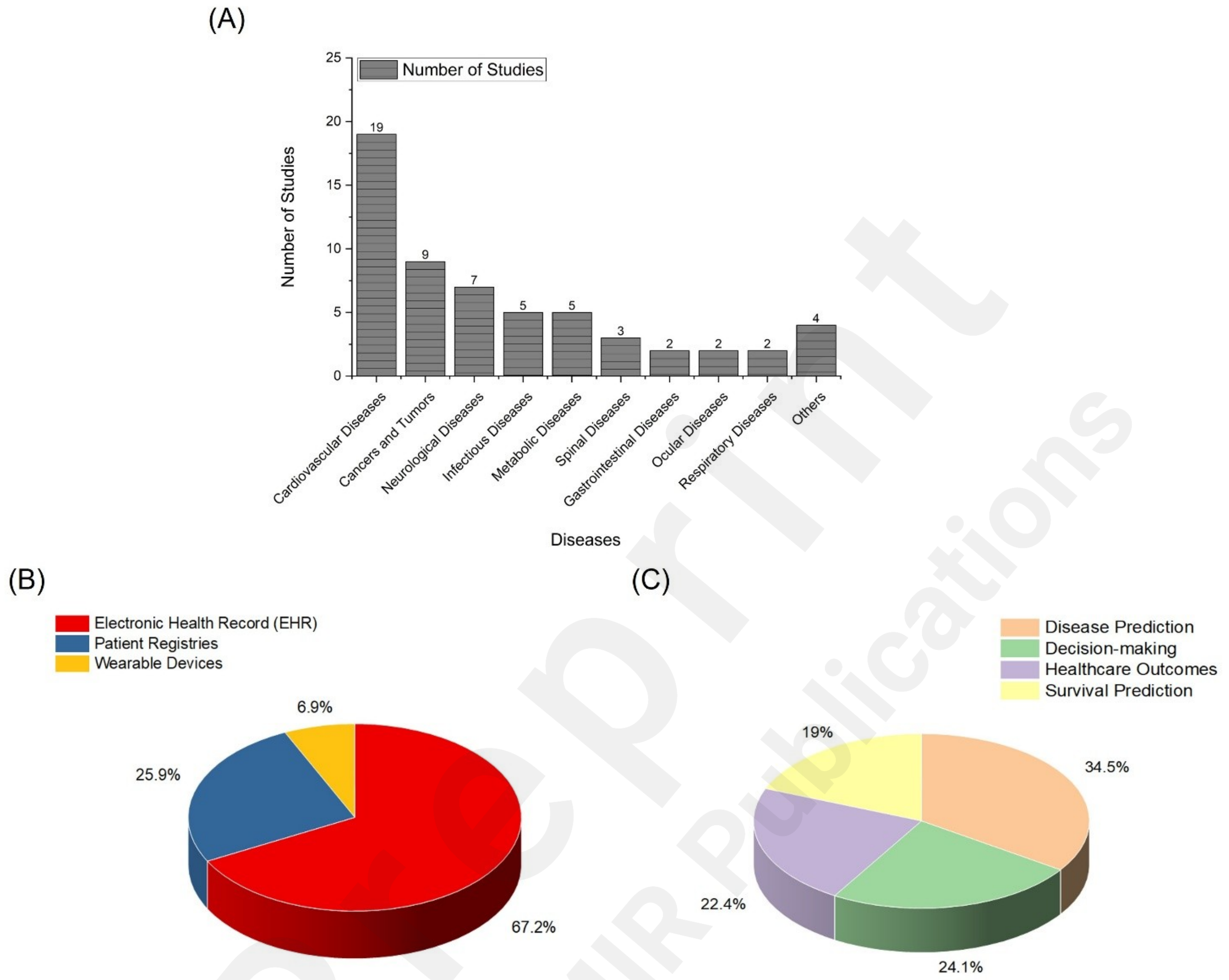


Figure 3b illustrates the distribution of the types of Real-World Evidence (RWE) utilized across the studies analyzed. The data reveals that Electronic Health Records (EHR) are the most frequently used form of RWE, accounting for 67.2% of the studies (39 out of 58). EHRs are a rich source of patient data, providing comprehensive records of patient health status, diagnoses, treatments, and outcomes over time. This makes EHRs particularly valuable for studies that require large-scale data to identify patterns, trends, and correlations in real-world clinical settings. The next most commonly used type of RWE is patient registries, which were utilized in 25.9% of the studies (15 out of 58). Patient registries typically collect data on specific patient populations with particular diseases or conditions, allowing for longitudinal tracking of disease progression and treatment outcomes. Wearable devices were the least utilized form of RWE, accounting for 6.9% of the studies (4 out of 58). Wearables are increasingly being used to collect real-time health data, including vital signs and activity levels, which can provide valuable insights into patients' health status outside of clinical environments. This distribution highlights the dominance of EHR as the primary data source in these studies, reflecting its accessibility and broad applicability in healthcare research.

