Information Fusion 123 (2025) 103283

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Full length article

Harnessing the potential of multimodal EHR data: A comprehensive survey of clinical predictive modeling for intelligent healthcare

Jialun Wu [a], Kai He [b], Rui Mao [c], Xuequn Shang [a], Erik Cambria [c,*]

[a] School of Computer Science, Northwestern Polytechnical University, China
[b] Saw Swee Hock School of Public Health, National University of Singapore, Singapore
[c] College of Computing and Data Science, Nanyang Technological University, Singapore

ARTICLE INFO

Keywords:
Intelligent healthcare
Medical intelligence
Electronic health records
Clinical predictive modeling

ABSTRACT

The digitization of healthcare has led to the accumulation of vast amounts of patient data through Electronic Health Records (EHRs) systems, creating significant opportunities for advancing intelligent healthcare. Recent breakthroughs in deep learning and information fusion techniques have enabled the seamless integration of diverse data sources, providing richer insights for clinical decision-making. This review offers a comprehensive analysis of predictive modeling approaches that leverage multimodal EHR data, focusing on the latest methodologies and their practical applications. We classify the current advancements from both task-driven and method-driven perspectives, while distilling key challenges and motivations that have fueled these innovations. This exploration examines the real-world impact of advanced technologies in healthcare, addressing issues from data integration to task formulation, challenges, and method refinement. The role of information fusion in enhancing model performance is also emphasized. Building on the discussions and findings, we highlight promising future research directions critical for advancing multimodal fusion technologies in clinical predictive modeling, addressing the complex challenges of real-world clinical environments, and moving toward universal intelligence in healthcare.

1. Introduction

Over the past two decades, the adoption of Electronic Health Records (EHR) systems has grown significantly. This growth has been driven by several factors. Policy initiatives, such as the HITECH Act in the United States [1], have supported their adoption. Global demographic changes, including an aging population [2], have also contributed. In addition, advancements in technology, like cloud computing and the Internet of Things (IoT), now allow for real-time health monitoring and large-scale data storage [3]. Together, these factors have made EHR systems a cornerstone of modern healthcare.

EHR systems not only serve as repositories of vast patient data but also play a pivotal role in advancing precision medicine, a healthcare model designed to tailor medical care to individual patient characteristics (see Fig. 1). By integrating structured data (*e.g.*, diagnoses, treatments, and lab results) with unstructured data (*e.g.*, medical images [4,5], clinical notes [6], and genomic information [7,8]), EHRs provide a comprehensive foundation for tasks such as medical event prediction [9,10], biomarker discovery [11,12], and patient-specific modeling [13,14].

Deep learning enhances these applications by extracting complex features from multimodal data and predicting clinical outcomes [15–17], enabling highly targeted interventions that optimize individual health outcomes, improve resource utilization, and reduce overall healthcare costs [18–20]. Building on these advancements, the transition from traditional machine learning methods to deep learning has marked a significant evolution in clinical predictive modeling. While conventional methods, such as linear regression [21], logistic regression [22], decision trees [23], and random forests [24], have been widely used, their limitations in capturing the complex, nonlinear relationships inherent in EHR data have become apparent. Deep learning and fusion strategy address these challenges by leveraging the rich, multimodal nature of EHR data, enabling tasks such as accurate diagnostics, personalized treatment, and tracking patient trajectories and disease progression [25]. Despite these advances, challenges persist [26], such as data heterogeneity [27,28], data imbalance [29,30], and issues of model interpretability and reliability [31, 32]. Tackling these challenges has spurred the development of innovative approaches, such as information fusion techniques, which integrate diverse data sources to provide a more holistic view of patients.

---

* Corresponding author.
E-mail address: cambria@ntu.edu.sg (E. Cambria).

https://doi.org/10.1016/j.inffus.2025.103283
Received 17 January 2025; Received in revised form 3 April 2025; Accepted 29 April 2025

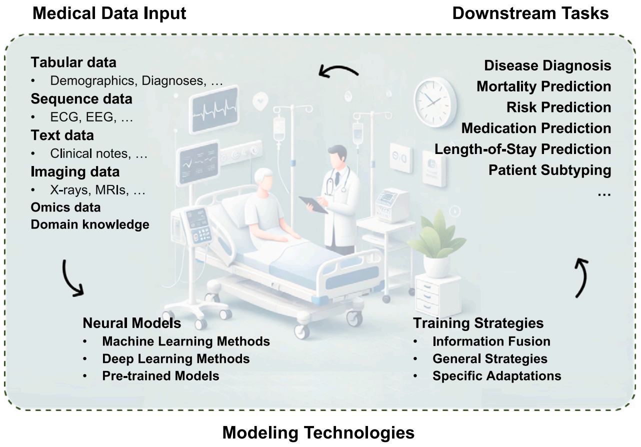


Fig. 1. Illustration of clinical predictive modeling in healthcare, from perspectives of medical data input and modeling technologies to downstream tasks.

These techniques enhance predictive accuracy, facilitate advanced representation learning [33,34], and optimize healthcare outcomes by enabling targeted treatments, faster disease detection, and efficient resource allocation. Numerous studies have investigated this area from various perspectives, revealing deep learning's transformative potential in clinical predictive modeling. For instance, Wang et al. [35] center their review on the predictive modeling of Electronic Health Record (EHR) data using deep learning techniques, emphasizing both the rigor of theoretical frameworks and methodologies as well as the practical tools and evaluation systems employed. In contrast, Nasarudin et al. [36] underscored the broad applications of deep learning in healthcare by focusing on its data-driven capabilities while omitting detailed technical methodologies. Si et al. [13] presented advanced techniques for developing patient profiles from EHR data, whereas Amirahmadi et al. [37] explored models designed to analyze complex medical trajectories. Addressing critical challenges, Shickel et al. [38] investigated issues such as model interpretability and data heterogeneity. Furthermore, innovations such as natural language processing have extended the role of deep learning in EHR-based research, as evidenced in immunology and allergy studies [39] and during the COVID-19 pandemic [40], thereby highlighting its broader impact on public health and crisis management. Unlike reviews that predominantly concentrate on machine learning algorithms or model performance, Ali et al. [41] focus specifically on the contribution of big data platforms and tools for managing massive heterogeneous datasets, and they discuss how these technologies facilitate efficient data management and real-time processing.

While the aforementioned surveys have made significant strides in categorizing studies by methodologies or tasks, they often provide fragmented perspectives that overlook the evolving interplay between research tasks, methodological innovations, and their implications for healthcare decision-making. Recognizing this gap, our review adopts a broader perspective that not only categorizes research by task typologies and methodologies but also critically examines the motivations, evolution, and correlation between methods. This approach highlights the transformative role of deep learning in healthcare by integrating datasets, methodologies, and practical applications, offering a unique perspective on both challenges and opportunities in the field (see Fig. 2). To the best of our knowledge, this is the first comprehensive review that holistically examines the datasets, techniques, challenges, and future directions of deep learning-driven clinical predictive modeling, aiming to inspire interdisciplinary innovations in healthcare. Importantly, our review identifies and categorizes the key motivations behind methodological innovations which can be grouped into the following key aspects: **(1) Data Quality** — Issues such as data scarcity, noise, and the complexity of multimodal and imbalanced datasets remain critical barriers to model performance. **(2) Time-Series Data Processing** — Irregular intervals and missing values in time-series data complicate accurate health predictions. **(3) Complex Relationships** — The nonlinear and intricate nature of medical data requires sophisticated models to effectively interpret relationships. **(4) Domain Knowledge Integration** — Incorporating domain-specific knowledge enhances model accuracy and interpretability, providing critical clinical context. **(5) Optimization Strategies** — Techniques such as pre-training and advanced loss functions are essential for handling limited data and novel patient scenarios. **(6) Model Transparency and Uncertainty** — Ensuring interpretability and improving uncertainty estimation are vital for building trust among clinicians and patients, ultimately enhancing decision-making reliability. By addressing these challenges, our review not only highlights the current limitations but also sets the stage for future innovations, aligning with our aim to unlock the full potential of EHR data in clinical predictive modeling.

In addition to the aforementioned topics, our review also investigates the application of foundation models (FMs) and large language models (LLMs) in EHR-based clinical predictive modeling [42–45]. These advanced models demonstrate significant potential for capturing complex patterns in multimodal medical data, enabling a new era of intelligent healthcare analytics [46–49]. Furthermore, we explore the integration of EHR data with multi-omics datasets, which include genomics, transcriptomics, and metabolomics [50,51]. This integration offers a complementary perspective by combining clinical and molecular insights, enhancing the understanding of disease mechanisms, biomarker discovery, and personalized medical interventions [52–55]. Building on these findings, we have identified seven critical challenges and corresponding future directions that are essential to unlocking the potential of EHR data and driving innovations in clinical predictive modeling. These challenges encompass data quality and heterogeneity, multimodal data integration, model scalability, transparency, ethical considerations, and computational infrastructure. By addressing these challenges, our review aims to provide a comprehensive roadmap that fosters interdisciplinary collaboration and opens new opportunities for advancing the field of clinical predictive modeling.

