electronics

Review

Review of Time Domain Electronic Medical Record Taxonomies in the Application of Machine Learning

Haider Ali [1], Imran Khan Niazi [1,2,3], Brian K. Russell [1,4], Catherine Crofts [1], Samaneh Madanian [1] and David White [1,*]

[1] BioDesign Lab, School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, Auckland 1010, New Zealand
[2] Centre for Chiropractic Research, New Zealand College of Chiropractic, Auckland 1060, New Zealand
[3] Center for Sensory-Motor Interaction, Department of Health Science and Technology, Aalborg University, 9220 Aalborg Øst, Denmark
[4] Ambient Cognition Limited, Auckland 1010, New Zealand
* Correspondence: david.white@biodesignlab.co.nz; Tel.: +64-211-23-5470

Abstract: Electronic medical records (EMRs) help in identifying disease archetypes and progression. A very important part of EMRs is the presence of time domain data because these help with identifying trends and monitoring changes through time. Most time-series data come from wearable devices monitoring real-time health trends. This review focuses on the time-series data needed to construct complete EMRs by identifying paradigms that fall within the scope of the application of artificial intelligence (AI) based on the principles of translational medicine. (1) Background: The question addressed in this study is: What are the taxonomies present in the field of the application of machine learning on EMRs? (2) Methods: Scopus, Web of Science, and PubMed were searched for relevant records. The records were then filtered based on a PRISMA review process. The taxonomies were then identified after reviewing the selected documents; (3) Results: A total of five main topics were identified, and the subheadings are discussed in this review; (4) Conclusions: Each aspect of the medical data pipeline needs constant collaboration and update for the proposed solutions to be useful and adaptable in real-world scenarios.

Keywords: time series; electronic medical records; systemic review; artificial intelligence; machine learning

Citation: Ali, H.; Niazi, I.K.; Russell, B.K.; Crofts, C.; Madanian, S.; White, D. Review of Time Domain Electronic Medical Record Taxonomies in the Application of Machine Learning. *Electronics* **2023**, *12*, 554. https://doi.org/10.3390/electronics12030554

Academic Editors: Mohammed Abdulhakim Al-Absi and Rui Fu

Received: 13 December 2022
Revised: 15 January 2023
Accepted: 19 January 2023
Published: 21 January 2023

1. Introduction

Translational medicine (TM) includes collaboration between medical clinicians, biomedical engineers, and scientists to develop artificial intelligence (AI) models that account for differing data sources, data collection methods, and other real-world factors. In some cases, TM reduces the time from solution development to deployment. It does so by allowing effective communication between different players for shared goals. It is characterized by integrating digital biomarkers, multi-omics profiling, model-based data, AI, biomarker-guided trial designs, and patient-centric companion diagnostics. Therefore, the taxonomies identified in this review are guided by translational medicine [1]. Ref [2] presents the following components of the complete medical record including: connected fitness devices, patient-focused personal health records, individual behavioral patterns, pharmacy-focused medical adherence data, provider-focused medical records, connected medical devices, and genomic information.

In these complete electronic medical records (EMRs), the use of time-series data is essential because most of the biomarkers are tracked as trends in time [3]. EMRs help in disease monitoring, pandemic monitoring, adjustment of lifestyles, hospitals, intensive care units, and integration of healthcare services. Machine learning (ML) is "The ability of computers to advise decisions based on the available data" [4]. ML has been used in applied clinical studies for some time now [5], and recent renewed interest has been driven by increased data availability [6] and an increase in computational capacity [7]. The most common medical data types are images, such as computerized tomography (CT) scans, time-series data, and blood, urine, and metabolic panels. Some noteworthy literature reviews cover computer vision techniques to assist clinical decision-making [8–10]. However, there is a need to review the broader research landscape relating to the application of ML in time-series data in EMRs to give the relevant taxonomies, patterns in literature, emerging trends, and knowledge streams in this field. Time series can be defined as repeated readings from a device over a period of time. The frequency of the readings can be periodic or non-periodic and range from thousands of times per second (e.g., accellerometry) to a few times a day (e.g., glucose monitoring).

Our review of the relevant studies has found numerous pertinent publications [11], including a previous work by Davy et al. (2015) that investigates the effectiveness of chronic care AI models employed in primary healthcare [12]. A more recent work by Chen et al. (2020) reviews the use of probabilistic ML models applied to healthcare data [13], while Wang et al. (2021) examined the latest advancements in graph-based analytics in healthcare [14]. The application of telemedicine in maintaining electronic health records was recently reviewed by Gu et al. (2019) using a cite space analysis [15]. Cite space was also used in a scientometric review of the application of latent discriminant analysis in healthcare data by Tean et al. (2019) [16]. Emerging challenges when using unstructured EMRs were deliberated by Adnan et al. (2020) in the context of using big unstructured data in healthcare [17].

In comparison to the studies above our study has three novel features:

1. It identifies taxonomies within the field after a systemic search of research databases.
2. It finds these taxonomies based on the principles of translational medicine so that the reader may find all the information needed for a translational solution in one place.
3. It identifies the core challenges and advancements in each taxonomy and provides a rigorous volume of the literature to serve as a baseline.

The following are the research questions we will try to address:

1. What paradigms fall under the umbrella of AI in time series and graph-based healthcare data?
2. What are the latest advances in these domains?
3. What are the latest challenges in these taxonomies?

Although earlier literature reviews have covered the application of specific algorithms and problems in EMRs, as shown in Table 1, a need exists to develop taxonomies and discuss recent findings.

Table 1. Comparison with earlier works.

References	Time Series	Disease Specific	Translational Medicine
[12]	✓	X	X
[18]	✓	✓	X
[19]	X	✓	X
[16]	X	✓	X
[15]	✓	✓	X
[17]	✓	✓	X
Our work	✓	X	✓

2. Materials and Methods

A paradigm is a collection of elements based on a common lexeme [20]. Identifying the paradigms of fields interacting with each other is imperative to achieving overarching solutions. In this review, these paradigms are based on an industry perspective of TM [1] and TM principles. The three Ts in TM are developing treatments and interventions, testing the proposed interventions' effectiveness, and deploying these applications in the real world [21]. In this review, all the paradigms in the field considered are presented by Figure 1, and the challenges in applying AI-based solutions to TM are also discussed.

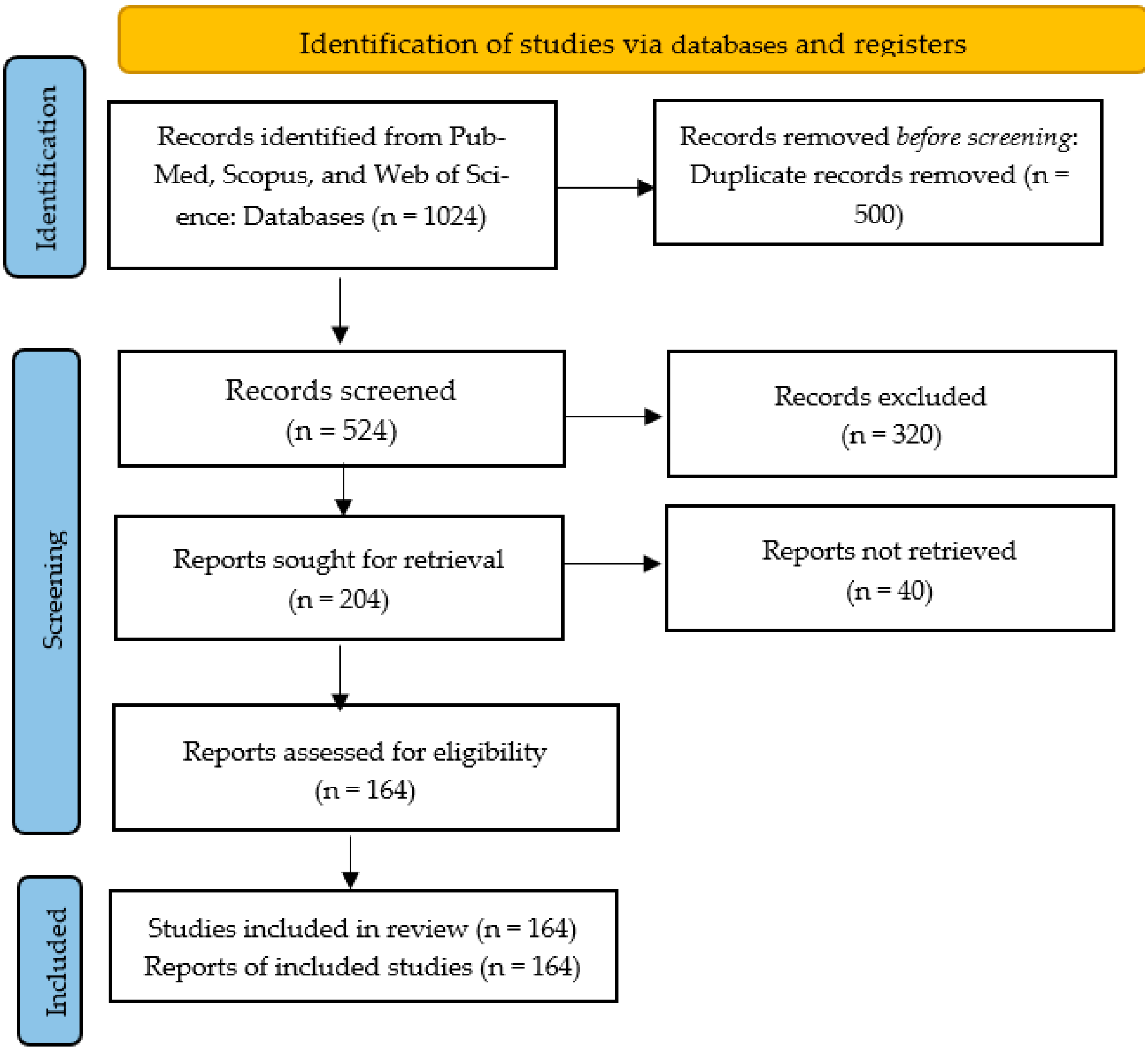

Figure 1. PRISMA review process for selection of records from research databases.

In this review, the literature was searched from the following databases: PubMed, Scopus, and Web of Science, and the papers between 2015–2022 were selected. The search terms were (("Physiological sensors" OR "Biomedical sensors" OR "Bio-medical Sensors") AND ("Machine Learning" OR "ML" OR "Artificial Intelligence" OR "AI" OR "Deep learning" OR "DL" OR "Reinforcement learning" OR "Electronic health records") NOT ("Security and Privacy") NOT ("images") NOT ("Robot")). The search was limited from 2015 to November 2022. Only articles were included, and reviews were not made part of the search criteria. The proceedings of various conferences were excluded from the search. Applying these filters and only selecting English language records resulted in 320 articles being removed, resulting in the number of articles selected for review totaling 164. These records were then collected and thoroughly read to answer the following questions:

- What is the type of data?
- What kind of algorithm is used?
- What pre-processing methods are used?
- What post-processing methods are used?
- What data privacy standards are observed?

- What interoperability or fusion techniques are used?