Figure 3c presents the categorization of study objectives based on the type of research focus. Disease prediction emerged as the most widely studied area, represented by 34.5% of the studies (20

Preprint JMIR Publications

out of 58). This suggests a strong emphasis on using machine learning and data analytics to predict the onset, progression, or outcomes of various diseases. The next most studied area is decision-making, with 14 studies (24.1%), which underscores the growing interest in leveraging data-driven insights to inform clinical decisions and treatment strategies. Healthcare outcomes, such as quality of life, recovery rates, and adverse events, were the focus of 22.4% of the studies (13 out of 58), reflecting the importance of understanding how diseases and treatments affect patients' overall well-being. Survival prediction, accounting for 19.0% of the studies (11 out of 58), is another critical area of research, particularly in oncology and chronic diseases, where predicting patient survival and the effectiveness of interventions can guide clinical decision-making. This distribution indicates that disease prediction and decision-making are central to applying real-world evidence in healthcare, with a significant focus on improving patient outcomes and guiding treatment strategies.

Discussion

Principal Results

The findings of this study underscore the growing application of ML techniques in RWD for disease prediction and management. The results reveal that machine learning methods, particularly ensemble models like RF, play a crucial role in enhancing prediction accuracy and addressing the complexities of large and high-dimensional datasets common in healthcare. Among the top ML methods employed, RF was the most widely used, featured in 24 studies, showcasing its adaptability to a variety of clinical datasets such as EHRs and patient registries. RF's ability to handle missing data, its resistance to overfitting, and its effectiveness in managing imbalanced datasets made it a powerful tool in predicting disease outcomes [115, 116], such as survival rates and complications in cardiovascular diseases and cancer. In terms of disease types, cardiovascular diseases dominated the studies, with 19 studies dedicated to predicting outcomes related to heart transplantation, atrial fibrillation, and peripheral artery disease. This concentration is likely attributed to the critical need for predictive tools in the early diagnosis and management of these conditions, which account for a significant burden on healthcare systems globally [117]. Machine learning applications, such as deep neural networks (DNN) and random survival forests (RSF), have been shown to improve the accuracy of survival predictions, assess treatment responses, and enhance patient stratification. Additionally, the study highlights the increasing application of ML in predicting conditions such as cancers, neurological disorders, and infectious diseases. These findings align with the broader trend of using RWD to bridge the gap between clinical trials and actual patient care by making predictions based on real-life data sources like EHRs and wearable devices. As evidenced in the studies reviewed, ML techniques can process vast amounts of medical data from various sources, facilitating early detection, timely intervention, and improved management of chronic conditions. Furthermore, these advancements in ML applications are subject to increasing regulatory oversight. Agencies such as the U.S. FDA and the EMA are actively exploring frameworks for the approval and regulation of ML-driven tools in healthcare. These regulations aim to ensure ML models' safety, efficacy, and transparency, especially in real-world applications where data variability and model interpretability remain key concerns. As regulatory bodies continue to define standards for using RWD and ML in clinical settings, ensuring compliance with FDA and EMA guidelines will be essential for the broader adoption and integration of these technologies into clinical practice.

Limitations

Despite the promising results, several limitations were observed across the studies reviewed. One key limitation is the heterogeneity in the data types used across studies. Many studies relied heavily on EHRs or patient registries, which are susceptible to biases such as missing data, errors in diagnosis coding, and lack of standardization across institutions. The quality and completeness of RWD can significantly influence the performance of machine learning models, leading to reduced generalizability across different healthcare settings. Additionally, while EHRs and registries are rich in data, they often lack detailed longitudinal information, which is crucial for predicting long-term outcomes and understanding disease progression over time. Another limitation is the generalizability of machine learning models across diverse patient populations. Many studies included in the review focused on specific patient groups or geographic regions, potentially limiting the applicability of findings to broader, more diverse populations. For example, certain ML algorithms might perform well in predicting outcomes for a specific ethnic group but may not be as effective when applied to a different demographic. Further research is needed to ensure these models are generalizable and equitable across diverse patient populations. Moreover, the interpretability of machine learning models remains a challenge, particularly for more complex algorithms such as deep neural networks. While these models can offer high predictive accuracy, their "black-box" nature makes it difficult for healthcare providers to understand the underlying decision-making process. This lack of transparency could impede the adoption of ML-based tools in clinical practice, where interpretability and trust in the model's decisions are crucial for clinical decision-making. Lastly, the application of machine learning models in clinical settings faces regulatory hurdles, particularly in relation to FDA and EMA guidelines. While these regulatory bodies are developing frameworks to approve ML-based tools, uncertainties about assessing model validity and ensuring their safety in real-world clinical use remain. Consequently, a lack of clear regulatory standards could delay the integration of ML tools into routine clinical practice.