To analyze advancements in deep learning methodologies for EHR data processing, we retrieved literature published between 2016 and October 2024 from Web of Science,[1] Google Scholar,[2] and PubMed.[3] Our search included studies with terms such as "Electronic Health Records (EHRs)" or "electronic medical records (EMRs)", and the selection process was systematically conducted following PRISMA guidelines [56], as illustrated in Fig. 3. Based on the screened literature, the key contributions of our review are as follows:

- We present a detailed and systematic review of clinical predictive modeling methods based on EHR data, offering a valuable reference for researchers and healthcare practitioners. Our review highlights the evolution of data mining methods and their diverse applications in academic research and clinical practice, addressing significant gaps in previous reviews.
- We propose a comprehensive review framework that maps the intricate relationships between EHR data, data mining techniques, problem-solving motivations, and biomedical applications. Our framework provides a systematic lens to understand the complexity of medical data, the integration of information fusion methods, and their transformative potential in biomedicine.
- We analyze influential research that has significantly advanced the application of information fusion techniques in the biomedical domain, including insights into the practical challenges and uncovering the transformative potential of these techniques.

---

[1] http://apps.webofknowledge.com

[2] https://scholar.google.com

[3] https://pubmed.ncbi.nlm.nih.gov

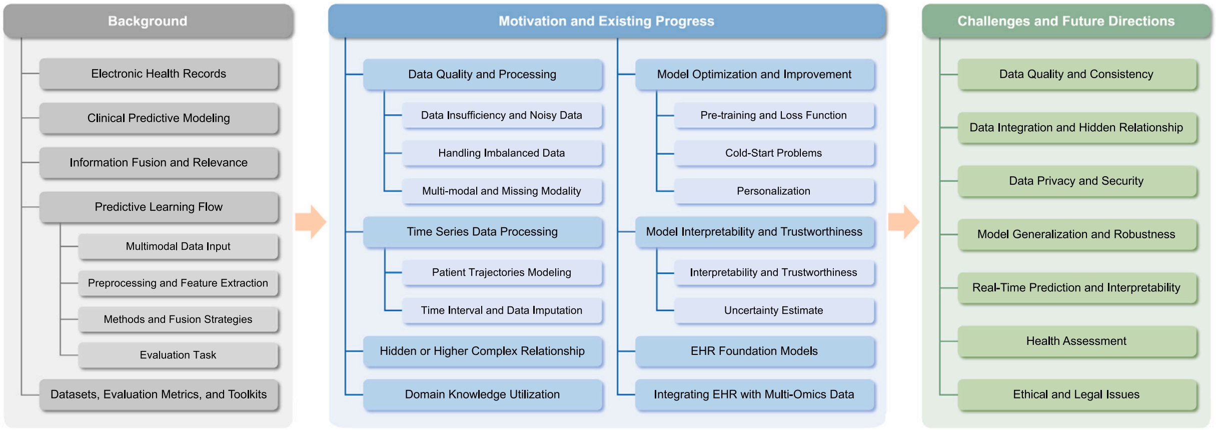


Fig. 2. Organizational framework of the content.

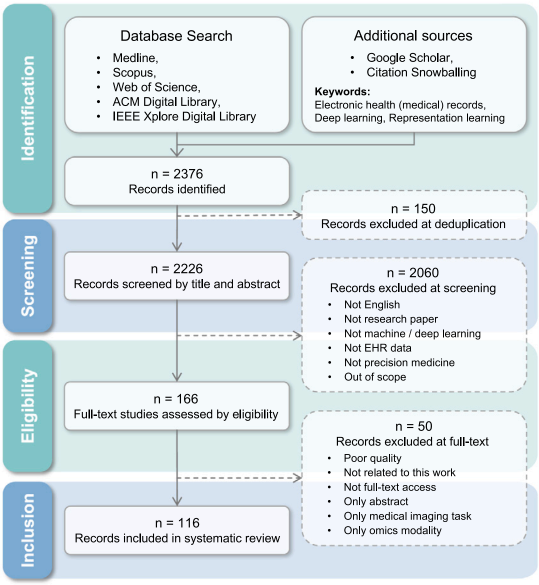


Fig. 3. PRISMA flowchart of the study selection process.

This review is structured as depicted in Fig. 2, beginning with an exploration of the foundational knowledge and key challenges associated with EHR data and predictive modeling. It then categorizes and critically analyzes the deep learning techniques employed in this field, highlighting significant advancements and identifying gaps that hinder further progress. By providing a systematic overview of methods, challenges, and opportunities, this review aims to serve as a roadmap for future innovations, offering actionable insights to amplify the potential of deep learning in transforming healthcare outcomes and advancing health informatics.

2. Background

EHRs serve as a cornerstone for advanced clinical predictive modeling by facilitating data-driven insights. Deep learning has significantly improved prediction accuracy and clinical decision-making through sophisticated techniques for patient representation and modeling. A key element in this progress is information fusion, which combines multi-level data to capture complex health trajectories. This section shows how the integration of deep learning with EHRs drives healthcare innovation, focusing on enhanced data integration, modeling strategies, and clinical applications.

2.1. Electronic Health Records (EHRs)

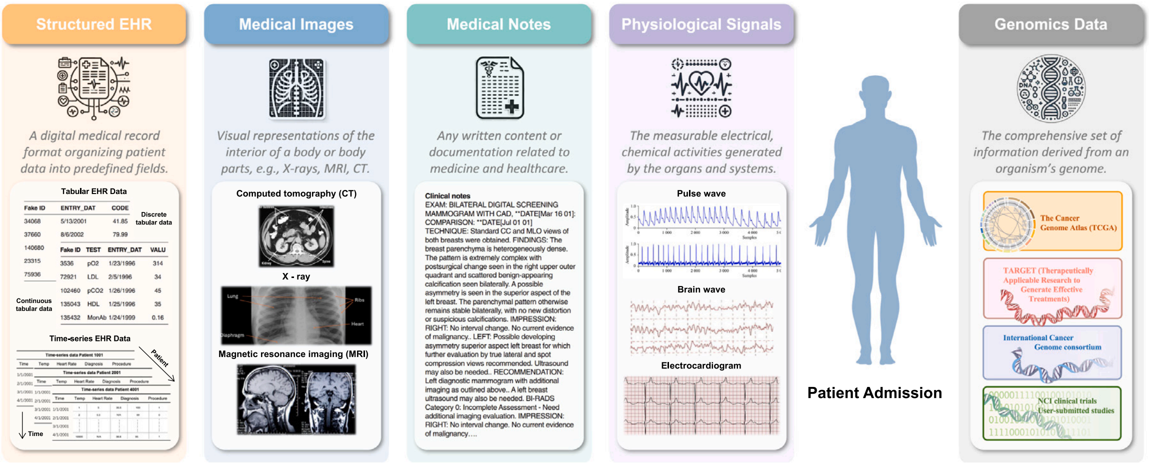
The transition to EHR systems marks a transformative shift in healthcare, replacing paper-based records with digital systems that improve accuracy, efficiency, and research capabilities. EHR systems minimize errors through consistent and accurate documentation, streamline data processes to enhance clinical decision-making, and provide comprehensive data for evidence-based research, advancing medical knowledge and policy-making.

Electronic Health Record (EHR) systems manage a diverse array of patient information, encompassing both structured and unstructured data [57] (see Fig. 4). Structured data can be further divided into tabular EHR and time-series EHR. Fig. 4 specifically depicts the tabular component — comprising dates, numerical values, and categorical variables (*e.g.*, diagnoses and medication classes) — which, along with time-series data that captures sequential information, is organized using predefined models for efficient analysis [58]. In contrast, unstructured data such as free-text clinical notes, medical images, and physiological signals require advanced processing techniques [4,59]. Moreover, although genomics data is sometimes integrated into EHR systems, it is more accurately classified as multi-omics data, derived from biological samples and necessitating specialized analytical approaches.

Despite their advantages, EHR systems face challenges due to the heterogeneity of coding systems used across institutions, as shown in Fig. 5. Standards like ICD, CPT, LOINC, and RxNorm support consistent data documentation and exchange but require integration tools like UMLS and SNOMED CT to address variations and promote interoperability [60]. Fig. 6 provides examples of these standards. Advances in technology and AI enable EHR systems to integrate multimodal data for personalized care and early disease detection, improving patient outcomes and resource allocation. Building on this foundation, we formalize EHR data representations to facilitate subsequent analysis and applications.