By answering these questions in the form of a table, the top three methods are identified within each paradigm. However, if the top three topics in a paradigm are repeated in any other paradigm, the next three topics are also discussed in the section to give a holistic overview of the topic. For example, the top three topics in the subtopic: time-series data and structured data are the same and, therefore, positioned in the structured data section.

An Ishikawa Fishbone Diagram (Figure 2) presents the different paradigms on the topic.

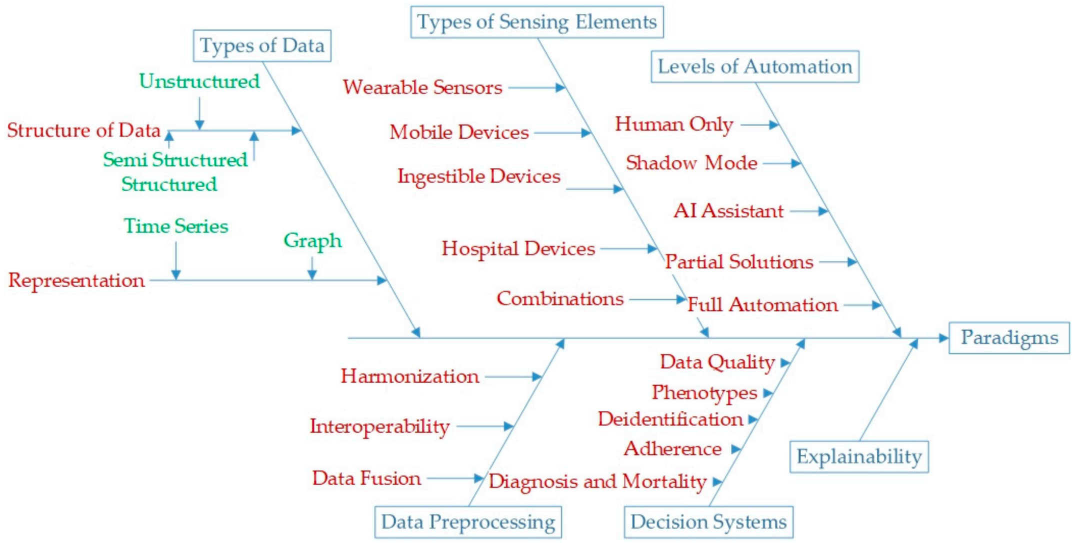

Figure 2. Ishikawa Fishbone diagram of the paradigms.

Figure 2 is a representation of the paradigms found in the literature under the topic at hand. The inclined axes represent the major topics in the field, for example, types of data and levels of automation. These inclined axes are also divided, shown as horizontal arrows, to identify the relevant subtopics.

3. Results

After collecting the records through the process previously described, then reading their methods and results, using TM principles, we enabled the paradigms to then be identified. The subsequent discussion is arranged as follows; first, we identify the most commonly occurring topics in the records that we have collected within each paradigm, and then the challenges within the paradigms are elaborated upon. This review is a combination of a narrative review built on systematic data collection.

3.1. Types of Data

One of the most obvious choices of paradigms is the type of data, which can depend upon the source of data (different types of sensors), organization of data (structured, unstructured), and representation of data (time series, images, graph representations).

Representation of Data

The way that the data are represented makes them suitable for a particular type of analysis to then be undertaken. The inductive biases of different ML models can be matched to different types of data representation. For example, as we will see in later parts of this review, the graph-based data representations are widely used for phenotyping (phenotype of the collection of observable effects of a disease), since the inductive bias of a graph-based ML model suits the relational nature of the problem of phenotyping. Thus, the first axis, along which we separated the data, is according to how it is being represented. This can take tabular, time series, or graph forms.

1. Time series (tabular representation): Time-series data contain information from physiological events in the form of time-varying biomarkers. Three leading solutions are specific to this data type: motif or pattern detection, data generation/imputation, and time-series forecasting. Generative AI models can potentially overcome the lack of access to time-series data by synthetically producing the missing and unknown data; however, accuracy by patient needs to be proven for each application, where missing data can be missing readings within a time series or the complete absence of a time-series recording in the EHR. Guided evolutionary networks (GENs) combine artificial neural networks and optimization algorithms such as genetic algorithms. These are used to fuse various information sources [22,23]. GENs are also used to discover time-series motifs in ECG data [24]. Ref. [25] uses a multilayer perceptron for time-series forecasting in healthcare data. The following Table 2 presents a comparison of the representative literature.

Table 2. Comparison of the time-series solution.

References	Applications	Sensors	Generative	Predictive	Clinical	Imputation
[26]	Motif Discovery	ECG	✓	X	✓	X
[22]	Motif Discovery	ECG and EEG	✓	✓	X	✓
[24]	Anomaly Detection	ECG	X	✓	✓	X
[25]	Expenditure Calculation	Healthcare data	X	✓	✓	X
[27]	Benchmarking	MIMIC-III	X	✓	✓	X
[28]	Imputation	ECG, MIMIC	X	✓	✓	✓

2. Graph representation: Healthcare data are relational, which makes them suitable for graphical representation. Relational data are characterized by the relations or dependence that exists amongst the rows and columns [29]. Graph-based techniques are used for developing graph-based representations of healthcare data, identifying clinical pathways and phenotypes of disease, and performing predictive modelling of disease and interventions. For example, refs. [30,31] are some typical graph representations of healthcare data. Ref. [32] determines the temporal phenotypes based on graph representations of healthcare data. Ref. [33] is a fog-based temporal network graph analysis for the Chikungunya virus in India. Ref. [34] uses a proximity-preserving graph embedding to represent electronic health records for hypertension. Ref. [35] incorporates metadata of the patients along with their vitals and lab results to learn a graph representation of electronic healthcare data. Ref. [36] is a study that employs cryptographic techniques for information embedding in the healthcare data. Ref. [37] is another knowledge-graph-based phenotyping technique for subarachnoid hemorrhage. Ref. [38] is a graph-based visualization for sensitive outcomes in medicine for healthcare data. Ref. [39] is a graph-based channel fusion for wrist pulse detection. Ref. [40] uses graphs for learning a lower dimensional representation of drug–disease interaction. As illustrated in [41], the main applications of graphs in medical interventions are drug–drug interaction, drug–disease interaction, protein–protein interaction, medical term classification, and protein function prediction. The three main methods to realize these ends are matrix factorization, random walk, and neural network-based methods. These include Laplacian methods, as demonstrated in [42], deep walk methods, as shown in [43], and neural networks, as illustrated in [40]. Graph algorithms commonly used can be categorized into temporal data mining [44], causal and contextual [45], and patient enteric graphs [46]. It is worth noting that there is no unique graph representation for sensor data or electronic medical records. Hence, most research focuses on developing graph-based presentations. One crucial research area is benchmarking and creating a numeric qualitative marker of adequate representation. There are several limitations of time-series- and graph-based healthcare

data; these include data sparsity [47], noise [48], limited generalizability [49], and lack of context [49].

The following Table 3 presents a comparison of the representative literature.

Table 3. Comparison of graph-based solutions.

References	Application	Techniques Used	Data	Contributions	Predictive	Descriptive
[32]	Temporal Phenotyping	Attention Models	MIMIC-III	10% greater than RNN in disease prediction and 3% improved areas under ROC	✓	✓
[38]		Hinge Loss		Predicted congestive health failure with an 80% accuracy. The area under the curve for patient readmission increased by over 50% from the spectral clustering	✓	✓
[36]	Graph representation	Note Binning	STRIDE	Developed term and concept mappings	X	✓
[39]	Feature fusion	Multi-Channel feature fusion	Pressure and Photo–electric Sensors	93.1% accuracy in predicting diabetes from pulse detection data.	X	X

3.2. Structure of Data

Another way to classify the type of healthcare data is the structure of available data. Most healthcare data are not structured against a set of rules. The structure of data dictates the kind of preprocessing required or the kind of algorithms that can be used.

A. Structured data: This follows a definite set of rules or schemes [50]. The main issues when using ML and structured data are data generation, data fusion, pattern detection, privacy preservation, and prediction of outcomes. Privacy preservation is guided by HIPAA rules [51]. Generative algorithms are used extensively to impute the missing data in the structured datasets [52,53]. Data fusion is another typical application of ML for combining two different kinds of structured data [54,55]. Federated learning that trains the models based on data from various decentralized devices is used extensively for privacy preservation of healthcare data [56–59]. ML and structured data are also valuable in predicting the outcomes of interventions, for example, [60] analyzes the user's choice in the event of alerts from clinical decision systems for potential drug–drug interference. Ref. [61] uses structured and unstructured data to find the social determinants of health characterized by social behavior, demographic features, and environmental factors of medical status and health care access. Ref. [62] is a systemic review of records from PubMed and Web of Science on the detection of strokes from structured data that found the leading keyword to be mortality and the most used algorithms to be neural networks, support vector machines, and XGBoost. Ref. [63] is another review that looked at the statistical and predictive machine learning models for cancer risk and found the cox model [64] is the most commonly used algorithm for predicting disease onset based on the input features. Ref. [65] used AI to auto-complete structured clinical records based on context. Ref. [66] is a model to detect probable cases of dementia using structured and unstructured data that uses a latent Dirichlet algorithm for feature extraction and a logistic regression model. The key issues of research for structured data in healthcare are detecting phenotypes from electronic health records [67,68], privacy and encoding of information [69–72], data harmonization from various sources [72], synthetic data generation for research [73–75], and fairness and bias in the structured data [76].

B. Semi-structured data: These EMRs have no specific structure, enabling categorical data, meta-data, and numerical data to be entered in any field. The key areas in application of ML in unstructured data is in the conversion to structured data, predictive modeling, and interoperability of different kinds of data sources. For example, an application of ML with unstructured data for predictive modeling is used [77] to derive contextual information to generate semi-structured data from electronic medical records. Ref. [78] is a method to allocate resources from the knowledge of semi-structured healthcare data. Ref. [79] uses HL7 standards to develop the interoperability of structured, semi-structured, and unstructured data to develop obesity phenotypes. Ref. [80] is another such system that uses open EMRs to this end. Ref. [81] detects autism from semi-structured and unstructured data using a combination of skip-gram models.

C. Unstructured data: Most EMRs are unstructured [82]. Key research areas for ML applications in unstructured data are conversion amongst the various kinds of data structure and predictive modeling. An example of predictive modeling using unstructured data [83] employs unstructured EMRs to phenotype depression in youth. Latent Dirichlet Analysis (LDA) and other dimensionality reduction methods are used to obtain the hidden information between different kinds of data and then leverage it for predictive modeling [84–87]. A priori algorithms and other Bayesian methods are used to convert unstructured data to structured data [37,88,89], and in so doing, these works can also combine with structured data to make predictions [90,91]. Another technique that is relevant to the conversion of unstructured data to structured data is distant supervision. Distant supervision is a method for labeling the data by utilizing the known structures of similar data [92,93]. Exploratory text analysis is also used for pattern analysis for predictive modeling in this [94,95].

The following Table 4 compares these techniques:

Table 4. Comparison of unstructured data.