Comparison with Prior Works

This systematic review aligns with and extends several recent literature reviews that have explored the application of ML to RWD in healthcare. Previous studies, such as those by Rajkomar et al. (2019) and Obermeyer et al. (2016), have highlighted the potential of ML models to transform healthcare by improving disease prediction and patient management [118, 119]. However, our review emphasizes a broader scope by including a wide variety of disease types, from cardiovascular diseases and cancer to neurological and infectious diseases, reflecting the growing versatility of ML tools in clinical settings. A notable comparison can be made with the work of Miotto et al. (2016), who focused on EHRs as the primary data source for ML models [120]. While their review identified the challenges associated with EHR-based studies, such as data sparsity and heterogeneity, our study similarly acknowledges these limitations but also expands the discussion to include wearable devices and patient registries as additional data sources. These emerging data types were shown to enhance model performance and enable more comprehensive patient monitoring in real-world settings. Another key comparison is with the review by Beam and Kohane (2018), which focused on ML's role in healthcare decision-making and its integration into clinical workflows [121]. While their work explored various ML algorithms in healthcare, our review places a stronger emphasis on the role of ensemble models like RF and their applicability across diverse healthcare datasets. Additionally, we underscore the need for regulatory clarity, particularly from the FDA and EMA, regarding the approval and safe deployment of ML models in clinical practice, an area that has been under-explored in prior reviews. Overall, this review builds upon the foundations set by previous literature but provides an updated and comprehensive analysis that incorporates new data sources, focuses on a broader range of diseases, and addresses the challenges of regulatory approval and model interpretability in the context of ML in healthcare.

Conclusions

In conclusion, integrating ML techniques into RWD has proven to be a transformative approach to disease prediction and management, offering significant improvements in predictive accuracy and patient stratification across various medical conditions. While methods like random forest, deep learning, and ensemble models have shown great promise in processing complex healthcare data, challenges related to data quality, generalizability, and model interpretability remain. Despite these limitations, the increasing adoption of ML in clinical settings, particularly for cardiovascular diseases and cancer, indicates its potential to revolutionize healthcare by enabling early diagnosis, personalized treatment plans, and optimized resource allocation. Future research should focus on enhancing model transparency, ensuring broader applicability across diverse populations, and addressing data inconsistencies to fully realize the potential of ML in clinical practice.

Data Availability

The original contributions presented in this study are included in the article/Multimedia Appendix, and further inquiries can be directed to the corresponding author.

Acknowledgements

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-RG23154).

Conflicts of Interest

None declared.

Multimedia Appendix 1

A dataset of selected studies, including details such as the year of publication, authors' names, study title, database, diseases/medical conditions, disease category, study type, real-world evidence (RWE) type, machine learning (ML) methods used, and key findings.

Abbreviations

AdaBoost	:	Adaptive Boosting
AHREs	:	Atrial High-rate Episodes
ANN	:	Artificial Neural Network
ARDS	:	Acute Respiratory Distress Syndrome
AUC	:	Area Under the Curve
BQR	:	Bayesian Quantile Regression
CNN	:	Convolutional Neural Network
DNN	:	Deep Neural Network
EHRs	:	Electronic Health Records
EMA	:	European Medicines Agency
FDA	:	Food and Drug Administration
GBM	:	Gradient Boosting Machine
GDPR	:	General Data Protection Regulation

JMIR Publications Preprint

HIPAA	:	Health Insurance Portability and Accountability Act
LR	:	Logistic Regression
LSCS	:	Lumbar Spinal Canal Stenosis
MeSH	:	Medical Subject Headings
ML	:	Machine Learning
MLR	:	Multivariable logistic regression
NB	:	Naive Bayes
NLP	:	Natural Language Processing
PULsE-AI	:	Prediction of Undiagnosed atriaL fibrillation using a machinE learning AlgorIthm
RF	:	Random Forest
PRISMA	:	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCTs	:	Randomized Controlled Trials
ResNet	:	Residual Neural Network
RSF	:	Random Survival Forest
RWD	:	Real-world Data
RWE	:	Real-world Evidence
SVM	:	Support Vector Machine
XGBoost	:	Extreme Gradient Boosting