Definition 2.1 (*EHR Data*)**.** The EHR system in hospitals organizes detailed medical data for each patient in a systematic manner. Let $P = \{p_1, \dots, p_n\}$ represent the set of patients within the EHR system. Each patient p_i has a chronological sequence of medical events, denoted as $E_i = \{r_{i1}, r_{i2}, \dots, r_{im}\}$, where r_{ij} is the medical record at the jth time point for patient p_i. Each record r_{ij} consists of structured and unstructured data, expressed as $r_{ij} = (S_{ij}, U_{ij})$. Structured data S_{ij} includes key medical details such as diagnoses (d_{ij}), symptoms (s_{ij}), procedures (p_{ij}), and medications (m_{ij}). Unstructured data U_{ij} comprises medical images (i_{ij}) and textual notes (t_{ij}).

The integration of structured and unstructured components forms a comprehensive EHR, offering a multidimensional view of the patient's medical history and health condition. This approach provides

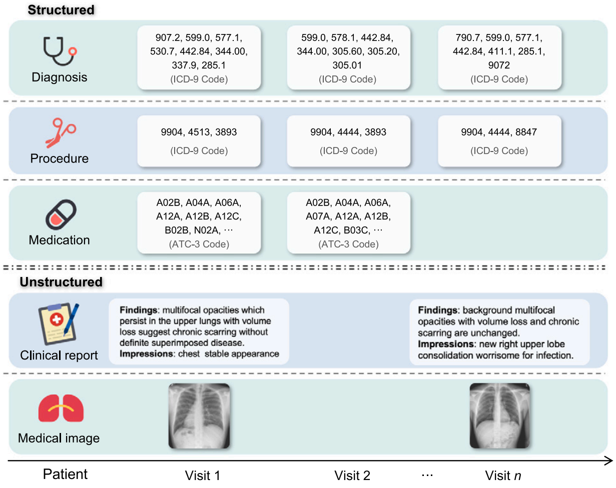


Fake ID	ENTRY_DAT	CODE
34068	5/13/2001	41.85
37660	8/6/2002	79.99
140680		
23315		
75936		

Fake ID	TEST	ENTRY_DAT	VALU
3536	pO2	1/23/1996	314
72921	LDL	2/5/1996	34
102460	pCO2	1/26/1996	45
135043	HDL	1/25/1996	35
135432	MonAb	1/24/1999	0.16

Clinical notes
EXAM: BILATERAL DIGITAL SCREENING MAMMOGRAM WITH CAD, **DATE[Mar 16 01]: COMPARISON: **DATE[Jul 01 01] TECHNIQUE: Standard CC and MLO views of both breasts were obtained. FINDINGS: The breast parenchyma is heterogeneously dense. The pattern is extremely complex with postsurgical change seen in the right upper outer quadrant and scattered benign-appearing calcification seen bilaterally. A possible asymmetry is seen in the superior aspect of the left breast. The parenchymal pattern otherwise remains stable bilaterally, with no new distortion or suspicious calcifications. IMPRESSION: RIGHT: No interval change. No current evidence of malignancy.. LEFT: Possible developing asymmetry superior aspect left breast for which further evaluation by true lateral and spot compression views recommended. Ultrasound may also be needed.. RECOMMENDATION: Left diagnostic mammogram with additional imaging as outlined above.. A left breast ultrasound may also be needed. BI-RADS Category 0: Incomplete Assessment - Need additional imaging evaluation. IMPRESSION: RIGHT: No interval change. No current evidence of malignancy….

Fig. 4. Categorization and examples of data types in multimodal EHR data and genomics data.



Findings: multifocal opacities which persist in the upper lungs with volume loss suggest chronic scarring without definite superimposed disease.
Impressions: chest stable appearance

Findings: background multifocal opacities with volume loss and chronic scarring are unchanged.
Impressions: new right upper lobe consolidation worrisome for infection.

Fig. 5. Example of multimodal time-series EHR data from a patient's health record.

healthcare providers with a holistic understanding of the patient's medical journey, enabling better clinical decision-making and personalized care.

2.2. Clinical predictive modeling

Clinical predictive modeling applies data analysis and machine learning to extract insights from complex healthcare data, enabling accurate predictions of patient health trajectories, treatment outcomes, and disease risks. By leveraging structured and unstructured EHR data, it identifies potential health risks and treatment opportunities, enhances diagnostic accuracy, supports early interventions, optimizes resource allocation, and facilitates personalized treatment plans, ultimately improving patient outcomes and quality of life. Here, we provide the mathematical definitions and symbolic representations of clinical predictive modeling.

Definition 2.2 (*Clinical Predictive Modeling*)**.** Clinical predictive modeling involves using predictive models to forecast future medical events. The input is the longitudinal visit records of the ith patient, represented as: $E_i = \{r_{i1}, r_{i2}, \dots, r_{im}\}$, where r_{im} denotes the medical record at the mth visit. The model aims to predict future health states or events, mathematically expressed as: $c_{m+1} = f(E_i; \theta)$, where c_{m+1} is the predicted clinical event, $f(\cdot)$ is the prediction model, and θ denotes the model parameters.

Deep neural networks provide a powerful framework for achieving these predictive objectives. By minimizing prediction errors through backpropagation, they iteratively optimize model parameters, enhancing their ability to accurately forecast future health states and events.

2.3. Information fusion and its relevance to clinical predictive modeling

Integrating deep learning methods with EHRs significantly improves clinical prediction performance by effectively leveraging diverse data sources through systematic information fusion strategies. These strategies can be broadly categorized based on the level of fusion, including basic data fusion, intermediate feature fusion, high-level decision fusion, multimodal fusion, and user feedback fusion.

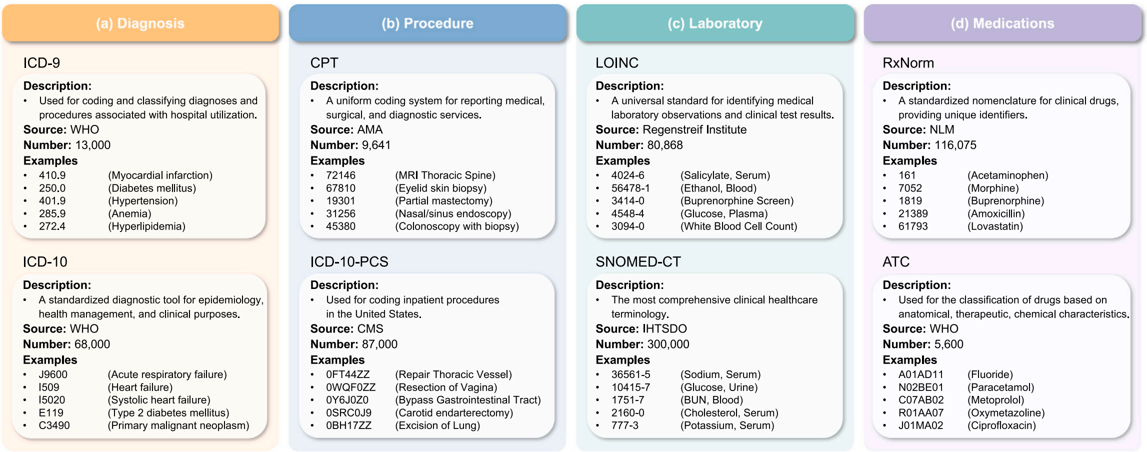
Basic Data Fusion involves preprocessing and merging raw data from heterogeneous clinical sources, ensuring data consistency through standardization, normalization, and noise removal. For instance, laboratory results obtained from different medical institutions may undergo uniform scaling and imputation techniques to manage missing or inconsistent values, thereby preparing reliable data inputs for predictive models.

Intermediate Feature Fusion integrates extracted features from distinct data modalities at intermediate layers within deep learning models, preserving essential semantic details while embedding contextual relevance. By combining structured clinical data (*e.g.*, vital signs and laboratory tests) with unstructured data (*e.g.*, physician notes or medical imaging features), models can better capture complex clinical scenarios and improve prediction accuracy.

High-Level Decision Fusion typically employs advanced modeling strategies, including ensemble learning, multi-task learning, and transfer learning. These methods integrate predictions from various modality-specific models or tasks to enhance the model's generalization and robustness. For example, an ensemble approach may aggregate decisions from independently trained classifiers focusing on radiological images and textual clinical notes, improving predictive reliability and reducing model bias.