References	Application	Techniques Used	Evaluation Metrics	Structured Data
[83]	Detection of clinical depression	NLP	Specificity:97%. Sensitivity:45%	X
[84]	Disease prediction	LDA	AUC 0.94, Sensitivity 0.87 and Specificity 0.87	✓
[94]	HPV detection	NLP	AUC: 0.861	X
[92]	Breast cancer detection	NLP	AUC 0.91, Sensitivity: 0.861, Specificity 0.878, Accuracy 0.870.	✓

The different ML techniques used in conjunction with unstructured data are clustering, classification, boosting, and a combination of these three. Clustering can help with phenotyping and grouping together different clinical pathways. Classification requires labeling the data, which can be taxing for a large volume of clinical notes. Boosting models can leverage the different structures present in unstructured data to make meaningful predictions, especially, risk and mortality.

Natural language processing techniques are extensively applied to unstructured data to detect disease onset. Data harmonization and standardization is also an essential topic of discussion in unstructured healthcare.

In the healthcare context, structured and semi-structured data are typically easier to work with and analyze because they have some inherent structure. Unstructured data, such as free-text notes in electronic health records, can be more challenging to work with because they require more effort to extract meaningful information.

3.3. Types of Sensing Elements

Types of data are dependent on the types of sensing elements used. There are many types of sensing elements including wearable sensors, mobile device sensors, ingestible sensors, medical devices from hospitals, and a combination of all or some of the factors mentioned earlier.

- Wearable sensors: These bridge the gap between assessment and onset prediction. The data sources measure the biomarkers from the physiological signals in real-time, making this a vital component of multi-omics profiling [96].
- Mobile devices: Along with real-time monitoring using mobile sensors, mobile devices also allow for input from the user, making them helpful in tracking medical adherence [97].
- Ingestible sensors: Drug adherence [98] and monitoring [99] are some applications of ingestible sensors.
- Medical devices from hospitals: These include connected medical devices intended to enhance healthcare quality for people in the hospital [100].
- Combinations: The combination of the sensors enables the Internet of Medical Devices [101].

The critical limitations of wearable sensors are the contextualization of data and integration with the existing clinical care pathways; hence, a challenge exists to show clinical efficacy. Most historic clinical data are taken from a patient at rest (e.g., resting heart rate) with the assumption that only disease can shift homeostasis, and most wearable data are ambulatory (e.g., heart rate during a workout) with confounders such as physical activity making traditional clinical interpretation challenging. Interoperability is another crucial aspect that needs to be addressed when deciding on different sensing elements. This will help increase the generalizability of models by allowing them access to various kinds of data.

3.4. Data Preprocessing

As we have seen previously, data can come from various sources and in various forms. For the successful application of ML, these data must be harmonized and standardized. Data harmonization standards and intelligent interoperability techniques are the two classes along the knowledge stream. Another axis to classify data preprocessing techniques is the data fusion methods, which include feature level, data level, and decision level fusion. One more way to organize the data prior to analysis is through preprocessing techniques. These include filtering, feature extraction, and natural language processing techniques.

1. Data harmonization standards: These standards describe the preprocessing technique that prepares different kinds of data to become compatible with each other. It allows the AI to access a diversity of information through access to researcher and institution knowledge [102]. Some standards are specific to the medical cases they deal with [103–105]; however, there exists a set of medical means to ensure interoperability. The most common standards are Health Level 7 (HL7), openEHR, and ISO/IEEE 11073 Personal Health Data (PHD) standards [106], International Statistical Classification of Diseases version 10 (ICD-10) [107] and Current Procedural Terminology (CPT) codes [108]
2. Intelligent interoperability: Here, ML or other algorithms are used to combine the information from different data sources, and particularly EMRs. In intelligent interoperability of healthcare components, artificial intelligence or some other rule-based systems are used to automatically draw the relevant information from the EMRs or

sensor data. These systems use different algorithms to ensure the interoperability of various data sources. The following Table 5 elucidates such strategies. Although these systems allow for effective data communication while ensuring information integrity, one key issue is allowing for the encoding of categorical features so that the information is stored effectively.

Table 5. Comparison of interoperability techniques.

Name	Properties	References
Blockchain technology	Focused on patients rather than healthcare providers. Data are linked to the patient, aggregated, and then sensitive information such as allergies is published on the blockchain, ensuring privacy and data immutability.	[109,110]
Internet of Things	It employs the principles of the internet of things for data interoperability. It uses the protocols of Message Queuing Telemetry Transport (MQTT) to publish the relevant patient information.	[111]
Dynamic Semantic Web services	It uses the dynamic semantic web to convert the data into the HL7 framework.	[112]
Cloud Based Interoperability	It uses cloud-based models, for example, amazon web services, Microsoft Azure, and IBM Watson, to convert it into an openEHR or HL7 standard.	[113]
Knowledge Graphs	Knowledge graphs are used for the interoperability of biomedical data.	[37]

The methods used for interoperability include NLP, data mapping and transformation, data quality assessment, predictive analytics, and anomaly detection. They are used to promote one or more of these: Standardization of data, using application programming interface (API), using middleware and frameworks such as the Da Vinci project, and health information exchanges (HIEs). While effective, NLP techniques are very resource intensive. Data mapping and transformations can be very narrow in application. Data quality assessments can be used to compare inconsistencies but require constant updates and maintenance. Predictive analytics can help improve care coordination and resource allocation, but this is also effective in a narrow range of situations. Anomaly detection can identify unusual or unexpected patterns in healthcare data, potentially flagging issues that may need to be addressed, but can suffer from alert fatigue if the sensitivity is too high. However, it requires certain contextual information to be more effective.

3. Data Fusion: A physiological event can be observed with the help of various sensors, each sensing a unique aspect of the physiological event. The system has to fuse or combine information from different sensing elements for a holistic understanding of the event. This is done at multiple levels. In industry 4.0, healthcare systems, these sensing elements are spread across time and space (wearable sensors, ambulances, and hospitals). Fusing information from multiple sensors provides a more holistic picture of healthcare, including detection, phenotyping, disease progression, and other related data-powered solutions. Ref. [114] exhibits a combination of different layers of data fusion in connected healthcare, from individual sensors to detect medical events, to a network of connected devices, and finally, fusing information amongst various institutions. Ref. [115] displays a sensor fusion model between communication systems. Ref. [116] defines different levels of data fusion. These include signal level fusion, feature level fusion, and decision level fusion. Kalman Filtering is a popular statistics method for signal level fusion and is widely used in biomedical sensor networks. Weighted averages are also widely used to penalize sensors with more

noise in a sensor network [117–119]. Particle filtering, amongst various other variants, is also used extensively for signal level fusion in sensor networks in healthcare [120]. Ref. [121] uses temporal evidence theory for signal level fusion for activity recognition. Feature level fusion means each sensing element's features are calculated and fused. Ref. [122] calculates a linear combination of features to obtain a new feature. Ref. [123] is a weakly supervised program for feature-level fusion. Decision level fusion is a way to fuse decisions based on different information streams. There exist many such systems in the context of healthcare [124,125]. The critical issue in all these is developing a plastic nature of fusion techniques. A plastic fusion technique would be flexible to change with the emerging problem because different features or data may have other significance for each model.

There are several key limitations to data harmonization standards for electronic medical records for example:

- Complexity—Data harmonization standards can be complex and may require significant resources to implement and maintain.
- Limited adoption—Not all electronic medical record systems may adopt the same data harmonization standards, which can limit the ability to exchange data between systems.
- Changing standards—Data standards can change over time, which can make it difficult to maintain compatibility with other systems.
- Privacy and security concerns—The exchange of patient data between systems can raise concerns about privacy and security. Careful measures must be taken to ensure that patient data are protected when they are shared between systems.
- Cost—Implementing and maintaining data harmonization standards can be expensive, particularly for smaller healthcare organizations.
- Intended use—some coding is designed for a different reason than it is used for, e.g., reimbursement versus treatment.

3.5. Decision Systems

The nature of decision systems is specific to the problem they deal with. One axis along which the decision systems can be classified is the medical problems they solve, which include data quality, phenotyping, medication adherence, graph representation of data, detection of disease, and mortality prediction. One more axis along which the decision systems can be classified is the nature of algorithms, natural language processing, time series analysis, and graph neural networks.

1. Data Quality: The quality of the data acquired in healthcare is essential for the credibility of the predicted outcomes. Data quality issues are hard to identify in data with varying structures, shapes, dimensions, and sources. The dimensions of data quality, as elaborated by [126], are completeness (whether the relevant information is present), correctness (are the data correct), concordance (are they relatable to other data sources), plausibility (is any element in the EHRs making sense in the presence of other evidence), and currency (meaning how old are the data). These solutions will help to identify data quality issues, log them, encode them in metadata for datasets, help develop exclusion criteria of data based on its quality, and record the number of such problems. Ref. [127] is one such work that creates a framework to carry out all the tasks and uses probabilistic models to detect temporal stability and plausibility in biomedical data. It employs probabilistic change detection using Jensen–Shannon distance principles of statistical control of posterior beta distribution. Ref. [128] uses probability distribution distance to the same end. Ref. [129] is a measure of completeness by flagging incomplete data sources using the Delphi method. It also measures the same DQ dimension using patterns in the number of patients and compares them. Ref. [130] considers the data quality of radio frequency identification (RFID) in nine phases within healthcare systems.

2. Phenotypes: Phenotypes are the combination of an individual's observable disease traits. The data from the electronic health record are a set of data points related to interventions and the change in the states measured in lab tests. The data help align heterogeneous disease progression into temporal phenotypes. This allows data science techniques to find the relation between disease, symptoms, and interventions. These are also linked to mortality prediction, disease progression, and observation of medically complex phenotypes. Most temporal phenotype identification methods deploy clustering techniques. Phenotypes are also used to identify rare diseases [131–133]. These methods are rule-based [133] and graph-theory-based [134].

One of the critical challenges in AI-based phenotype is the representation of data. The data are being presented to domain experts, but developing a metric that identifies the visual tools' efficacy to represent the temporal phenotypes is worthwhile. For example, in encoding information in the edges and nodes of a graph, silhouette diagrams [135] are very different in richness and application compared to graphs. Graph theory is widely used in these systems as it is very suitable for the relational nature of phenotypes. One key concept is called category theory, which is a directed graph. The data used in temporal phenotypes are time-based (hence, directed), comprising nodes and morphisms. It is different from other graph representations as the morphisms encode the information of the mappings [136]. Very little focused work in this domain comes from EMRs.

3. Deidentification: De-identification of electronic medical records in an automatic manner is an active area of research where blockchain has recently been widely used [137,138]. Ref. [139] compares deep learning, rule-based systems, and shallow learning for de-identifying EMRs and argues that stacked learning is the most efficient ensemble technique. Ref. [140] deploys self-attention networks and stacked recurrent neural networks to de-identify the medical records. The main de-identification methods are neural networks, blockchain technology, and rule-based systems [140]. Some Internet of Medical Things (IoMT) schemes uses IoT protocols to preserve privacy while ensuring that critical information is relayed to the relevant stakeholder [141].