References

1. Liu, F. and D. Panagiotakos, *Real-world data: a brief review of the methods, applications, challenges and opportunities.* BMC Med Res Methodol, 2022. **22**(1): p. 287.
2. Dang, A., *Real-World Evidence: A Primer.* Pharmaceut Med, 2023. **37**(1): p. 25-36.
3. Klonoff, D.C., *The Expanding Role of Real-World Evidence Trials in Health Care Decision Making.* J Diabetes Sci Technol, 2020. **14**(1): p. 174-179.
4. Junaid, S.B., et al., *Recent Advancements in Emerging Technologies for Healthcare Management Systems: A Survey.* Healthcare (Basel), 2022. **10**(10).
5. Williamson, S.M. and V. Prybutok, *Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare.* Applied Sciences, 2024. **14**(2): p. 675.
6. Derman, B.A., et al., *Reality check: Real-world evidence to support therapeutic development in hematologic malignancies.* Blood Reviews, 2022. **53**: p. 100913.
7. Rane, N., et al., *Machine Learning and Deep Learning for Big Data Analytics: A Review of Methods and Applications.* 2024.
8. Saeed, S., S. Ahmed, and s. joseph, *Machine Learning in the Big Data Age: Advancements, Challenges, and Future Prospects.* 2024.
9. Javaid, M., et al., *Significance of machine learning in healthcare: Features, pillars and applications.* International Journal of Intelligent Networks, 2022. **3**: p. 58-73.
10. Ogrizović, M., D. Drašković, and D. Bojić, *Quality assurance strategies for machine learning applications in big data analytics: an overview.* Journal of Big Data, 2024. **11**(1): p. 156.
11. Badalotti, D., et al., *Development of a Natural Language Processing (NLP) model to automatically extract clinical data from electronic health records: results from an Italian comprehensive stroke center.* International Journal of Medical Informatics, 2024. **192**: p. 105626.

Preprint JMIR Publications

12. Turchin, A. and L. Builes, *Using Natural Language Processing to Measure and Improve Quality of Diabetes Care: A Systematic Review.* Journal of Diabetes Science and Technology, 2021. **15**: p. 193229682110008.
13. Dominguez-Catena, I., et al., *Less can be more: representational vs. stereotypical gender bias in facial expression recognition.* Progress in Artificial Intelligence, 2024.
14. Ferrara, C., et al., *Fairness-aware machine learning engineering: how far are we?* Empirical Software Engineering, 2023. **29**(1): p. 9.
15. Hassija, V., et al., *Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence.* Cognitive Computation, 2024. **16**(1): p. 45-74.
16. Ali, S., et al., *Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence.* Information Fusion, 2023. **99**: p. 101805.
17. El Mestari, S.Z., G. Lenzini, and H. Demirci, *Preserving data privacy in machine learning systems.* Computers & Security, 2024. **137**: p. 103605.
18. lu, H., et al., *Patient Autonomy in Medical Education: Navigating Ethical Challenges in the Age of Artificial Intelligence.* INQUIRY: The Journal of Health Care Organization, Provision, and Financing, 2024. **61**.
19. Tucker, K., et al., *Protecting patient privacy when sharing patient-level data from clinical trials.* BMC Med Res Methodol, 2016. **16 Suppl 1**(Suppl 1): p. 77.
20. Franklin, G., et al., *The Sociodemographic Biases in Machine Learning Algorithms: A Biomedical Informatics Perspective.* Life (Basel), 2024. **14**(6).
21. FDA, *Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products*, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, and O.C.o. Excellence, Editors. 2023, FDA.
22. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* PLoS Med, 2009. **6**(7): p. e1000097.
23. Alotaiq, N. and D. Dermawan, *Advancements in Virtual Bioequivalence: A Systematic Review of Computational Methods and Regulatory Perspectives in the Pharmaceutical Industry.* Pharmaceutics, 2024. **16**(11): p. 1414.
24. Wissel, B.D., et al., *Automated, machine learning-based alerts increase epilepsy surgery referrals: A randomized controlled trial.* Epilepsia, 2023. **64**(7): p. 1791-1799.
25. Ayers, B., et al., *Using machine learning to improve survival prediction after heart transplantation.* J Card Surg, 2021. **36**(11): p. 4113-4120.
26. Nadarajah, R., et al., *Future Innovations in Novel Detection for Atrial Fibrillation (FIND-AF): pilot study of an electronic health record machine learning algorithm-guided intervention to identify undiagnosed atrial fibrillation.* Open Heart, 2023. **10**(2).
27. Yadgir, S.R., et al., *Machine learning-assisted screening for cognitive impairment in the emergency department.* J Am Geriatr Soc, 2022. **70**(3): p. 831-837.
28. Liu, Y., J. Xue, and J. Jiang, *Application of machine learning algorithms in electronic medical records to predict amputation-free survival after first revascularization in patients with peripheral artery disease.* Int J Cardiol, 2023. **383**: p. 175-184.
29. Hill, N.R., et al., *Identification of undiagnosed atrial fibrillation using a machine learning risk prediction algorithm and diagnostic testing (PULsE-AI) in primary care: cost-effectiveness of a screening strategy evaluated in a randomized controlled trial in England.* J Med Econ, 2022. **25**(1): p. 974-983.
30. Sheth, S.A., et al., *Machine Learning-Enabled Automated Determination of Acute Ischemic Core From Computed Tomography Angiography.* Stroke, 2019. **50**(11): p. 3093-3100.
31. Barton, C., et al., *Evaluation of a machine learning algorithm for up to 48-hour advance prediction of sepsis using six vital signs.* Comput Biol Med, 2019. **109**: p. 79-84.
32. Kao, Y.T., et al., *Machine Learning-Based Prediction of Atrial Fibrillation Risk Using*