Multimodal Fusion explicitly targets the integration of different data types — structured (*e.g.*, numerical lab results), textual (*e.g.*, clinical narratives), and imaging data (*e.g.*, X-rays or MRI scans) — to achieve comprehensive patient representation. Such integration can substantially enhance diagnostic precision and clinical decision-making. To effectively manage multimodal data, three primary fusion strategies are widely utilized, as illustrated in Fig. 7: early fusion, late fusion, and mid-level (joint) fusion.

Early Fusion combines multimodal features at the initial stage of model training, enabling direct interaction across data modalities during the entire training process [61,62]. Conversely, *Late Fusion*



(a) Diagnosis

ICD-9

Description:
- Used for coding and classifying diagnoses and procedures associated with hospital utilization.

Source: WHO
Number: 13,000
Examples
- 410.9 (Myocardial infarction)
- 250.0 (Diabetes mellitus)
- 401.9 (Hypertension)
- 285.9 (Anemia)
- 272.4 (Hyperlipidemia)

ICD-10

Description:
- A standardized diagnostic tool for epidemiology, health management, and clinical purposes.

Source: WHO
Number: 68,000
Examples
- J9600 (Acute respiratory failure)
- I509 (Heart failure)
- I5020 (Systolic heart failure)
- E119 (Type 2 diabetes mellitus)
- C3490 (Primary malignant neoplasm)

(b) Procedure

CPT

Description:
- A uniform coding system for reporting medical, surgical, and diagnostic services.

Source: AMA
Number: 9,641
Examples
- 72146 (MRI Thoracic Spine)
- 67810 (Eyelid skin biopsy)
- 19301 (Partial mastectomy)
- 31256 (Nasal/sinus endoscopy)
- 45380 (Colonoscopy with biopsy)

ICD-10-PCS

Description:
- Used for coding inpatient procedures in the United States.

Source: CMS
Number: 87,000
Examples
- 0FT44ZZ (Repair Thoracic Vessel)
- 0WQF0ZZ (Resection of Vagina)
- 0Y6J0Z0 (Bypass Gastrointestinal Tract)
- 0SRC0J9 (Carotid endarterectomy)
- 0BH17ZZ (Excision of Lung)

(c) Laboratory

LOINC

Description:
- A universal standard for identifying medical laboratory observations and clinical test results.

Source: Regenstreif Institute
Number: 80,868
Examples
- 4024-6 (Salicylate, Serum)
- 56478-1 (Ethanol, Blood)
- 3414-0 (Buprenorphine Screen)
- 4548-4 (Glucose, Plasma)
- 3094-0 (White Blood Cell Count)

SNOMED-CT

Description:
- The most comprehensive clinical healthcare terminology.

Source: IHTSDO
Number: 300,000
Examples
- 36561-5 (Sodium, Serum)
- 10415-7 (Glucose, Urine)
- 1751-7 (BUN, Blood)
- 2160-0 (Cholesterol, Serum)
- 777-3 (Potassium, Serum)

(d) Medications

RxNorm

Description:
- A standardized nomenclature for clinical drugs, providing unique identifiers.

Source: NLM
Number: 116,075
Examples
- 161 (Acetaminophen)
- 7052 (Morphine)
- 1819 (Buprenorphine)
- 21389 (Amoxicillin)
- 61793 (Lovastatin)

ATC

Description:
- Used for the classification of drugs based on anatomical, therapeutic, chemical characteristics.

Source: WHO
Number: 5,600
Examples
- A01AD11 (Fluoride)
- N02BE01 (Paracetamol)
- C07AB02 (Metoprolol)
- R01AA07 (Oxymetazoline)
- J01MA02 (Ciprofloxacin)

Fig. 6. Example classification schema for diagnosis, procedure, laboratory tests, and medications.

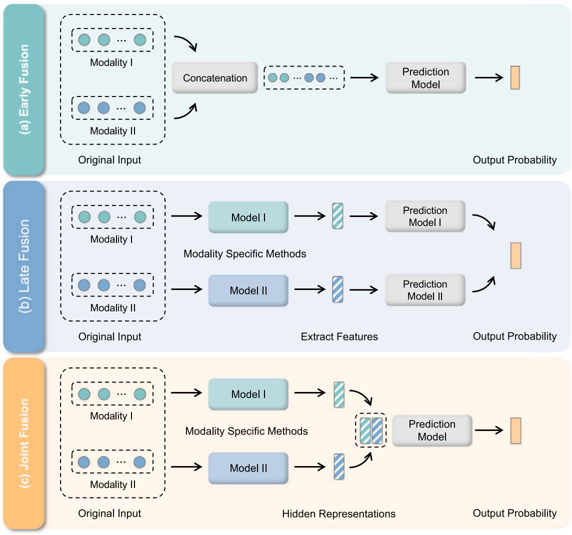


Fig. 7. Fusion strategies using deep learning. Model architecture for different fusion strategies. (a) Early fusion concatenates original or extracted features at the input level. (b) Late fusion aggregates predictions at the decision level. (c) Joint fusion also joins features at the input level, but the loss is propagated back to the feature extracting model.

(decision-level fusion) trains separate models for each modality independently, then integrates their final predictions using methods such as majority voting, weighted averaging, or meta-classifiers [63,64]. This approach allows each modality-specific model to specialize in its own domain, with the final decision made by combining the outputs from all models. *Mid-level (Joint) Fusion* merges intermediate representations generated by different modality-specific networks within shared hidden layers. This method facilitates joint optimization across modalities via backpropagation, effectively enhancing modality-specific feature extraction and model interpretability [65–67]. Empirical studies have demonstrated that multimodal fusion surpasses unimodal models in clinical predictive tasks. For instance, in the prediction of survival outcomes for non-small cell lung cancer (NSCLC) patients, Deng et al. [68] proposed a cross-modality attention-based multimodal fusion approach that integrates histopathological image data and RNA sequencing data. The results indicated that, compared to unimodal models, the multimodal fusion model improved the concordance index (C-index), signifying a substantial enhancement in predictive performance. Similarly, Braman et al. [69] developed a Deep Orthogonal Fusion (DOF) model that integrates radiological, pathological, genomic, and clinical data to predict overall survival in glioma patients. This multimodal model significantly outperforming the best unimodal model's C-index, further corroborating the advantage of multimodal fusion in clinical prediction.

Despite their effectiveness, fusion techniques encounter challenges such as data heterogeneity, computational complexity, and reduced interpretability. Variations in data formats, scales, and missing values necessitate advanced preprocessing strategies for effective integration. The computational resources required for multimodal fusion can lead to increased training times, prompting researchers to explore model compression, parallel processing, and hardware acceleration solutions. Moreover, complex multimodal models often sacrifice interpretability, challenging clinical trust. To mitigate this, methods such as attention visualization, feature importance analysis, and explanatory model components are increasingly utilized to improve transparency and clinician confidence. Future research should further refine data handling, computational efficiency, and interpretability to develop robust, clinically actionable predictive systems.

2.4. *Clinical predictive learning flow*

The integration of EHR with deep learning offers transformative potential for improving diagnostic accuracy and treatment efficiency. This process involves several stages, including consolidating structured and unstructured data with domain knowledge, preprocessing the data to ensure quality, training deep learning models with optimized features, and evaluating model performance for clinical predictions (Fig. 8). By leveraging this streamlined workflow, deep learning significantly enhances medical analysis and decision-making.

2.4.1. *Data input*

Data input is foundational for healthcare modeling, involving structured data, unstructured data, and domain knowledge.

Structured data, such as patient information, diagnostic codes, and test results, facilitates statistical analysis but often requires cleaning and standardization to address issues like inconsistency and missing values. Models such as AdaCare [70], INPREM [71], and GNDP [72] focus on disease diagnosis using structured inputs, while GAMENet [73] and SafeDrug [19] incorporate drug data for medication prediction, while 4SDrug [74] and DPGNN [75] integrate symptom data into their predictive frameworks.

Unstructured data, including physicians' notes and medical images, presents challenges like heterogeneity and complexity, requiring techniques like natural language processing (NLP) and convolutional neural networks (CNN) for feature extraction. Examples include CGL [76] and MedRetriever [77], which use text embeddings, and SCNet [66], which analyzes medical images.