Challenges in this field remain the interplay of structured, unstructured, and semi-structured data. These data come from various sources and categories and, in the case of categorical features with other features, must be collated before solutions can be designed.

4. Adherence: Adherence to suggested and prescribed medical regimens is a crucial component of healthcare. Healthcare is an integrated process; hence, adherence is monitored by different sensing and AI techniques to ensure the efficacy of the interventions. The following Table 6 represents the various AI methods used to this end.

Table 6. Recent Works: AI in Adherence.

Name	Summary	Application	References
Conversational Robot	Chatbot used for drug adherence	Drug Adherence	[142,143]
Ethics	Deliberates over the ethical questions arising from the usage of AI in Norm Adherence	Ethics	[144]
Lifestyle Modification	It uses a web app to help monitor adherence, lifestyle modifications, for Example, in the case of cancer.	Drug Adherence	[145]
Medication Adherence	It uses machine learning to perform binary classification of the medication adherence for Parkinson's disease patients.	Remote Monitoring	[146]
Excercise Adherence	Uses machine learning models to estimate likelihood to adhere to a physical exercise regimen using accelerators and other data sources.	Predictive healthcare	[147]

Table 6. *Cont.*

Name	Summary	Application	References
Medication Adherence	Uses machine learning models to identify the likelihood of non-adherence to medication from electronic health records	Predictive healthcare	[148]
Medication Adherence	Uses data from wearable sensors to measure drug adherence for a specific cause.	Remote Monitoring	[149]
Medication Adherence	Uses cloud-based applications for medication adherence in home hospitalizations	Remote Monitoring	[150]

The key challenge in this domain is access to relevant data as the disease progresses. Here, the importance of different features coming from the same sensors and additional sensors can change as the condition changes its phase.

5. Diagnosis and mortality prediction: Disease prediction can help speed up the process of health care and increase the prediction accuracy, leading to the correct treatment being administered earlier. In the case of critical systems, the idea of mortality prediction and their interplay with demographic information and phenotype can help save lives. It can also help in understanding the progression of the disease and can direct healthcare resources in the right direction. Ref. [151] contains a process for disease prediction using electronic health records. It uses convolutional neural networks (CNN) to this end. Ref. [152] uses hybrid machine learning techniques to predict cardiovascular diseases. It uses a combination of random forest and linear classification models. Ref. [153] develops a naive Bayes analytic model for disease prediction using electronic health records.

Machine learning has been used to predict mortality for some time [154]. It has significant implications for different phenotypes [155,156]. These algorithms are used widely in brain injuries [157,158].

The critical challenges in disease and mortality prediction are the development of explainable machine learning models. As these models make predictions, they need to be explainable and validated as accurate for each patient prediction. Another crucial issue in this domain is the development of ethical frameworks to enable them to be deployed in the real world. Moral dilemmas such as those explained in [159] for self-driving cars should be identified.

3.6. Explainability

The literature uses many exciting techniques in time-series, EHR, and graph-based data. These techniques include feature significance and their interplay-based methods. Deep learning important features, or DeepLIFT, is widely used to this end [160,161] as it combines the importance of a feature as it passes through the layers of the neural network. local interpretable model agnostic explanation, or LIME, introduced in [162], is also widely used in electronic health records [163,164].

Attention mechanisms find the relevant neurons or the dataset components that are the most pertinent information needed for classification. DeepSOFA, DeepHINT, and Grad-CAM are such systems [165–167].

Least absolute shrinkage and selection operator (LASSO) is an explainability technique that uses dimensionality reduction techniques to explain the outcomes of a neural network. They are also used to describe healthcare outcomes [168].

Some explainability techniques draw the rules from the networks, and these systems are also applied in healthcare [169,170].

Deep Taylor decomposition is one such explainability technique used in these systems [171]. Shapley values are also used in such scenarios. The key challenges in developing these systems for graph neural networks are primarily encountered when this method is used for phenotyping.

There is currently a lack of standardized evaluation methods for interpretability techniques, making it difficult to compare and contrast the effectiveness of different approaches. Clinical relevance at present is limited to the identification of which traditional inputs are significant.

3.7. Levels of Automation

Levels of automation for the said topic are discussed as follows:

1. Human Only: Here, there is no AI involved, for example, the calculation of muscle atrophy using electromyogram (EMG) signals [172]. This process, however, involves the signal processing techniques for the representation of data.
2. Shadow Mode: In shadow mode, the data generated by the interaction of the medical practitioner and other sources are logged, and the data are labeled using the judgment of a qualified physician. These data are used to train a machine learning or an optimization algorithm. One such system developed by the ICL team is a reinforcement learning framework optimizing interventions retrospectively that allows a regulatory compliant pathway to clinical testing. This technique is used for sepsis treatment in the ICU [54].
3. AI Assistant: This level of decision making assistance provides the physician with suggestions. Some systems use these to detect cancers; for example, one such system uses biomedical images and structured data to detect hepatocellular carcinoma in the AI assistant model [173].
4. Partial Solutions: Based on the data, the AI comes up with a diagnosis independently, but needs a physician's input.
5. Full Automation: All the tasks in healthcare are provided by AI alone.

4. Conclusions

This review presents a paradigm of the application of AI in times-series and graph-based healthcare data that is driven by translational medicine. It looks at the complete pipeline, starting from data collection, harmonization, and quality dimensions. The decision systems are deliberated over, including various kinds of phenotyping, mortality detection, and other methods. We looked at the components related to the data, classifying them into multiple axes. Recent advances and state of the art technology in the various lexemes of the paradigms found were also reviewed.

Data can be classified along multiple axes, including structure, source, and dimension. Most healthcare data are unstructured, which has been used in conjunction with structured data to predict healthcare outcomes. Data preprocessing techniques can help combine different types of data, denoising and harmonizing to increase the reusability.

Another issue that needs to be tackled in data collection and preprocessing is interoperability of various devices and sensors, and this review has elaborated on different interoperability methods.

There are issues where collected data are fed to different decision systems. This part of the pipeline was discussed by this review, especially where the graph-based solutions, such as temporal phenotyping used to help identify risks for various morbidities and help cluster disease presentation into various groups, are concerned. The most recent works and reviewed literature focus more on applying these solutions in the real world. This application becomes easier when the AI process making predictions can be explained, hence, different explainability and interpretability techniques are compared here while highlighting the lack of standard metrics of evaluation for such methods. The validation of accuracy for each individual patient is an open area of research.

Based on the advances mentioned in this review, any future review may include the identification of ethical dilemmas in healthcare interventions and personalized healthcare: continuous healthcare monitoring and better intervention methods. Clinical use of ambulatory data continues to be a challenge for traditional medical practice. The debate between generalizable AI models with the required precision to achieve individual specific health outcomes will likely continue. This work can influence the current healthcare system in a positive manner; however, a way of combining these issues can first develop individual specific models, then explain them using explainability techniques, and then cluster them for general exploratory studies.

There are many challenges associated with healthcare data collection for the so-called disease X [174,175]. The more evolved diseases can be stopped from progressing in real-time using multi-omics-profiling and outlier detection [176]. Another challenge in dealing with data derived from time-based sensor data is the integration of advancements in real-time systems. To this end, translational medicine is already defining some solutions. Another major challenge is generating data for groups for which these data are unavailable using generative AI.

Author Contributions: Conceptualization, H.A., I.K.N., B.K.R., C.C. and D.W.; methodology, I.K.N., B.K.R. and S.M.; software, H.A., I.K.N., B.K.R. and S.M.; validation, I.K.N., B.K.R. and S.M.; formal analysis, H.A., I.K.N., B.K.R. and S.M.; investigation, H.A.; resources, I.K.N. and D.W.; data curation, H.A.; writing—original draft preparation, H.A.; writing—review and editing, H.A., I.K.N., B.K.R., C.C., S.M. and D.W.; visualization, H.A.; supervision, I.K.N., B.K.R., S.M. and D.W.; project administration, D.W.; funding acquisition, I.K.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the New Zealand College of Chiropractic student scholarship, funding number NZCC 20126384.

Institutional Review Board Statement: Ethical approval was not required for publication review.

Informed Consent Statement: Informed consent was not required for publication review.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hartl, D.; de Luca, V.; Kostikova, A.; Laramie, J.; Kennedy, S.; Ferrero, E.; Siegel, R.; Fink, M.; Ahmed, S.; Millholland, J.; et al. Translational precision medicine: An industry perspective. *J. Transl. Med.* **2021**, *19*, 245. [CrossRef] [PubMed]
2. Jordan, L. The problem with Big Data in Translational Medicine. A review of where we've been and the possibilities ahead. *Appl. Transl. Genom.* **2015**, *6*, 3–6. [CrossRef] [PubMed]
3. Ewusie, J.; Soobiah, C.; Blondal, E.; Beyene, J.; Thabane, L.; Hamid, J.S. Methods, Applications and Challenges in the Analysis of Interrupted Time Series Data: A Scoping Review. *J. Multidiscip. Health* **2020**, *13*, 411–423. [CrossRef]
4. Ahmad, M.A.; Eckert, C.; Teredesai, A. Interpretable machine learning in healthcare. In Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, Washington, DC, USA, 29 August–1 September 2018.
5. Baum, E.B. On the capabilities of multilayer perceptrons. *J. Complex.* **1988**, *4*, 193–215. [CrossRef]
6. Paganelli, A.I.; Mondéjar, A.G.; da Silva, A.C.; Silva-Calpa, G.; Teixeira, M.F.; Carvalho, F.; Raposo, A.; Endler, M. Real-time data analysis in health monitoring systems: A comprehensive systematic literature review. *J. Biomed. Inform.* **2022**, *127*, 104009. [CrossRef] [PubMed]
7. Chen, J.X. The Evolution of Computing: AlphaGo. *Comput. Sci. Eng.* **2016**, *18*, 4–7. [CrossRef]
8. Singh, S.P.; Wang, L.; Gupta, S.; Goli, H.; Padmanabhan, P.; Gulyás, B. 3D Deep Learning on Medical Images: A Review. *Sensors* **2020**, *20*, 5097. [CrossRef]
9. Taghanaki, S.A.; Abhishek, K.; Cohen, J.P.; Cohen-Adad, J.; Hamarneh, G. Deep semantic segmentation of natural and medical images: A review. *Artif. Intell. Rev.* **2020**, *54*, 137–178. [CrossRef]
10. Kumar, M.; Mishra, S.K. A comprehensive review on nature inspired neural network based adaptive filter for eliminating noise in medical images. *Curr. Med. Imaging* **2020**, *16*, 278. [CrossRef] [PubMed]
11. Pavlič, J.; Tomažič, T.; Kožuh, I. The impact of emerging technology influences product placement effectiveness: A scoping study from interactive marketing perspective. *J. Res. Interact. Mark.* **2021**, *16*, 551–568. [CrossRef]
12. Davy, C.; Bleasel, J.; Liu, H.; Tchan, M.; Ponniah, S.; Brown, A. Effectiveness of chronic care models: Opportunities for improving healthcare practice and health outcomes: A systematic review. *BMC Health Serv. Res.* **2015**, *15*, 1–11. [CrossRef] [PubMed]
13. Chen, I.Y.; Joshi, S.; Ghassemi, M.; Ranganath, R. Probabilistic Machine Learning for Healthcare. *Annu. Rev. Biomed. Data Sci.* **2021**, *4*, 393–415. [CrossRef]