Electronic Medical Records in Older Aged Patients. Am J Cardiol, 2023. **198**: p. 56-63.

33. Kim, M., et al., *Artificial intelligence predicts clinically relevant atrial high-rate episodes in patients with cardiac implantable electronic devices.* Sci Rep, 2022. **12**(1): p. 37.
34. Park, H.B., et al., *Risk factors based vessel-specific prediction for stages of coronary artery disease using Bayesian quantile regression machine learning method: Results from the PARADIGM registry.* Clin Cardiol, 2023. **46**(3): p. 320-327.
35. Hilbert, A., et al., *Data-efficient deep learning of radiological image data for outcome prediction after endovascular treatment of patients with acute ischemic stroke.* Comput Biol Med, 2019. **115**: p. 103516.
36. Chen, W., et al., *Using machine learning techniques predicts prognosis of patients with Ewing sarcoma.* J Orthop Res, 2021. **39**(11): p. 2519-2527.
37. Koutsouleris, N., et al., *Multisite prediction of 4-week and 52-week treatment outcomes in patients with first-episode psychosis: a machine learning approach.* Lancet Psychiatry, 2016. **3**(10): p. 935-946.
38. Strömblad, C.T., et al., *Effect of a Predictive Model on Planned Surgical Duration Accuracy, Patient Wait Time, and Use of Presurgical Resources: A Randomized Clinical Trial.* JAMA Surg, 2021. **156**(4): p. 315-321.
39. Wang, S.V., et al., *Stepped-wedge randomised trial to evaluate population health intervention designed to increase appropriate anticoagulation in patients with atrial fibrillation.* BMJ Qual Saf, 2019. **28**(10): p. 835-842.
40. Tan, T.H., et al., *Predicting outcomes in older ED patients with influenza in real time using a big data-driven and machine learning approach to the hospital information system.* BMC Geriatr, 2021. **21**(1): p. 280.
41. Goerigk, S., et al., *Predicting instructed simulation and dissimulation when screening for depressive symptoms.* European Archives of Psychiatry and Clinical Neuroscience, 2020. **270**(2): p. 153-168.
42. Kijpaisalratana, N., et al., *Real-time machine learning-assisted sepsis alert enhances the timeliness of antibiotic administration and diagnostic accuracy in emergency department patients with sepsis: a cluster-randomized trial.* Internal and Emergency Medicine, 2024. **19**(5): p. 1415-1424.
43. Sharma, A., et al., *Patient Phenotypes, Cardiovascular Risk, and Ezetimibe Treatment in Patients After Acute Coronary Syndromes (from IMPROVE-IT).* Am J Cardiol, 2019. **123**(8): p. 1193-1201.
44. Singhal, L., et al., *eARDS: A multi-center validation of an interpretable machine learning algorithm of early onset Acute Respiratory Distress Syndrome (ARDS) among critically ill adults with COVID-19.* PLoS One, 2021. **16**(9): p. e0257056.
45. Kanchanatawan, B., et al., *Episodic memory and delayed recall are significantly more impaired in younger patients with deficit schizophrenia than in elderly patients with amnestic mild cognitive impairment.* PLoS One, 2018. **13**(5): p. e0197004.
46. Huang, J., et al., *Twenty-eight-day in-hospital mortality prediction for elderly patients with ischemic stroke in the intensive care unit: Interpretable machine learning models.* Front Public Health, 2022. **10**: p. 1086339.
47. She, H., et al., *Metabolomics and machine learning approaches for diagnostic and prognostic biomarkers screening in sepsis.* BMC Anesthesiol, 2023. **23**(1): p. 367.
48. Sundar, R., et al., *Machine-learning model derived gene signature predictive of paclitaxel survival benefit in gastric cancer: results from the randomised phase III SAMIT trial.* Gut, 2022. **71**(4): p. 676-685.
49. Alaa, A.M., et al., *Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants.* PLoS One, 2019. **14**(5): p. e0213653.