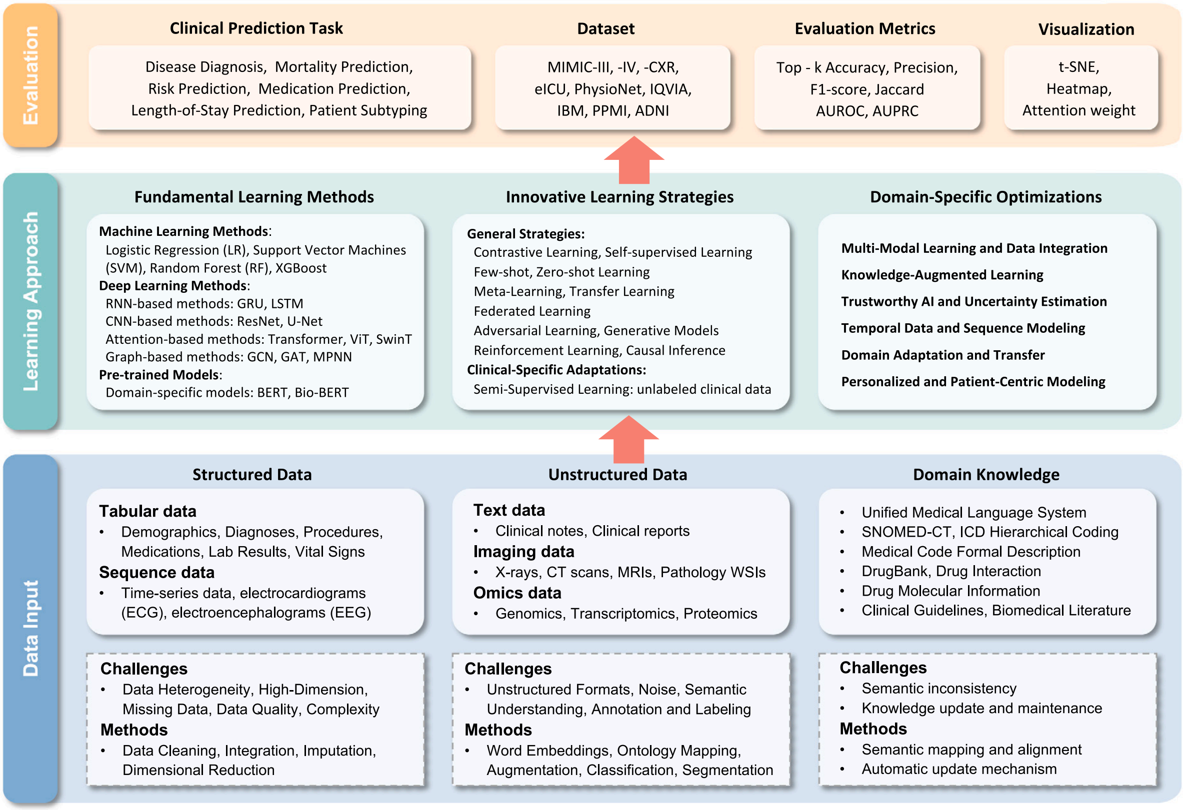


Fig. 8. Overview of EHR learning flow. The figure illustrates a general flow of steps for clinical decision support systems learning EHR data. The typical data analysis pipeline consists of data input, learning methods, and model evaluation.

Domain knowledge enhances model accuracy and interpretability. ICD coding hierarchies improve understanding of disease relationships, as seen in ADORE [78] and JMRL [79]. Semantic descriptions of medical codes, utilized in models like MEGACare [80], aid in processing detailed medical information. Drug molecular data from sources like DrugBank [81] supports treatment reliability, exemplified by SafeDrug [19] and MoleRec [82]. Clinical guidelines and biomedical literature provide evidence-based insights, leveraged by MMORE [83], HAP [84], and UNITE [85]. Integrating domain knowledge as data input poses challenges, notably semantic inconsistencies arising from diverse coding systems and terminologies across data sources [86,87]. Additionally, the fast-paced evolution of medical knowledge necessitates continuous updates to maintain data accuracy and relevance [88–90]. Robust integration of these data inputs is essential for accurate and reliable predictions.

2.4.2. Data preprocessing

Data preprocessing and feature extraction are critical steps to optimize model performance, tailored to different data types:

(I) Structured data processing. *(1) Handling missing data* is essential for maintaining the validity of healthcare analyses [91]. Approaches range from basic imputation methods (mean, median) to more advanced ones like multiple imputation and regression. Chen et al. [92] highlight how interpretable machine learning can effectively address this issue.

(2) Data normalization is also crucial for ensuring consistent scaling across variables, thus improving model accuracy. Techniques such as z-score normalization, min–max scaling, and mean normalization are commonly used, with Butwall's study showing its impact on healthcare model performance [93].

(3) Detecting and addressing outliers is important to prevent skewed analyses. Methods like box plots and isolation forests are used, and Cao et al. [94] propose a robust algorithm that combines outlier detection and data scaling for improved biomedical classification. For categorical data, encoding methods like One-Hot and Label Encoding are standard, although research on healthcare-specific encoding remains limited.

(4) Feature selection and transformation are vital for enhancing model performance, with techniques such as correlation analysis, variance filtering, and L1 regularization. Liu et al. [95] review deep learning-based imputation methods, emphasizing the importance of feature selection in handling missing healthcare data.

(II) Unstructured data processing. *(1) Text Data Processing* encompasses several key stages: First, text cleaning and normalization refine raw clinical data by removing noise (*e.g.*, special characters, excess spaces) and standardizing text (*e.g.*, stemming, lemmatization). This ensures uniformity, essential for clinical NLP, as highlighted by Kreimeyer et al. [96]. Second, tokenization and representation transform text into semantic vectors using techniques like Word2Vec, GloVe, and BERT. Methods like BPE and SentencePiece optimize embeddings for LLMs, as shown by Alsentzer et al. with clinical BERT [97]. Domain-specific preprocessing involves standardizing terminology (*e.g.*, UMLS, SNOMED CT), correcting spelling errors, and applying NER to extract entities like diseases and medications. Syntactic parsing enhances sentence-level understanding, aiding concept extraction and relationships. Demner-Fushman et al. demonstrate its role in clinical decision support systems [98]. Finally, LLM preparation tailors data for biomedical models like BioBERT, as shown by Lee et al. emphasizing tasks such as prompt design and fine-tuning [99].

(2) Image Data Processing improves quality and consistency through contrast enhancement techniques like histogram equalization and CLAHE, along with deep learning-driven noise reduction methods, such as Denoising Autoencoders [100]. Data augmentation enhances robustness using traditional methods (*e.g.*, rotation, scaling) and automated strategies (*e.g.*, AutoAugment), with generative models like

GANs and Diffusion Models enriching datasets [101]. Feature extraction employs CNNs, pre-trained models, and Transformer architectures, incorporating self-supervised learning and multimodal integration for comprehensive analysis [102]. Deep learning-based segmentation ensures precise anatomical delineation, while reinforcement learning-driven registration aligns images across times or devices for accurate spatial correspondence [103].

(3) Physiological Signals Processing, such as Butterworth filters, wavelet transformations, and Independent Component Analysis (ICA), improve the clarity of medical signals like ECG and EEG for clinical use [104]. Clinically significant features are extracted using methods like wavelet decompositions to enhance diagnostic accuracy [105]. Noise-robust solutions, including adaptive filters and deep learning-based denoising networks, enable precise and real-time medical signal processing [106]. Signal standardization and Principal Component Analysis (PCA) further ensure consistency and efficiency in processing medical data [107].

(4) Genomics Data Processing encompasses critical tasks such as sequence alignment, multi-omics integration, epigenomic analysis, and comparative genomics. Li and Durbin's work [108] on the Burrows–Wheeler transform revolutionized short-read alignment, ensuring speed and accuracy. Multi-omics integration, as reviewed by Picard et al. [109], enhances the holistic understanding of biological interactions. Kundaje et al. [110] provide valuable insights into regulatory mechanisms through an analysis of 111 human epigenomes. Comparative genomics, highlighted by Alkan et al. [111], underscores structural variations' role in genome diversity and evolution.

In summary, effective data preprocessing and feature extraction are critical for optimizing model performance across various data types. While structured data emphasizes data quality, unstructured data — including text and images — requires specialized transformations. Physiological signals benefit from robust noise reduction techniques, and genomics data processing focuses on sequence alignment, the integration of multi-omics data, and regulatory analysis, all of which enable the derivation of accurate and meaningful clinical insights.

2.4.3. Model learning approaches

Structured and unstructured data preprocessing underpins clinical prediction by converting healthcare datasets into analyzable formats. Beyond this, selecting appropriate learning methods is crucial for extracting actionable insights. This section reviews key learning approaches tailored to diverse data types and tasks and critically discusses the rationale behind employing each methodology.