14. Wang, F.; Cui, P.; Pei, J.; Song, Y.; Zang, C. Recent Advances on Graph Analytics and Its Applications in Healthcare. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Virtual Event, CA, USA, 6–10 July 2020.
15. Gu, D.; Li, T.; Wang, X.; Yang, X.; Yu, Z. Visualizing the intellectual structure and evolution of electronic health and telemedicine research. *Int. J. Med. Inform.* **2019**, *130*, 103947. [CrossRef] [PubMed]
16. Tran, B.X.; Nghiem, S.; Sahin, O.; Vu, T.M.; Ha, G.H.; Vu, G.T.; Pham, H.Q.; Do, H.T.; Latkin, C.; Tam, W.; et al. Modeling Research Topics for Artificial Intelligence Applications in Medicine: Latent Dirichlet Allocation Application Study. *J. Med. Internet Res.* **2019**, *21*, e15511. [CrossRef]
17. Adnan, K.; Akbar, R.; Khor, S.W.; Ali, A.B.A. Role and Challenges of Unstructured Big Data in Healthcare. In *Data Management, Analytics and Innovation*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 301–323.
18. Krishna, S.; Boren, S.; Balas, E.A. Healthcare via Cell Phones: A Systematic Review. *Telemed. e-Health* **2009**, *15*, 231–240. [CrossRef]
19. Bellamy, D.; Celi, L.; Beam, A. Evaluating progress on machine learning for longitudinal electronic healthcare data. *arXiv* **2020**, arXiv:2010.01149.
20. McCarthy, J.J. Optimal paradigms. In *Paradigms in Phonological Theory*; Linguistics Department Faculty Publication: Amherst, MA, USA, 2005; p. 55.
21. Adithan, C. Principles of translational science in medicine: From bench to bedside. *Indian J. Med. Res.* **2017**, *145*, 408–409. [CrossRef]
22. Liu, B.; Li, J.; Chen, C.; Tan, W.; Chen, Q.; Zhou, M. Efficient Motif Discovery for Large-Scale Time Series in Healthcare. *IEEE Trans. Ind. Informatics* **2015**, *11*, 583–590. [CrossRef]
23. Balasubramanian, A.; Wang, J.; Prabhakaran, B. Discovering multidimensional motifs in physiological signals for personalized healthcare. *IEEE J. Sel. Top. Signal Process.* **2016**, *10*, 832–841. [CrossRef]
24. Pereira, J.; Silveira, M. Learning representations from healthcare time series data for unsupervised anomaly detection. In Proceedings of the 2019 IEEE International Conference on Big Data and Smart Computing (BigComp), Kyoto, Japan, 27 February–02 March 2019.
25. Kaushik, S.; Choudhury, A.; Sheron, P.K.; Dasgupta, N.; Natarajan, S.; Pickett, L.A.; Dutt, V. AI in healthcare: Time-series forecasting using statistical, neural, and ensemble architectures. *Front. Big Data* **2020**, *3*, 4. [CrossRef] [PubMed]
26. Maweu, B.M.; Shamsuddin, R.; Dakshit, S.; Prabhakaran, B. Generating Healthcare Time Series Data for Improving Diagnostic Accuracy of Deep Neural Networks. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 2508715. [CrossRef]
27. Harutyunyan, H.; Khachatrian, H.; Kale, D.C.; Ver Steeg, G.; Galstyan, A. Multitask learning and benchmarking with clinical time series data. *Sci. Data* **2019**, *6*, 96. [CrossRef] [PubMed]
28. Lipton, Z.C.; Kale, D.; Wetzel, R. Directly modeling missing data in sequences with rnns: Improved classification of clinical time series. In Proceedings of the 1st Machine Learning for Healthcare Conference, Los Angeles, CA, USA, 9–20 August 2016.
29. Whitney, V.K.M. Relational data management implementation techniques. In Proceedings of the ACM SIGFIDET (now SIGMOD) Workshop on Data description, Access and Control, Ann Arbor, MI, USA, 1–5 May 1974.
30. Zhang, Y.; Sheng, M.; Zhou, R.; Wang, Y.; Han, G.; Zhang, H.; Xing, C.; Dong, J. Hkgb: An inclusive, extensible, intelligent, semi-auto-constructed knowledge graph framework for healthcare with clinicians' expertise incorporated. *Inf. Process. Manag.* **2020**, *57*, 102324. [CrossRef]
31. Choi, E.; Bahadori, M.T.; Song, L.; Stewart, W.F.; Sun, J. GRAM: Graph-based attention model for healthcare representation learning. In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, Halifax, NS, Canada, 13–17 August 2017.
32. Liu, C.; Wang, F.; Hu, J.; Xiong, H. Temporal phenotyping from longitudinal electronic health records: A graph based framework. In Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining, Sydney, NSW, Australia, 10–13 August 2015.
33. Sood, S.K.; Mahajan, I. A Fog-Based Healthcare Framework for Chikungunya. *IEEE Internet Things J.* **2017**, *5*, 794–801. [CrossRef]
34. Wu, T.; Wang, Y.; Wang, Y.; Zhao, E.; Yuan, Y. Leveraging graph-based hierarchical medical entity embedding for healthcare applications. *Sci. Rep.* **2021**, *11*, 5858. [CrossRef] [PubMed]
35. Winter, A.; Brigl, B.; Wendt, T. Modeling hospital information systems (part 1): The revised three-layer graph-based meta model 3LGM2. *Methods Inf. Med.* **2003**, *42*, 544–551. [PubMed]
36. Sharma, N.; Bhatt, R. Privacy Preservation in WSN for Healthcare Application. *Procedia Comput. Sci.* **2018**, *132*, 1243–1252. [CrossRef]
37. Malik, K.M.; Krishnamurthy, M.; Alobaidi, M.; Hussain, M.; Alam, F.; Malik, G. Automated domain-specific healthcare knowledge graph curation framework: Subarachnoid hemorrhage as phenotype. *Expert Syst. Appl.* **2019**, *145*, 113120. [CrossRef]
38. Kalamaras, I.; Glykos, K.; Megalooikonomou, V.; Votis, K.; Tzovaras, D. Graph-based visualization of sensitive medical data. *Multimed. Tools Appl.* **2021**, *81*, 209–236. [CrossRef]
39. Zhang, Q.; Zhou, J.; Zhang, B. Graph Based Multichannel Feature Fusion for Wrist Pulse Diagnosis. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 3732–3743. [CrossRef] [PubMed]
40. Finlayson, S.G.; LePendu, P.; Shah, N.H. Building the graph of medicine from millions of clinical narratives. *Sci. Data* **2014**, *1*, 140032. [CrossRef]

41. Yue, X.; Wang, Z.; Huang, J.; Parthasarathy, S.; Moosavinasab, S.; Huang, Y.; Lin, S.M.; Zhang, W.; Zhang, P.; Sun, H. Graph embedding on biomedical networks: Methods, applications and evaluations. *Bioinformatics* **2019**, *36*, 1241–1251. [CrossRef] [PubMed]
42. Zhang, W.; Chen, Y.; Li, D.; Yue, X. Manifold regularized matrix factorization for drug-drug interaction prediction. *J. Biomed. Inform.* **2018**, *88*, 90–97. [CrossRef]
43. Kulmanov, M.; Khan, M.A.; Hoehndorf, R. DeepGO: Predicting protein functions from sequence and interactions using a deep ontology-aware classifier. *Bioinformatics* **2017**, *34*, 660–668. [CrossRef]
44. Chen, L.; Li, X.; Sheng, Q.Z.; Peng, W.-C.; Bennett, J.; Hu, H.-Y.; Huang, N. Mining Health Examination Records—A Graph-Based Approach. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 2423–2437. [CrossRef]
45. Kaur, K.; Rani, R. Managing Data in Healthcare Information Systems: Many Models, One Solution. *Computer* **2015**, *48*, 52–59. [CrossRef]
46. Thews, O.; Rohrbach, C.; Sergl, M.; Pommerening, K.; Müller, R. A Graph-Grammar Approach to Represent Causal, Temporal and Other Contexts in an Oncological Patient Record. *Methods Inf. Med.* **1996**, *35*, 127–141. [CrossRef]
47. Liu, Y.; Song, Z.; Xu, X.; Rafique, W.; Zhang, X.; Shen, J.; Khosravi, M.R.; Qi, L. Bidirectional GRU networks-based next POI category prediction for healthcare. *Int. J. Intell. Syst.* **2021**, *37*, 4020–4040. [CrossRef]
48. Rodeheaver, N.; Kim, H.; Herbert, R.; Seo, H.; Yeo, W.H. Breathable, Wireless, Thin-Film Wearable Biopatch Using Noise-Reduction Mechanisms. *ACS Appl. Electron. Mater.* **2022**, *4*, 503–512. [CrossRef]
49. Yang, J.; Soltan, A.A.S.; Clifton, D.A. Machine learning generalizability across healthcare settings: Insights from multi-site COVID-19 screening. *NPJ Digit. Med.* **2022**, *5*, 69. [CrossRef] [PubMed]
50. Palanisamy, V.; Thirunavukarasu, R. Implications of big data analytics in developing healthcare frameworks—A review. *J. King Saud Univ. Comput. Inf. Sci.* **2019**, *31*, 415–425. [CrossRef]
51. Moore, W.; Frye, S. Review of HIPAA, part 1: History, protected health information, and privacy and security rules. *J. Nucl. Med. Technol.* **2019**, *47*, 269–272. [CrossRef]
52. Zhang, Z.; Yan, C.; Lasko, T.A.; Sun, J.; Malin, B. SynTEG: A framework for temporal structured electronic health data simulation. *J. Am. Med. Inform. Assoc.* **2020**, *28*, 596–604. [CrossRef]
53. Abedi, M.; Hempel, L.; Sadeghi, S.; Kirsten, T. GAN-Based Approaches for Generating Structured Data in the Medical Domain. *Appl. Sci.* **2022**, *12*, 7075. [CrossRef]
54. Li, L.; Albert-Smet, I.; Faisal, A. Optimizing medical treatment for sepsis in intensive care: From reinforcement learning to pre-trial evaluation. *arXiv* **2020**, arXiv:2003.06474.
55. Klompas, M.; Kulldorff, M.; Vilk, Y.; Bialek, S.R.; Harpaz, R. Herpes Zoster and Postherpetic Neuralgia Surveillance Using Structured Electronic Data. *Mayo Clin. Proc.* **2011**, *86*, 1146–1153. [CrossRef] [PubMed]
56. Aminifar, A.; Lamo, Y.; Pun, K.I.; Rabbi, F. *A Practical Methodology for Anonymization of Structured Health Data*; Linköping University Electronic Press: Linköping, Sweden, 2019.
57. Kanwal, T.; Anjum, A.; Khan, A. Privacy preservation in e-health cloud: Taxonomy, privacy requirements, feasibility analysis, and opportunities. *Clust. Comput.* **2020**, *24*, 293–317. [CrossRef]
58. Xu, R.; Baracaldo, N.; Zhou, Y.; Anwar, A.; Ludwig, H. Hybridalpha: An efficient approach for privacy-preserving federated learning. In Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security, London, UK, 15 November 2019.
59. Wibawa, F.; Catak, F.O.; Kuzlu, M.; Sarp, S.; Cali, U. Homomorphic Encryption and Federated Learning based Privacy-Preserving CNN Training: COVID-19 Detection Use-Case. In Proceedings of the 2022 European Interdisciplinary Cybersecurity Conference, Barcelona, Spain, 15–16 June 2022.
60. Wright, A.; McEvoy, D.S.; Aaron, S.; McCoy, A.; Amato, M.G.; Kim, H.; Ai, A.; Cimino, J.J.; Desai, B.R.; El-Kareh, R.; et al. Structured override reasons for drug-drug interaction alerts in electronic health records. *J. Am. Med. Inform. Assoc.* **2019**, *26*, 934–942. [CrossRef] [PubMed]
61. Vest, J.R.; Grannis, S.J.; Haut, D.P.; Halverson, P.K.; Menachemi, N. Using structured and unstructured data to identify patients' need for services that address the social determinants of health. *Int. J. Med. Inform.* **2017**, *107*, 101–106. [CrossRef] [PubMed]
62. Wang, W.; Kiik, M.; Peek, N.; Curcin, V.; Marshall, I.J.; Rudd, A.G.; Wang, Y.; Douiri, A.; Wolfe, C.D.; Bray, B. A systematic review of machine learning models for predicting outcomes of stroke with structured data. *PLoS ONE* **2020**, *15*, e0234722.
63. Richter, A.N.; Khoshgoftaar, T.M. A review of statistical and machine learning methods for modeling cancer risk using structured clinical data. *Artif. Intell. Med.* **2018**, *90*, 1–14. [CrossRef] [PubMed]
64. Therneau, T.M.; Grambsch, P.M. *Modeling Survival Data: Extending the Cox Model*; Springer: New York, NY, USA, 2000; pp. 39–77.
65. Gopinath, D.; Agrawal, M.; Murray, L.; Horng, S.; Karger, D.; Sontag, D. Fast, Structured Clinical Documentation via Contextual Autocomplete. In Proceedings of the 5th Machine Learning for Healthcare Conference, Virtual, 7–8 August 2020.
66. Shao, Y.; Zeng, Q.T.; Chen, K.K.; Shutes-David, A.; Thielke, S.M.; Tsuang, D.W. Detection of probable dementia cases in undiagnosed patients using structured and unstructured electronic health records. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 128. [CrossRef] [PubMed]
67. Banda, J.M.; Seneviratne, M.; Hernandez-Boussard, T.; Shah, N.H. Advances in Electronic Phenotyping: From Rule-Based Definitions to Machine Learning Models. *Annu. Rev. Biomed. Data Sci.* **2018**, *1*, 53–68. [CrossRef] [PubMed]
68. Sung, S.-F.; Lin, C.-Y.; Hu, Y.-H. EMR-Based Phenotyping of Ischemic Stroke Using Supervised Machine Learning and Text Mining Techniques. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 2922–2931. [CrossRef] [PubMed]