Preprint JMIR Publications

50. Azimi, P., et al., *Use of artificial neural networks to decision making in patients with lumbar spinal canal stenosis.* J Neurosurg Sci, 2017. **61**(6): p. 603-611.
51. Baxter, S.L., et al., *Machine Learning-Based Predictive Modeling of Surgical Intervention in Glaucoma Using Systemic Data From Electronic Health Records.* Am J Ophthalmol, 2019. **208**: p. 30-40.
52. Anderson, J.P., et al., *Reverse Engineering and Evaluation of Prediction Models for Progression to Type 2 Diabetes: An Application of Machine Learning Using Electronic Health Records.* J Diabetes Sci Technol, 2015. **10**(1): p. 6-18.
53. Bannister, C.A., et al., *A genetic programming approach to development of clinical prediction models: A case study in symptomatic cardiovascular disease.* PLoS One, 2018. **13**(9): p. e0202685.
54. Scheer, J.K., et al., *Development of a preoperative predictive model for major complications following adult spinal deformity surgery.* J Neurosurg Spine, 2017. **26**(6): p. 736-743.
55. Rau, H.H., et al., *Development of a web-based liver cancer prediction model for type II diabetes patients by using an artificial neural network.* Comput Methods Programs Biomed, 2016. **125**: p. 58-65.
56. Ramezankhani, A., et al., *Decision tree-based modelling for identification of potential interactions between type 2 diabetes risk factors: a decade follow-up in a Middle East prospective cohort study.* BMJ Open, 2016. **6**(12): p. e013336.
57. Pei, D., et al., *Identification of Potential Type II Diabetes in a Chinese Population with a Sensitive Decision Tree Approach.* J Diabetes Res, 2019. **2019**: p. 4248218.
58. Oviedo, S., et al., *Risk-based postprandial hypoglycemia forecasting using supervised learning.* Int J Med Inform, 2019. **126**: p. 1-8.
59. Mubeen, A.M., et al., *A six-month longitudinal evaluation significantly improves accuracy of predicting incipient Alzheimer's disease in mild cognitive impairment.* J Neuroradiol, 2017. **44**(6): p. 381-387.
60. Lopez-de-Andres, A., et al., *Predictors of in-hospital mortality following major lower extremity amputations in type 2 diabetic patients using artificial neural networks.* BMC Med Res Methodol, 2016. **16**(1): p. 160.
61. Kwon, J.M., et al., *Deep-learning-based out-of-hospital cardiac arrest prognostic system to predict clinical outcomes.* Resuscitation, 2019. **139**: p. 84-91.
62. Kim, I., et al., *A predictive model for high/low risk group according to oncotype DX recurrence score using machine learning.* Eur J Surg Oncol, 2019. **45**(2): p. 134-140.
63. Khanji, C., et al., *Lasso Regression for the Prediction of Intermediate Outcomes Related to Cardiovascular Disease Prevention Using the TRANSIT Quality Indicators.* Med Care, 2019. **57**(1): p. 63-72.
64. Karhade, A.V., et al., *Development of machine learning algorithms for prediction of prolonged opioid prescription after surgery for lumbar disc herniation.* Spine J, 2019. **19**(11): p. 1764-1771.
65. Jovanovic, P., N.N. Salkic, and E. Zerem, *Artificial neural network predicts the need for therapeutic ERCP in patients with suspected choledocholithiasis.* Gastrointest Endosc, 2014. **80**(2): p. 260-8.
66. Kang, A.R., et al., *Development of a prediction model for hypotension after induction of anesthesia using machine learning.* PLOS ONE, 2020. **15**(4): p. e0231172.
67. Isma'eel, H.A., et al., *Artificial neural network-based model enhances risk stratification and reduces non-invasive cardiac stress imaging compared to Diamond-Forrester and Morise risk assessment models: A prospective study.* J Nucl Cardiol, 2018. **25**(5): p. 1601-1609.
68. Hill, N.R., et al., *Predicting atrial fibrillation in primary care using machine learning.* PLoS One, 2019. **14**(11): p. e0224582.
69. Dong, Y., et al., *A novel surgical predictive model for Chinese Crohn's disease patients.*