(I) Fundamental learning methods provide the backbone for clinical predictions and are selected based on the nature of the data and specific clinical tasks. Their use is motivated by their interpretability, efficiency, and the ability to capture key patterns in structured and unstructured EHR data. (1) *Machine learning methods*, such as Logistic Regression (LR), Support Vector Machines (SVM), Random Forest (RF), and XGBoost, are typically employed for structured clinical data (*e.g.*, demographics, lab results, and diagnostic codes). LR is often favored for risk assessment due to its straightforward interpretability through coefficient analysis. SVMs are chosen when dealing with high-dimensional spaces, offering robust classification boundaries for disease state discrimination. RF and XGBoost, on the other hand, excel in capturing nonlinear interactions and complex relationships in heterogeneous data, though they may require extensive hyperparameter tuning to mitigate overfitting in noisy clinical datasets [112–115]. (2) *Deep learning approaches* have transformed the analysis of unstructured data by automating feature extraction. For example, RNN variants (*e.g.*, GRU [116] and LSTM [117]) are widely adopted in clinical predictive tasks due to their ability to model temporal dependencies in sequential EHR data such as visit histories or time-stamped lab results. Their sequential nature aligns well with the inherently chronological structure of medical records. However, they often struggle with long-range dependencies, which has led to the adoption of *attention-based architectures* for capturing long-term contextual information more effectively. CNNs, including ResNet [118] and U-Net [119], are primarily chosen for medical image analysis because of their ability to capture spatial hierarchies and local patterns in imaging data. They are highly effective for tasks such as disease detection and segmentation. Nonetheless, CNNs may fall short when it comes to processing data with strong temporal dependencies, where sequential context is crucial. (3) *Attention-based methods* (*e.g.*, Transformers [120], ViT [121], and SwinT [122]) have emerged as powerful alternatives to RNNs by enabling models to focus on the most relevant parts of the input data, thereby overcoming some limitations of traditional sequential models. Their ability to handle long-range dependencies and integrate multimodal information makes them particularly suitable for tasks that involve combining clinical notes with imaging data. (4) *Graph-based methods*, including GCN [123], GAT [124], and MPNN [125], are selected when the clinical task involves relational or network data, such as modeling interactions among patients, drugs, and diseases. These methods excel in capturing the underlying graph structures inherent in biomedical data. However, their computational complexity can be high, and constructing an effective graph representation from heterogeneous clinical data can be challenging. (5) *Pre-trained models*, such as BERT [126] and Bio-BERT [99], leverage large-scale biomedical corpora to achieve superior performance in natural language processing tasks, including clinical note summarization and entity recognition. Their ability to transfer knowledge from vast amounts of text data is invaluable; however, fine-tuning these models to specific clinical domains is often necessary to achieve optimal performance. In summary, the choice among these methods is largely determined by the specific characteristics of the clinical data and the targeted predictive task. Each method offers distinct strengths and faces unique challenges, necessitating careful consideration of their suitability for different applications in clinical predictive modeling.

(II) Innovative learning strategies are adopted to overcome challenges associated with high-dimensional, sparse, or limited clinical data, and to enhance model adaptability in dynamic healthcare settings. These strategies are particularly valuable when traditional supervised methods are constrained by the scarcity of labeled data or the complexity of EHR-derived tasks. (1) *Contrastive learning and self-supervised learning* leverage unlabeled data to reduce reliance on manual annotations. Contrastive learning, by bringing similar samples closer in the feature space while pushing dissimilar ones apart, is especially effective in tasks like medical image segmentation and patient clustering, where subtle differences can be crucial [127,128]. Self-supervised learning, on the other hand, generates pseudo-labels directly from inherent data structures, which is advantageous for tasks with limited labeled examples [127,129]. However, while these methods can significantly improve feature representation, they require careful design of the pretext tasks to ensure that the learned representations are clinically relevant. (2) *Few-shot and zero-shot learning* approaches are critical when dealing with rare diseases or emerging clinical conditions that lack ample labeled data. These methods allow models to generalize from minimal examples or even to predict unseen classes by leveraging inter-class relationships or external knowledge [130,131]. Their strength lies in the ability to quickly adapt to new scenarios, but they often face challenges in maintaining robustness and accuracy when the available examples are extremely limited. (3) *Meta-learning and transfer learning* enhance model adaptability by enabling rapid learning across diverse tasks. Meta-learning frameworks train models to quickly adjust to new, unseen tasks—a property that is particularly beneficial in the dynamic environment of clinical prediction where data distributions can vary significantly over time. Transfer learning, which often involves fine-tuning pre-trained models like BERT for EHR analysis, allows models to benefit from knowledge acquired on large external datasets [132–134]. However, both strategies can be sensitive to domain shifts, and their success often depends on the degree of similarity

between the pre-training and target domains. (4) *Federated learning* addresses data privacy concerns by allowing multiple institutions to collaboratively train models without sharing raw data. This approach is well-suited for multi-institutional studies where patient data privacy is paramount [135,136]. While federated learning enhances privacy and leverages a broader dataset, it introduces challenges related to data heterogeneity and communication overhead, which can affect convergence and model performance. (5) *Adversarial learning and generative models*, such as GANs, are utilized to improve data robustness and diversity. These methods are particularly effective for augmenting scarce datasets and addressing class imbalances in medical imaging [137,138]. The adversarial framework encourages the generation of realistic synthetic data, which can enrich training sets. Nonetheless, these methods often require complex training procedures and careful balancing of generator and discriminator dynamics to avoid mode collapse and maintain data fidelity. (6) *Reinforcement learning and causal inference* contribute to decision-making in clinical pathways by optimizing treatment strategies and uncovering causal relationships. Reinforcement learning models are designed to learn optimal policies through trial and error, making them ideal for sequential decision tasks such as treatment planning. Causal inference methods, meanwhile, help in identifying underlying cause–effect relationships, which are essential for risk prediction and therapeutic evaluations [139–141]. The primary challenge with these approaches is ensuring that the learned policies or inferred causal relationships are both reliable and interpretable in a clinical context. Each of these innovative strategies is selected based on the specific characteristics of the data and the clinical tasks at hand. While they offer substantial advantages in terms of flexibility and performance, they also bring challenges related to model complexity, interpretability, and robustness that must be carefully managed.

(III) Domain-specific optimizations tailor predictive models to the unique characteristics of clinical data, enhancing reliability, interpretability, and overall applicability. Below, we discuss several specialized strategies, providing motivation, data/task matching, and a balanced view of their strengths and challenges. *Multimodal Learning and Data Integration* leverages heterogeneous sources such as clinical text, imaging, and multi-omics data to construct a comprehensive patient view. This approach is motivated by the need to capture complementary information—while clinical notes offer contextual details, imaging and omics data can provide objective biomarkers. Models like vision–language frameworks (*e.g.*, BLIP and CLIP) are specifically chosen because they can correlate visual patterns with textual descriptions, thereby enhancing diagnostic accuracy [142–144]. However, integrating diverse data types poses challenges in terms of data alignment, normalization, and increased computational complexity. *Knowledge-Augmented Learning* incorporates medical ontologies and expert-curated rules into the modeling process to improve performance in tasks such as disease classification and treatment recommendation [145]. By embedding structured domain knowledge (*e.g.*, from ICD, SNOMED CT), these methods offer enhanced interpretability and context awareness that purely data-driven models might lack. While rule-based systems can refine predictions, they may require constant updates to remain current with evolving clinical guidelines, and overly rigid structures can sometimes limit the model's flexibility to learn novel patterns. *Trustworthy AI and Uncertainty Estimation* address the critical need for model transparency in clinical settings. Techniques such as attention maps, SHAP values, and LIME provide insights into model decision-making, which are crucial for clinician acceptance and trust [146,147]. Bayesian uncertainty estimation quantifies prediction confidence, guiding risk assessment in treatment planning. The trade-off, however, is that these methods can add layers of complexity and sometimes reduce computational efficiency, necessitating a balance between interpretability and performance. *Temporal Data and Sequence Modeling* focuses on capturing the dynamic aspects of clinical data, such as vitals and medication histories, which are inherently sequential. Transformer-based models and sequence-to-sequence architectures are preferred for their ability to model long-term dependencies, overcoming limitations seen in traditional RNNs that struggle with long-range contextual information [148,149]. Despite their superior performance in capturing temporal dynamics, these models typically require larger datasets and more computational resources, and they may be less interpretable compared to simpler models. *Domain Adaptation and Transfer Learning* are employed to ensure that predictive models remain robust and generalizable across different institutions and datasets. This strategy is particularly important given the variability in EHR systems and coding practices. By leveraging pre-trained models and fine-tuning them on domain-specific data, researchers can bridge discrepancies between diverse datasets [150,151]. However, adapting models to new domains often involves challenges in aligning feature spaces and may require additional calibration to maintain high accuracy in low-resource settings. *Personalized and Patient-Centric Modeling* aims to customize predictions based on individual patient characteristics, thus advancing precision medicine. This approach is motivated by the recognition that clinical decisions benefit from models that account for patient-specific variables. Techniques in this domain focus on tailoring outputs to reflect personal risk profiles and treatment responses [152,153]. The main challenge lies in effectively integrating individualized data without compromising model robustness or overfitting to idiosyncratic patterns. Together, these domain-specific optimizations enhance the capability of predictive models by aligning methodological choices with the specific needs of clinical data and tasks. They provide a more nuanced and targeted approach, although each comes with its own set of challenges that must be carefully managed to maximize clinical utility.