69. Boxwala, A.; Kim, J.; Grillo, J.M.; Ohno-Machado, L. Using statistical and machine learning to help institutions detect suspicious access to electronic health records. *J. Am. Med. Inform. Assoc.* **2011**, *18*, 498–505. [CrossRef] [PubMed]
70. Lantz, E. Machine Learning for Risk Prediction and Privacy in Electronic Health Records. Ph.D. Thesis, The University of Wisconsin-Madison, Madison, WI, USA, 2016.
71. Kim, S.; Lee, H.; Chung, Y.D. Privacy-preserving data cube for electronic medical records: An experimental evaluation. *Int. J. Med. Inform.* **2017**, *97*, 33–42. [CrossRef] [PubMed]
72. Marble, H.D.; Huang, R.; Dudgeon, S.N.; Lowe, A.; Herrmann, M.D.; Blakely, S.; Leavitt, M.O.; Isaacs, M.; Hanna, M.G.; Sharma, A.; et al. A Regulatory Science Initiative to Harmonize and Standardize Digital Pathology and Machine Learning Processes to Speed up Clinical Innovation to Patients. *J. Pathol. Inform.* **2020**, *11*, 22. [CrossRef] [PubMed]
73. Guan, J.; Li, R.; Yu, S.; Zhang, X. A method for generating synthetic electronic medical record text. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2019**, *23*, 173–182. [CrossRef]
74. Chin-Cheong, K.; Sutter, T.M.; Vogt, J.E. Generation of heterogeneous synthetic electronic health records using GANs. In Proceedings of the Workshop on Machine Learning For Health (ML4H) at the 33rd Conference on Neural Information Processing Systems (NeurIPS), Vancouver, BC, Canada, 8–14 December 2019.
75. Walonoski, J.A.; Kramer, M.; Nichols, J.; Quina, A.; Moesel, C.; Hall, D.; Duffett, C.; Dube, K.; Gallagher, T.; McLachlan, S. Synthea: An approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record. *J. Am. Med. Inform. Assoc.* **2017**, *25*, 230–238. [CrossRef]
76. Chen, I.Y.; Pierson, E.; Rose, S.; Joshi, S.; Ferryman, K.; Ghassemi, M. Ethical Machine Learning in Healthcare. *Annu. Rev. Biomed. Data Sci.* **2020**, *4*, 123–144. [CrossRef] [PubMed]
77. Aggarwal, A.; Garhwal, S.; Kumar, A. HEDEA: A Python Tool for Extracting and Analysing Semi-structured Information from Medical Records. *Health Inform. Res.* **2018**, *24*, 148–153. [CrossRef] [PubMed]
78. Makarova, E.; Lagerev, D. Methodology for Preprocessing Semi-Structured Data for Making Managerial Decisions in the Healthcare. In Proceedings of the InCEUR Workshop Proceedings of the 30th International Conference on Computer Graphics and Machine Vision, Saint Petersburg, Russia, 22–25 September 2020.
79. Hong, N.; Wen, A.; Stone, D.J.; Tsuji, S.; Kingsbury, P.R.; Rasmussen, L.V.; Pacheco, J.A.; Adekkanattu, P.; Wang, F.; Luo, Y.; et al. Developing a FHIR-based EHR phenotyping framework: A case study for identification of patients with obesity and multiple comorbidities from discharge summaries. *J. Biomed. Inform.* **2019**, *99*, 103310. [CrossRef]
80. Batra, S.; Sachdeva, S. Organizing standardized electronic healthcare records data for mining. *Health Policy Technol.* **2016**, *5*, 226–242. [CrossRef]
81. Yuan, J.; Holtz, C.; Smith, T.H.; Luo, J. Autism spectrum disorder detection from semi-structured and unstructured medical data. *EURASIP J. Bioinform. Syst. Biol.* **2016**, *2017*, 3. [CrossRef]
82. Miled, Z.B.; Haas, K.; Black, C.M.; Khandker, R.K.; Chandrasekaran, V.; Lipton, R.; Boustani, M.A. Predicting dementia with routine care EMR data. *Artif. Intell. Med.* **2020**, *102*, 101771. [CrossRef]
83. Geraci, J.; Wilansky, P.; de Luca, V.; Roy, A.; Kennedy, J.L.; Strauss, J. Applying deep neural networks to unstructured text notes in electronic medical records for phenotyping youth depression. *Evid.-Based Ment. Health* **2017**, *20*, 83–87. [CrossRef] [PubMed]
84. Goh, K.H.; Wang, L.; Yeow, A.Y.K.; Poh, H.; Li, K.; Yeow, J.J.L.; Tan, G.Y.H. Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. *Nat. Commun.* **2021**, *12*, 1–10. [CrossRef] [PubMed]
85. Zuo, Z.; Li, J.; Xu, H.; Al Moubayed, N. Curvature-based feature selection with application in classifying electronic health records. *Technol. Forecast. Soc. Chang.* **2021**, *173*, 12112. [CrossRef]
86. Xu, Z.; Chou, J.; Zhang, X.S.; Luo, Y.; Isakova, T.; Adekkanattu, P.; Ancker, J.S.; Jiang, G.; Kiefer, R.C.; Pacheco, J.A.; et al. Identifying sub-phenotypes of acute kidney injury using structured and unstructured electronic health record data with memory networks. *J. Biomed. Inform.* **2020**, *102*, 103361. [CrossRef] [PubMed]
87. Zhao, M.; Havrilla, J.; Peng, J.; Drye, M.; Fecher, M.; Guthrie, W.; Tunc, B.; Schultz, R.; Wang, K.; Zhou, Y. Development of a phenotype ontology for autism spectrum disorder by natural language processing on electronic health records. *J. Neurodev. Disord.* **2022**, *14*, 32. [CrossRef] [PubMed]
88. Song, B.; Feng, Y.; Li, X.; Sun, Z.; Yang, Y. Un-apriori: A novel association rule mining algorithm for unstructured EMRs. In Proceedings of the IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom), Dalian, China, 12–15 October 2017.
89. Kim, J.-C.; Chung, K. Associative Feature Information Extraction Using Text Mining from Health Big Data. *Wirel. Pers. Commun.* **2018**, *105*, 691–707. [CrossRef]
90. Boustani, M.; Perkins, A.J.; Khandker, R.K.; Duong, S.; Dexter, P.R.; Lipton, R.; Black, C.M.; Chandrasekaran, V.; Solid, C.A.; Monahan, P. Passive digital signature for early identification of Alzheimer's disease and related dementia. *J. Am. Geriatr. Soc.* **2020**, *68*, 511–518. [CrossRef]
91. Chung, K.; Yoo, H.; Choe, D.-E. Ambient context-based modeling for health risk assessment using deep neural network. *J. Ambient. Intell. Humaniz. Comput.* **2018**, *11*, 1387–1395. [CrossRef]
92. Ling, A.Y.; Kurian, A.W.; Caswell-Jin, J.; Sledge, G.W.; Shah, N.H.; Tamang, S.R. Using natural language processing to construct a metastatic breast cancer cohort from linked cancer registry and electronic medical records data. *JAMIA Open* **2019**, *2*, 528–553. [CrossRef] [PubMed]