Medicine (Baltimore), 2019. **98**(46): p. e17510.
70. Bowman, A., et al., *A prognostic model for the patient-reported outcome of surgical treatment of carpal tunnel syndrome.* Muscle Nerve, 2018. **58**(6): p. 784-789.
71. Bertsimas, D., et al., *Applied Informatics Decision Support Tool for Mortality Predictions in Patients With Cancer.* JCO Clin Cancer Inform, 2018. **2**: p. 1-11.
72. Manz, C.R., et al., *Effect of Integrating Machine Learning Mortality Estimates With Behavioral Nudges to Clinicians on Serious Illness Conversations Among Patients With Cancer: A Stepped-Wedge Cluster Randomized Clinical Trial.* JAMA Oncol, 2020. **6**(12): p. e204759.
73. Tian, D., et al., *Machine Learning-Based Prognostic Model for Patients After Lung Transplantation.* JAMA Netw Open, 2023. **6**(5): p. e2312022.
74. Li, E., et al., *Oncologist phenotypes and associations with response to a machine learning-based intervention to increase advance care planning: Secondary analysis of a randomized clinical trial.* PLoS One, 2022. **17**(5): p. e0267012.
75. Tedeschi, S.K., et al., *Classifying Pseudogout Using Machine Learning Approaches With Electronic Health Record Data.* Arthritis Care Res (Hoboken), 2021. **73**(3): p. 442-448.
76. Ambwani, G., et al., *PPM8 A MACHINE LEARNING MODEL FOR CANCER BIOMARKER IDENTIFICATION IN ELECTRONIC HEALTH RECORDS.* Value in Health, 2019. **22**: p. S334.
77. Jorge, A., et al., *Identifying lupus patients in electronic health records: Development and validation of machine learning algorithms and application of rule-based algorithms.* Semin Arthritis Rheum, 2019. **49**(1): p. 84-90.
78. Shimabukuro, D.W., et al., *Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial.* BMJ Open Respir Res, 2017. **4**(1): p. e000234.
79. Sarraju, A., et al., *Machine learning approaches improve risk stratification for secondary cardiovascular disease prevention in multiethnic patients.* Open Heart, 2021. **8**(2).
80. Ye, B., et al., *Discrimination of indoor versus outdoor environmental state with machine learning algorithms in myopia observational studies.* J Transl Med, 2019. **17**(1): p. 314.
81. Sun, Z., et al., *An improved random forest based on the classification accuracy and correlation measurement of decision trees.* Expert Systems with Applications, 2024. **237**: p. 121549.
82. Bandyopadhyay, A., et al., *Using real-world electronic health record data to predict the development of 12 cancer-related symptoms in the context of multimorbidity.* JAMIA Open, 2024. **7**(3): p. ooae082.
83. Devendra, M., et al., *Enhancing Prediction Accuracy in Healthcare Data Using Advanced Data Mining Techniques.* 2024.
84. Ghosh, D. and J. Cabrera, *Enriched Random Forest for High Dimensional Genomic Data.* IEEE/ACM Trans Comput Biol Bioinform, 2022. **19**(5): p. 2817-2828.
85. Maalouf, M., *Logistic regression in data analysis: An overview.* International Journal of Data Analysis Techniques and Strategies, 2011. **3**: p. 281-299.
86. Boateng, E.Y. and D. Abaye, *A Review of the Logistic Regression Model with Emphasis on Medical Research.* Journal of Data Analysis and Information Processing, 2019. **07**: p. 190-207.
87. Shipe, M.E., et al., *Developing prediction models for clinical use using logistic regression: an overview.* J Thorac Dis, 2019. **11**(Suppl 4): p. S574-s584.
88. Zabor, E., et al., *Logistic Regression in Clinical Studies.* International Journal of Radiation Oncology*Biology*Physics, 2021. **112**.
89. Guido, R., et al., *An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review.* Information, 2024. **15**(4): p. 235.