2.4.4. Downstream tasks

Building on the foundational concepts of Clinical Predictive Learning Flow, this section explores advanced predictive applications that tackle critical clinical challenges, such as disease diagnosis and personalized treatment recommendations.

Diagnosis prediction aims to determine the presence of specific conditions like diabetes or heart disease through EHR analysis, enabling preventive care and timely interventions. Notable works in this area include Dipole [154], which uses bidirectional recurrent neural networks (RNNs) and attention mechanisms to enhance predictive accuracy by analyzing temporal data; KAME [155], which integrates knowledge graph embeddings to incorporate external medical information; Stradegy [156], which applies convolutional neural networks to transform online medical descriptions into vector embeddings for disease prediction; and MEGACare [80], which builds an EHR hypergraph and multi-view learning framework to improve patient representations and model robustness by identifying high-order correlations.

Risk prediction estimates the likelihood of adverse events, such as heart attacks or hospital readmissions, to guide preventive measures and chronic disease management. ConCare [157] handles irregular EMR data by extracting inter-feature relationships for personalized risk prediction. StageNet [158] uses a stage-aware LSTM and stage-adaptive convolutional modules to improve predictive performance by modeling health stage transitions and progression patterns. UNITE [85] enhances disease risk prediction by integrating multi-sourced health data and incorporating uncertainty estimation. SCEHR [159] refines clinical risk prediction with supervised contrastive learning applied to longitudinal EHR data, improving representation learning. Furthermore, Al et al. [160] applies deep learning methods to analyze the Haberman's survival dataset in order to predict the long-term survival prospects of breast cancer patients following surgery, thereby providing a basis for tailoring treatment plans. In contrast, researchers utilize advanced ensemble classifiers combined with feature selection strategies to meticulously classify tumor viability in osteosarcoma histological slides, aiming to enhance diagnostic accuracy and robustness for optimized therapeutic decision-making [161].

Mortality prediction estimates patient survival probabilities, particularly in critical care settings, helping with resource allocation and prioritization. Several approaches have been proposed: GRASP [162] identifies patient similarities to learn from comparable cases, SAFARI [163] applies a sparsity prior for compact, interpretable representations that improve generalizability, and CTCL [164] integrates multi-view graph convolutional networks to combine structured and unstructured data, offering a unified patient representation for more accurate mortality predictions.

Readmission Prediction assesses the probability of disease recurrence or hospital readmission, aiding post-treatment monitoring and personalized care plans. HVec [165] uses multitask learning with a hierarchical vectorizer model to predict both mortality and readmission by leveraging shared patient information. TADEL [20] focuses on trajectory-based readmission prediction by modeling dynamic medical histories, offering insights into patient trajectories.

Medication prediction is essential for personalized therapy and drug recommendations. GAMENet [73] combines a drug–drug interaction graph with graph convolutional networks and longitudinal patient records to predict medications and ensure safe prescribing. SafeDrug [19] uses a global message-passing neural network and a local bipartite learning module to encode drug molecular connectivity and functionality. DAPSNet [166] leverages historical patient records to predict accurate drug combinations by evaluating diagnoses, procedures, and medications. PROMISE [167] integrates structured and unstructured data streams in a multimodal framework, improving medication prediction through separate but complementary learning pathways.

Length-of-Stay Prediction estimates the duration of a patient's hospitalization, which helps optimize hospital resource allocation and management strategies. Doctor AI [168] uses RNNs to analyze time-stamped EHR data for predicting hospital stay durations. RAIM [169] incorporates an attention mechanism to analyze continuous monitoring data like ECG signals, guided by discrete clinical events such as medication usage. CRXMDL [170] combines data from multiple modalities and uses an attention-based 1D convolutional neural network for accurate hospitalization duration predictions.

Similar Patient Retrieval helps identify historical cases that resemble the current patient, supporting clinical decisions and personalized treatment plans. GRASP [162] customizes patient similarity metrics for various tasks, aiding in knowledge extraction from relevant cases. Temporal Tree [171] employs a temporal hierarchical structure to capture multi-level information from temporal co-occurrence in EHR data, allowing for detailed patient comparisons.

2.4.5. Real-world case studies

Multimodal EHR integrates structured data, unstructured text, imaging features, and omics data, thereby significantly enhancing the performance of clinical prediction models across various application scenarios. In the field of oncology, the joint modeling of pathological images and pathology report texts not only captures immune infiltration and proliferative activity within the tumor microenvironment but also augments the biological interpretability of risk stratification for breast cancer recurrence [172]. Moreover, integrating metabolomics, transcriptomics, and genomics data further improves the sensitivity and specificity for detecting recurrent cases, with particularly notable enhancements in predicting high-risk subgroups such as patients with lymph node metastasis [173]. In addition, for gastric cancer treatment response prediction, the fusion of endoscopic images, histopathological reports, and patient baseline characteristics enables more accurate forecasts of response to anti-HER2 therapy, thereby reducing patients' exposure to ineffective treatments. Cross-cancer prognostic modeling, achieved through constructing cross-modal knowledge graphs that interrelate clinical records, genomic mutations, and imaging features, significantly enhances the robustness of overall survival predictions [174].

In the domains of cardiovascular medicine and critical care, data fusion based on chest imaging and dynamic vital signs (*e.g.*, NT-proBNP, oxygen saturation) has effectively improved the early detection of acute heart failure, thereby shortening diagnosis time and reducing misdiagnosis rates [175]. Furthermore, analysis of long-sequence EHR data — including medication adjustment records, progression of complications, and wearable device monitoring — has yielded dynamic risk assessment models that perform exceptionally well in predicting long-term cardiovascular event risks, thus providing real-time support for clinical interventions [176].

In the area of infectious diseases and public health, the integration of mobile phone location trajectories, environmental sensor data, and symptom reports from EHR not only facilitates the early identification of epidemic hotspots and close-contact populations but also significantly enhances the responsiveness of public health departments during emergencies. Simultaneously, combining wearable device monitoring data with clinical history enables digital twin models to simulate individual responses to specific interventions (*e.g.*, medication dosage adjustments), thereby optimizing personalized treatment strategies [172].

In summary, the integration of multimodal EHR data across various clinical applications, achieved through the collaborative and synergistic combination of heterogeneous data sources, substantially improves the accuracy of risk assessment, treatment response prediction, and prognostic evaluation, thereby providing a robust foundation for clinical decision support and healthcare resource optimization.

2.5. Datasets, evaluation metrics, and toolkits

This section outlines key publicly available EHR datasets and highlights essential toolkits and libraries designed to aid deep learning applications in healthcare.

2.5.1. Publicly available EHR datasets

Publicly available EHR datasets are essential in clinical research, offering extensive data for various studies. Key datasets include MIMIC-III, MIMIC-IV, MIMIC-CXR, eICU, IBM MarketScan, PCORnet, THIN, PPMI, and ADNI, each playing a significant role in advancing healthcare research. A summary of these datasets can be found in Table 1, with further details in the supplementary materials. These datasets provide invaluable resources for biomedical research, facilitating progress in understanding, preventing, and treating complex diseases. With data spanning ICU records to biomarker studies, they support innovation in fields like machine learning, epidemiology, and precision medicine, driving clinical research forward.

2.5.2. Evaluation metrics

Evaluation of predictive models in medical contexts relies on several key metrics tailored to different tasks. For disease prediction, AUROC *(Area Under the Receiver Operating Characteristic Curve)* [177], AUPRC *(Area Under the Precision–Recall Curve)* [178], and *accuracy* are commonly used to assess discriminative ability, performance under class imbalance, and overall prediction accuracy, respectively. *Top K recall and precision* further evaluate the model's ability to identify correct positives among the top predictions. Mortality prediction includes *F1 score* alongside the above metrics to balance precision and recall. For length-of-stay prediction, *MSE (Mean Squared Error)* [179] and *MAPE (Mean Absolute Percentage Error)* [180] are used to evaluate continuous prediction accuracy. Readmission prediction uses *AUROC*, *AUPRC*, *accuracy*, and *Top K metrics*. Prescription prediction additionally utilizes the *Jaccard Index* [181] to measure set similarity. For multi-class or multi-label tasks, macro and micro averaging provide a balanced view across classes. These metrics, collectively, offer a comprehensive framework to assess model performance in EHR-based predictive modeling.