93. Wallace, B.C.; Kuiper, J.; Sharma, A.; Zhu, M.B.; Marshall, I.J. Extracting PICO Sentences from Clinical Trial Reports using Supervised Distant Supervision. *J. Mach. Learn. Res.* **2016**, *17*, 4572–4596.
94. Lin, F.P.-Y.; Pokorny, A.; Teng, C.; Epstein, R.J. TEPAPA: A novel in silico feature learning pipeline for mining prognostic and associative factors from text-based electronic medical records. *Sci. Rep.* **2017**, *7*, 6918. [CrossRef] [PubMed]
95. Bjarnadottir, R.I.; Lucero, R.J. What Can We Learn about Fall Risk Factors from EHR Nursing Notes? A Text Mining Study. *eGEMs* **2018**, *6*, 21. [CrossRef]
96. Wang, L.; Lou, Z.; Jiang, K.; Shen, G. Bio-Multifunctional Smart Wearable Sensors for Medical Devices. *Adv. Intell. Syst.* **2019**, *1*, 1900040. [CrossRef]
97. Sempionatto, J.R.; Montiel, V.R.-V.; Vargas, E.; Teymourian, H.; Wang, J. Wearable and Mobile Sensors for Personalized Nutrition. *ACS Sens.* **2021**, *6*, 1745–1760. [CrossRef]
98. Chai, P.R.; Goodman, G.; Bustamante, M.; Mendez, L.; Mohamed, Y.; Mayer, K.H.; Boyer, E.W.; Rosen, R.K.; O'Cleirigh, C. Design and Delivery of Real-Time Adherence Data to Men Who Have Sex with Men Using Antiretroviral Pre-exposure Prophylaxis via an Ingestible Electronic Sensor. *AIDS Behav.* **2020**, *25*, 1661–1674. [CrossRef] [PubMed]
99. Weitschies, W.; Müller, L.; Grimm, M.; Koziolek, M. Ingestible devices for studying the gastrointestinal physiology and their application in oral biopharmaceutics. *Adv. Drug Deliv. Rev.* **2021**, *176*, 113853. [CrossRef]
100. Li, G.; Lian, W.; Qu, H.; Li, Z.; Zhou, Q.; Tian, J. Improving patient care through the development of a 5G-powered smart hospital. *Nat. Med.* **2021**, *27*, 936–937. [CrossRef]
101. Muhammad, G.; Alshehri, F.; Karray, F.; El Saddik, A.; Alsulaiman, M.; Falk, T.H. A comprehensive survey on multimodal medical signals fusion for smart healthcare systems. *Inf. Fusion* **2021**, *76*, 355–375.
102. Lucas, C.; Wong, P.; Klein, J.; Castro, T.B.R.; Silva, J.; Sundaram, M.; Ellingson, M.K.; Mao, T.; Oh, J.E.; Israelow, B.; et al. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **2020**, *584*, 463–469. [CrossRef] [PubMed]
103. Batra, G.; Aktaa, S.; Wallentin, L.; Maggioni, A.P.; Wilkinson, C.; Casadei, B.; Gale, C.P. Methodology for the development of international clinical data standards for common cardiovascular conditions: European Unified Registries for Heart Care Evaluation and Randomised Trials (EuroHeart). *Eur. Heart J. Qual. Care Clin. Outcomes* **2021**, *2021*, qcab052. [CrossRef] [PubMed]
104. Baxter, S.L.; Lee, A.Y. Gaps in standards for integrating artificial intelligence technologies into ophthalmic practice. *Curr. Opin. Ophthalmol.* **2021**, *32*, 431–438. [CrossRef]
105. American Diabetes Association. Diabetes technology: Standards of medical care in diabetes—2021. *Diabetes Care* **2021**, *44*, S85–S99. [CrossRef] [PubMed]
106. Laleci, G.B.; Dogac, A. A Semantically Enriched Clinical Guideline Model Enabling Deployment in Heterogeneous Healthcare Environments. *IEEE Trans. Inf. Technol. Biomed.* **2009**, *13*, 263–273. [CrossRef]
107. Cartwright, D.J. ICD-9-CM to ICD-10-CM Codes: What? Why? How? *Adv. Wound Care* **2013**, *2*, 588–592. [CrossRef]
108. Dotson, P. *CPT® Codes: What Are They, Why Are They Necessary, and How Are They Developed?* Mary Ann Liebert: New Rochelle, NY, USA, 2013.
109. Gordon, W.J.; Catalini, C. Blockchain Technology for Healthcare: Facilitating the Transition to Patient-Driven Interoperability. *Comput. Struct. Biotechnol. J.* **2018**, *16*, 224–230. [CrossRef] [PubMed]
110. Jabbar, R.; Fetais, N.; Krichen, M.; Barkaoui, K. Blockchain technology for healthcare: Enhancing shared electronic health record interoperability and integrity. In Proceedings of the IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), Doha, Qatar, 2–5 February 2020; pp. 310–317.
111. Pathak, N.; Misra, S.; Mukherjee, A.; Kumar, N. HeDI: Healthcare Device Interoperability for IoT-Based e-Health Platforms. *IEEE Internet Things J.* **2021**, *8*, 16845–16852. [CrossRef]
112. Balakrishna, S.; Thirumaran, M. Semantic Interoperability in IoT and Big Data for Health Care: A Collaborative Approach. In *Handbook of Data Science Approaches for Biomedical Engineering*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 185–220.
113. Joshi, R.; Negi, S.; Sachdeva, S. *Cloud Based Interoperability in Healthcare, in Computational Methods and Data Engineering*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 599–611.
114. Dautov, R.; Distefano, S.; Buyya, R. Hierarchical data fusion for Smart Healthcare. *J. Big Data* **2019**, *6*, 19. [CrossRef]
115. Hall, D.L.; Llinas, J. An introduction to multisensor data fusion. *Proc. IEEE* **1997**, *85*, 6–23. [CrossRef]
116. Lee, H.; Park, K.; Lee, B.; Choi, J.; Elmasri, R. Issues in data fusion for healthcare monitoring. In Proceedings of the 1st International Conference on PErvasive Technologies Related to Assistive Environments, Athens, Greece, 16–18 July 2008.
117. Djenouri, D.; Balasingham, I. New QoS and geographical routing in wireless biomedical sensor networks. In Proceedings of the Sixth International Conference on Broadband Communications, Networks, and Systems, Madrid, Spain, 14–16 September 2009.
118. Choi, S.; Han, S.I.; Jung, D.; Hwang, H.J.; Lim, C.; Bae, S.; Park, O.K.; Tschabrunn, C.M.; Lee, M.; Bae, S.Y.; et al. Highly conductive, stretchable and biocompatible Ag–Au core–sheath nanowire composite for wearable and implantable bioelectronics. *Nat. Nanotechnol.* **2018**, *13*, 1048–1056. [CrossRef] [PubMed]
119. Nathan, V.; Jafari, R. Particle Filtering and Sensor Fusion for Robust Heart Rate Monitoring Using Wearable Sensors. *IEEE J. Biomed. Health Inform.* **2017**, *22*, 1834–1846. [CrossRef]
120. Brady, K.; Gwon, Y.; Khorrami, P.; Godoy, E.; Campbell, W.; Dagli, C.; Huang, T.S. Multi-modal audio, video and physiological sensor learning for continuous emotion prediction. In Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge, Amsterdam, The Netherlands, 16 October 2016.

121. McKeever, S.; Ye, J.; Coyle, L.; Bleakley, C.; Dobson, S. Activity recognition using temporal evidence theory. *J. Ambient. Intell. Smart Environ.* **2010**, *2*, 253–269. [CrossRef]
122. Cai, H.; Qu, Z.; Li, Z.; Zhang, Y.; Hu, X.; Hu, B. Feature-level fusion approaches based on multimodal EEG data for depression recognition. *Inf. Fusion* **2020**, *59*, 127–138. [CrossRef]
123. Miao, F.; Liu, Z.-D.; Liu, J.-K.; Wen, B.; He, Q.-Y.; Li, Y. Multi-Sensor Fusion Approach for Cuff-Less Blood Pressure Measurement. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 79–91. [CrossRef] [PubMed]
124. Hossain, M.S.; Muhammad, G. Emotion-Aware Connected Healthcare Big Data Towards 5G. *IEEE Internet Things J.* **2017**, *5*, 2399–2406. [CrossRef]
125. Chen, C.; Jafari, R.; Kehtarnavaz, N. A Real-Time Human Action Recognition System Using Depth and Inertial Sensor Fusion. *IEEE Sens. J.* **2015**, *16*, 773–781. [CrossRef]
126. Weiskopf, N.G.; Weng, C. Methods and dimensions of electronic health record data quality assessment: Enabling reuse for clinical research. *J. Am. Med. Inform. Assoc.* **2013**, *20*, 144–151. [CrossRef]
127. Fox, F.; Aggarwal, V.R.; Whelton, H.; Johnson, O. A data quality framework for process mining of electronic health record data. In Proceedings of the 2018 IEEE International Conference on Healthcare Informatics (ICHI), New York, NY, USA, 4–7 June 2018.
128. Sáez, C.; Rodrigues, P.P.; Gama, J.; Robles, M.; Garcia-Gomez, J.M. Probabilistic change detection and visualization methods for the assessment of temporal stability in biomedical data quality. *Data Min. Knowl. Discov.* **2014**, *29*, 950–975. [CrossRef]
129. Puttkammer, N.; Baseman, J.; Devine, E.; Valles, J.; Hyppolite, N.; Garilus, F.; Honoré, J.; Matheson, A.; Zeliadt, S.; Yuhas, K.; et al. An assessment of data quality in a multi-site electronic medical record system in Haiti. *Int. J. Med. Inform.* **2016**, *86*, 104–116. [CrossRef] [PubMed]
130. Taggart, J.; Liaw, S.-T.; Yu, H. Structured data quality reports to improve EHR data quality. *Int. J. Med. Inform.* **2015**, *84*, 1094–1098. [CrossRef] [PubMed]
131. Li, Q.; Zhao, K.; Bustamante, C.D.; Ma, X.; Wong, W.H. Xrare: A machine learning method jointly modeling phenotypes and genetic evidence for rare disease diagnosis. *Anesthesia Analg.* **2019**, *21*, 2126–2134. [CrossRef]
132. Jia, J.; Wang, R.; An, Z.; Guo, Y.; Ni, X.; Shi, T. RDAD: A Machine Learning System to Support Phenotype-Based Rare Disease Diagnosis. *Front. Genet.* **2018**, *9*, 587. [CrossRef] [PubMed]
133. Morley, K.I.; Wallace, J.; Denaxas, S.C.; Hunter, R.J.; Patel, R.S.; Perel, P.; Shah, A.D.; Timmis, A.D.; Schilling, R.J.; Hemingway, H. Defining Disease Phenotypes Using National Linked Electronic Health Records: A Case Study of Atrial Fibrillation. *PLoS ONE* **2014**, *9*, e110900. [CrossRef] [PubMed]
134. Ash, J.A.; Rapp, P.R. A quantitative neural network approach to understanding aging phenotypes. *Ageing Res. Rev.* **2014**, *15*, 44–50. [CrossRef]
135. Lee, C.; Rashbass, J.; van der Schaar, M. Outcome-Oriented Deep Temporal Phenotyping of Disease Progression. *IEEE Trans. Biomed. Eng.* **2020**, *68*, 2423–2434. [CrossRef]
136. Tuyéras, R. Category theory for genetics II: Genotype, phenotype and haplotype. *arXiv* **2018**, arXiv:1805.07004.
137. Mayer, A.H.; da Costa, C.; Righi, R. Electronic health records in a blockchain: A systematic review. *Health Inform. J.* **2020**, *26*, 1273–1288. [CrossRef]
138. Shi, S.; He, D.; Li, L.; Kumar, N.; Khan, M.K.; Choo, K.-K.R. Applications of blockchain in ensuring the security and privacy of electronic health record systems: A survey. *Comput. Secur.* **2020**, *97*, 101966. [CrossRef]
139. Kim, Y.; Heider, P.; Meystre, S. Ensemble-based Methods to Improve De-identification of Electronic Health Record Narratives. *AMIA Annu. Symp. Proceedings AMIA Symp.* **2018**, *2018*, 663–672.
140. Ahmed, T.; Al Aziz, M.; Mohammed, N. De-identification of electronic health record using neural network. *Sci. Rep.* **2020**, *10*, 18600. [CrossRef]
141. Guan, Z.; Lv, Z.; Du, X.; Wu, L.; Guizani, M. Achieving data utility-privacy tradeoff in Internet of Medical Things: A machine learning approach. *Futur. Gener. Comput. Syst.* **2019**, *98*, 60–68. [CrossRef]
142. Vaidyam, A.N.; Wisniewski, H.; Halamka, J.D.; Kashavan, M.S.; Torous, J.B. Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. *Can. J. Psychiatry* **2019**, *64*, 456–464. [CrossRef] [PubMed]
143. Abd-Alrazaq, A.; Safi, Z.; Alajlani, M.; Warren, J.; Househ, M.; Denecke, K. Technical Metrics Used to Evaluate Health Care Chatbots: Scoping Review. *J. Med. Internet Res.* **2020**, *22*, e18301. [CrossRef] [PubMed]
144. Campbell, J.I.; Eyal, N.; Musiimenta, A.; Haberer, J.E. Ethical Questions in Medical Electronic Adherence Monitoring. *J. Gen. Intern. Med.* **2015**, *31*, 338–342. [CrossRef]
145. Golshahi, J.; Ahmadzadeh, H.; Sadeghi, M.; Mohammadifard, N.; Pourmoghaddas, A. Effect of self-care education on lifestyle modification, medication adherence and blood pressure in hypertensive adults: Randomized controlled clinical trial. *Adv. Biomed. Res.* **2015**, *4*, 204–209.
146. Molugulu, N.; Gubbiyappa, K.S.; Murthy, C.R.V.; Lumae, L.; Mruthyunjaya, A.T. Evaluation of self-reported medication adherence and its associated factors among epilepsy patients in Hospital Kuala Lumpur. *J. Basic Clin. Pharm.* **2016**, *7*, 105–109. [CrossRef] [PubMed]
147. Bavan, L.; Surmacz, K.; Beard, D.; Mellon, S.; Rees, J. Adherence monitoring of rehabilitation exercise with inertial sensors: A clinical validation study. *Gait Posture* **2019**, *70*, 211–217. [CrossRef] [PubMed]