90. Huang, S., et al., *Applications of Support Vector Machine (SVM) Learning in Cancer Genomics.* Cancer Genomics Proteomics, 2018. **15**(1): p. 41-51.
91. Pisner, D. and D. Schnyer, *Support vector machine*. 2020. p. 101-121.
92. Son, Y.J., et al., *Application of support vector machine for prediction of medication adherence in heart failure patients.* Healthc Inform Res, 2010. **16**(4): p. 253-9.
93. Manikandan, G., et al., *Classification models combined with Boruta feature selection for heart disease prediction.* Informatics in Medicine Unlocked, 2024. **44**: p. 101442.
94. Chen, T. and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*. 2016. 785-794.
95. Nadkarni, S.B., G.S. Vijay, and R.C. Kamath, *Comparative Study of Random Forest and Gradient Boosting Algorithms to Predict Airfoil Self-Noise.* Engineering Proceedings, 2023. **59**(1): p. 24.
96. Adeola, O. and W. Qing-Guo, *XGBoost Model for Chronic Kidney Disease Diagnosis.* IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2019. **PP**: p. 1-1.
97. Chakraborty, A. and C.P. Tsokos, *An AI-driven Predictive Model for Pancreatic Cancer Patients Using Extreme Gradient Boosting.* Journal of Statistical Theory and Applications, 2023. **22**(4): p. 262-282.
98. Zou, J., Y. Han, and S.-S. So, *Overview of Artificial Neural Networks.* Methods in molecular biology (Clifton, N.J.), 2009. **458**: p. 14-22.
99. Kufel, J., et al., *What Is Machine Learning, Artificial Neural Networks and Deep Learning?-Examples of Practical Applications in Medicine.* Diagnostics (Basel), 2023. **13**(15).
100. Ahmed, F.E., *Artificial neural networks for diagnosis and survival prediction in colon cancer.* Mol Cancer, 2005. **4**: p. 29.
101. Abbas, S., et al., *Emerging research trends in artificial intelligence for cancer diagnostic systems: A comprehensive review.* Heliyon, 2024. **10**(17): p. e36743.
102. Sirocchi, C., A. Bogliolo, and S. Montagna, *Medical-informed machine learning: integrating prior knowledge into medical decision systems.* BMC Med Inform Decis Mak, 2024. **24**(Suppl 4): p. 186.
103. Rostami, M. and M. Oussalah, *A novel explainable COVID-19 diagnosis method by integration of feature selection with random forest.* Informatics in Medicine Unlocked, 2022. **30**: p. 100941.
104. Mohsin, S.N., et al., *The Role of Artificial Intelligence in Prediction, Risk Stratification, and Personalized Treatment Planning for Congenital Heart Diseases.* Cureus, 2023. **15**(8): p. e44374.
105. Natekin, A. and A. Knoll, *Gradient Boosting Machines, A Tutorial.* Frontiers in neurorobotics, 2013. **7**: p. 21.
106. Ji, G.W., et al., *Development and validation of a gradient boosting machine to predict prognosis after liver resection for intrahepatic cholangiocarcinoma.* BMC Cancer, 2022. **22**(1): p. 258.
107. Hossain, E., et al., *Natural Language Processing in Electronic Health Records in relation to healthcare decision-making: A systematic review.* 2023. **155**: p. 106649.
108. Rojas-Carabali, W., et al., *Natural Language Processing in medicine and ophthalmology: A review for the 21st-century clinician.* Asia-Pacific Journal of Ophthalmology, 2024. **13**(4): p. 100084.
109. Sheikhalishahi, S., et al., *Natural Language Processing of Clinical Notes on Chronic Diseases: Systematic Review.* JMIR Med Inform, 2019. **7**(2): p. e12239.
110. George, B., S. Seals, and I. Aban, *Survival analysis and regression models.* J Nucl Cardiol, 2014. **21**(4): p. 686-94.
111. Beis, G., A. Iliopoulos, and I. Papasotiriou, *An Overview of Introductory and Advanced Survival Analysis Methods in Clinical Applications: Where Have we Come so far?* Anticancer Research, 2024. **44**(2): p. 471.

112. Wassan, S., et al., *Deep convolutional neural network and IoT technology for healthcare.* DIGITAL HEALTH, 2024. **10**: p. 20552076231220123.
113. Nazir, S., D.M. Dickson, and M.U. Akram, *Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks.* Computers in Biology and Medicine, 2023. **156**: p. 106668.
114. Lundervold, A.S. and A. Lundervold, *An overview of deep learning in medical imaging focusing on MRI.* Zeitschrift für Medizinische Physik, 2019. **29**(2): p. 102-127.
115. Genuer, R., et al., *Random Forests for Big Data.* Big Data Research, 2015. **9**.
116. Khalilia, M., S. Chakraborty, and M. Popescu, *Predicting disease risks from highly imbalanced data using random forest.* BMC Med Inform Decis Mak, 2011. **11**: p. 51.
117. Kumar, Y., et al., *Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda.* J Ambient Intell Humaniz Comput, 2023. **14**(7): p. 8459-8486.
118. Rajkomar, A., J. Dean, and I. Kohane, *Machine Learning in Medicine.* N Engl J Med, 2019. **380**(14): p. 1347-1358.
119. Obermeyer, Z., et al., *Dissecting racial bias in an algorithm used to manage the health of populations.* Science, 2019. **366**(6464): p. 447-453.
120. Miotto, R., et al., *Deep learning for healthcare: review, opportunities and challenges.* Brief Bioinform, 2018. **19**(6): p. 1236-1246.
121. Beam, A.L. and I.S. Kohane, *Big Data and Machine Learning in Health Care.* Jama, 2018. **319**(13): p. 1317-1318.

Preprint JMIR Publications

Supplementary Files

Preprint JMIR Publications

Multimedia Appendixes

A dataset of selected studies, including details such as the year of publication, authors' names, study title, database, diseases/medical conditions, disease category, study type, real-world evidence (RWE) type, machine learning (ML) methods used, and key findings.
URL: http://asset.jmir.pub/assets/4bf13f7216431aa9971e8752d6d0c0fe.xlsx