148. Wang, L.; Fan, R.; Zhang, C.; Hong, L.; Zhang, T.; Chen, Y.; Liu, K.; Wang, Z.; Zhong, J. Applying Machine Learning Models to Predict Medication Nonadherence in Crohn's Disease Maintenance Therapy. *Patient Prefer. Adherence* **2020**, *14*, 917–926. [CrossRef] [PubMed]
149. Aldeer, M.; Javanmard, M.; Martin, R.P. A Review of Medication Adherence Monitoring Technologies. *Appl. Syst. Innov.* **2018**, *1*, 14. [CrossRef]
150. Chai, P.R.; Castillo-Mancilla, J.; Buffkin, E.; Darling, C.; Rosen, R.K.; Horvath, K.J.; Boudreaux, E.D.; Robbins, G.K.; Hibberd, P.L.; Boyer, E.W. Utilizing an Ingestible Biosensor to Assess Real-Time Medication Adherence. *J. Med. Toxicol.* **2015**, *11*, 439–444. [CrossRef] [PubMed]
151. Chen, M.; Hao, Y.; Hwang, K.; Wang, L.; Wang, L. Disease Prediction by Machine Learning Over Big Data from Healthcare Communities. *IEEE Access* **2017**, *5*, 8869–8879. [CrossRef]
152. Mohan, S.; Thirumalai, C.; Srivastava, G. Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques. *IEEE Access* **2019**, *7*, 81542–81554. [CrossRef]
153. Venkatesh, R.; Balasubramanian, C.; Kaliappan, M. Development of Big Data Predictive Analytics Model for Disease Prediction using Machine learning Technique. *J. Med. Syst.* **2019**, *43*, 272. [CrossRef] [PubMed]
154. Cooper, G.F.; Aliferis, C.F.; Ambrosino, R.; Aronis, J.; Buchanan, B.G.; Caruana, R.; Fine, M.J.; Glymour, C.; Gordon, G.; Hanusa, B.H.; et al. An evaluation of machine-learning methods for predicting pneumonia mortality. *Artif. Intell. Med.* **1997**, *9*, 107–138. [CrossRef]
155. Rose, S. Mortality Risk Score Prediction in an Elderly Population Using Machine Learning. *Am. J. Epidemiol.* **2013**, *177*, 443–452. [CrossRef]
156. van Doorn, W.P.; Stassen, P.M.; Borggreve, H.F.; Schalkwijk, M.J.; Stoffers, J.; Bekers, O.; Meex, S.J. A comparison of machine learning models versus clinical evaluation for mortality prediction in patients with sepsis. *PLoS ONE* **2021**, *16*, e0245157. [CrossRef] [PubMed]
157. Raj, R.; Luostarinen, T.; Pursiainen, E.; Posti, J.P.; Takala, R.S.K.; Bendel, S.; Konttila, T.; Korja, M. Machine learning-based dynamic mortality prediction after traumatic brain injury. *Sci. Rep.* **2019**, *9*, 17672. [CrossRef]
158. Rau, C.-S.; Kuo, P.-J.; Chien, P.-C.; Huang, C.-Y.; Hsieh, H.-Y.; Hsieh, C.-H. Mortality prediction in patients with isolated moderate and severe traumatic brain injury using machine learning models. *PLoS ONE* **2018**, *13*, e0207192. [CrossRef] [PubMed]
159. Awad, E.; Dsouza, S.; Kim, R.; Schulz, J.; Henrich, J.; Shariff, A.; Bonnefon, J.-F.; Rahwan, I. The Moral Machine experiment. *Nature* **2018**, *563*, 59–64. [CrossRef] [PubMed]
160. Wang, J.; Ji, J.; Zhang, M.; Lin, J.-W.; Zhang, G.; Gong, W.; Cen, L.-P.; Lu, Y.; Huang, X.; Huang, D.; et al. Automated Explainable Multidimensional Deep Learning Platform of Retinal Images for Retinopathy of Prematurity Screening. *JAMA Netw. Open* **2021**, *4*, e218758. [CrossRef] [PubMed]
161. Zuallaert, J.; Godin, F.; Kim, M.; Soete, A.; Saeys, Y.; De Neve, W. SpliceRover: Interpretable convolutional neural networks for improved splice site prediction. *Bioinformatics* **2018**, *34*, 4180–4188. [CrossRef]
162. Ribeiro, M.T.; Singh, S.; Guestrin, C. Model-agnostic interpretability of machine learning. *arXiv* **2016**, arXiv:1606.05386.
163. Visani, G.; Bagli, E.; Chesani, F. OptiLIME: Optimized LIME explanations for diagnostic computer algorithms. *arXiv* **2020**, arXiv:2006.05714.
164. Salih, A.; Galazzo, I.B.; Raisi-Estabragh, Z.; Petersen, S.E.; Gkontra, P.; Lekadir, K.; Menegaz, G.; Radeva, P. A new scheme for the assessment of the robustness of Explainable Methods Applied to Brain Age estimation. In Proceedings of the 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS), Aveiro, Portugal, 7–9 June 2021.
165. Shickel, B.; Loftus, T.J.; Adhikari, L.; Ozrazgat-Baslanti, T.; Bihorac, A.; Rashidi, P. DeepSOFA: A Continuous Acuity Score for Critically Ill Patients using Clinically Interpretable Deep Learning. *Sci. Rep.* **2019**, *9*, 1879. [CrossRef]
166. Hartono, P. A transparent cancer classifier. *Health Inform. J.* **2018**, *26*, 190–204. [CrossRef] [PubMed]
167. Park, S.; Kim, Y.J.; Kim, J.W.; Park, J.J.; Ryu, B.; Ha, J.-W. Interpretable Prediction of Vascular Diseases from Electronic Health Records via Deep Attention Networks. In Proceedings of the IEEE 18th International Conference on Bioinformatics and Bioengineering (BIBE), Taichung, Taiwan, 29–31 October 2018; pp. 110–117.
168. Bernardini, M.; Romeo, L.; Misericordia, P.; Frontoni, E. Discovering the Type 2 Diabetes in Electronic Health Records Using the Sparse Balanced Support Vector Machine. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 235–246. [CrossRef] [PubMed]
169. Ming, Y.; Qu, H.; Bertini, E. RuleMatrix: Visualizing and Understanding Classifiers with Rules. *IEEE Trans. Vis. Comput. Graph.* **2018**, *25*, 342–352. [CrossRef] [PubMed]
170. Xiao, C.; Ma, T.; Dieng, A.B.; Blei, D.M.; Wang, F. Readmission prediction via deep contextual embedding of clinical concepts. *PLoS ONE* **2018**, *13*, e0195024. [CrossRef] [PubMed]
171. Montavon, G.; Lapuschkin, S.; Binder, A.; Samek, W.; Müller, K.-R. Explaining nonlinear classification decisions with deep Taylor decomposition. *Pattern Recognit.* **2017**, *65*, 211–222. [CrossRef]
172. Silva, P.E.; Maldaner, V.; Vieira, L.; de Carvalho, K.L.; Gomes, H.; Melo, P.; Babault, N.; Cipriano, G.; Durigan, J.L.Q. Neuromuscular electrophysiological disorders and muscle atrophy in mechanically-ventilated traumatic brain injury patients: New insights from a prospective observational study. *J. Crit. Care* **2018**, *44*, 87–94. [CrossRef]
173. Menegotto, A.B.; Becker, C.D.L.; Cazella, S.C. Computer-aided diagnosis of hepatocellular carcinoma fusing imaging and structured health data. *Health Inf. Sci. Syst.* **2021**, *9*, 20. [CrossRef]

174. Simpson, S.; Kaufmann, M.C.; Glozman, V.; Chakrabarti, A. Disease X: Accelerating the development of medical countermeasures for the next pandemic. *Lancet Infect. Dis.* **2020**, *20*, e108–e115. [CrossRef]
175. Higgins, M.K. Can we AlphaFold our way out of the next pandemic? *J. Mol. Biol.* **2021**, *433*, 167093. [CrossRef] [PubMed]
176. Li, J.; Zhang, S.; Li, B.; Hu, Y.; Kang, X.-P.; Wu, X.-Y.; Huang, M.-T.; Li, Y.-C.; Zhao, Z.-P.; Qin, C.-F.; et al. Machine Learning Methods for Predicting Human-Adaptive Influenza A Viruses Based on Viral Nucleotide Compositions. *Mol. Biol. Evol.* **2019**, *37*, 1224–1236. [CrossRef]