Table 1
Overview of healthcare datasets.

Dataset	Count	Modalities	Period	Applications	Region
MIMIC-III	60,000 ICU admissions	Demographics, vital signs, lab tests (*e.g.*, hematology, biochemistry), medications, clinical notes	2001–2012	Epidemiology, machine learning, clinical decision-making, ICU outcomes	USA
MIMIC-IV	60,000 ICU admissions	Demographics, vital signs, lab tests (*e.g.*, microbiology, blood gases), medications, clinical notes, out-of-hospital mortality	2008–2019	Epidemiology, machine learning, predictive analytics, ICU research	USA
MIMIC-CXR	65,079 patients	Chest X-ray images, radiology reports, metadata (*e.g.*, age, gender, hospital ID)	2001–2008	Medical imaging, diagnostic algorithm development, imaging quality assessment	USA
eICU	200,000 ICU admissions	Demographics, diagnoses, treatments, outcomes, procedures	2014–2015	ICU practices, patient outcomes, multi-institutional benchmarking	USA
IQVIA	100+ million patients	Prescription trends, medical claims, real-world evidence, healthcare costs, demographics, treatment pathways	Ongoing	Epidemiology, health economics, treatment patterns, market analysis	Global
IBM MarketScan	273 million patients	Healthcare claims, patient demographics, medical conditions, treatments, costs, insurance enrollment data	1995–present	Health economics, outcomes research, payer-provider impact analysis	USA
PCORnet	30 million patients	EHRs, claims, patient-reported outcomes, lab results (*e.g.*, hematology, imaging), demographics, medications	2010-present	Large-scale clinical research, patient-centered outcomes research, comparative effectiveness studies	USA
THIN	69 million patients	Diagnoses, treatments, outcomes from primary care, prescriptions, patient demographics	1994-present	Epidemiological research, health policy analysis, chronic disease tracking	Europe
PPMI	5000 participants	Clinical, imaging, genetic, biomarker data, longitudinal follow-ups	Ongoing	Biomarker discovery, early diagnosis, disease progression monitoring	Global
ADNI	1700 participants	Clinical, imaging, biomarker, cognitive assessments, genetic data	2004–present	Diagnostic tools, treatment evaluation for Alzheimer's	USA

Table 2
Summary of various medical toolkits.

Toolkit	Primary focus	Framework	Key features	Integration	Deployment	User community	Data handling	License
PyHealth	EHRs predictive modeling, signals, text, and images	Python	Multiple modalities, advanced ML	EHRs, physiological signals, medical images	Local, cloud	Researchers, data scientists	Structured and unstructured data, multimodal data	MIT
MedCAT	Named Entity Recognition and Linking in biomedical text	Python	Self-supervised UMLS and SNOMED-CT integration	EHRs, biomedical text	Local, cloud	Biomedical researchers, clinicians	Unstructured text	Apache 2.0
Fasten Health	Personal and family EMRs aggregation	Self-hosted	Self-hosted, privacy-focused	Insurance companies, healthcare providers, laboratories	Self-hosted	Individuals, privacy advocates	Structured EHRs	GPL
MONAI	Medical imaging	PyTorch	Data preprocessing, AutoML, GPU acceleration	TensorBoard, NVIDIA FLARE	Local, cloud	Researchers, developers, clinicians	High-dimensional medical imaging	Apache 2.0
NiftyNet	Medical imaging	TensorFlow	Modular pipeline, regression, image generation, segmentation	TensorFlow, TensorBoard	Local, cloud	Researchers, developers	High-dimensional medical imaging	Apache 2.0
MedPy	Medical imaging	Python	Segmentation, feature extraction	Computing libraries	Local	Researchers, developers	High-dimensional medical imaging	GPL

2.5.3. Toolkits and libraries

Several specialized toolkits and libraries have been developed to facilitate deep learning applications in EHR data analysis, addressing both textual and imaging data. These toolkits, summarized in Table 2, include PyHealth [182], MedCAT [183], Fasten Health, MONAI [184], NiftyNet [185], and MedPy. PyHealth and MedCAT focus on processing structured EHRs and clinical texts, employing advanced machine learning and NLP techniques for tasks like patient risk prediction and medical concept linking. On the other hand, MONAI, NiftyNet, and MedPy specialize in medical imaging, offering tools for segmentation, classification, and image synthesis. These toolkits are optimized for different frameworks such as PyTorch (MONAI and MedPy) and TensorFlow (NiftyNet), and address unique challenges in their respective domains. Fasten Health, distinct from the others, is a self-hosted platform that emphasizes data security and HIPAA-compliant management of personal health records. These resources collectively support innovation in healthcare, catering to diverse research needs and clinical applications. To effectively support these high-level applications, robust underlying big data platforms are essential [186].

2.5.4. Application of big data platforms and tools

The integration of big data platforms into personalized medicine facilitates the management and analysis of vast amounts of heterogeneous healthcare data [41]. *Hadoop Distributed File System (HDFS)*: Used for distributed storage of massive healthcare data, ensuring data persistence and high availability. *Apache Spark*: As a fast distributed computing engine, Spark enables real-time processing and complex computations on large datasets, supporting online analytics and rapid feedback.

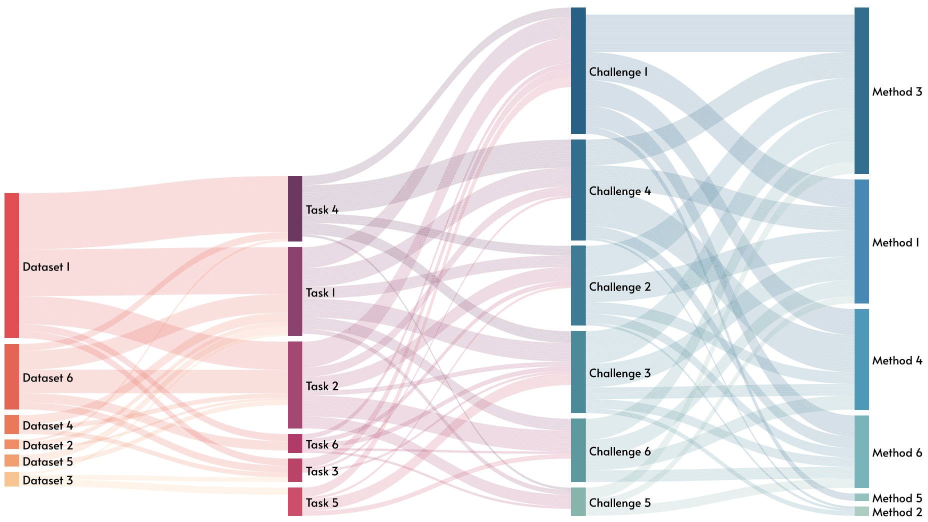


Fig. 9. Motivation-Driven Sankey Diagram. The figure illustrates the relationship between different datasets, tasks, challenges, and methods. **Datasets**: Dataset1: MIMIC, Dataset2: IQVIA, Dataset3: Physionet, Dataset4: eICU, Dataset5: Diabetes, Dataset6: Others. **Tasks**: Task1: Diagnosis Prediction, Task2: Risk Prediction, Task3: Mortality Prediction, Task4: Medication Prediction, Task5: Length-of-Stay Prediction, Task6: Readmission Prediction. **Challenges**: Challenge1: Data Quality and Processing, Challenge2: Time Series Data Processing, Challenge3: Hidden or Complex Relationships, Challenge4: Domain Knowledge Utilization, Challenge5: Model Optimization and Improvement, Challenge6: Model Interpretability and Trustworthiness. **Methods**: Method1: Recurrent Neural Networks (RNNs) and Variants, Method2: Convolutional Neural Networks (CNN), Method3: Attention Mechanisms and Transformer Models, Method4: Knowledge Graphs and Graph Neural Networks (GNN), Method5: Generative Adversarial Networks (GANs), Method6: Other Methods (*e.g.*, Hierarchical Propagation, Memory Networks, Pre-training, Meta-Learning, Contrastive Learning).

Apache Hive: Through its SQL interface, Hive allows efficient querying and analysis of data stored in HDFS, helping researchers extract key indicators and uncover hidden patterns in the data.

3. Motivation and existing progress

In the previous section, we classified and summarized existing research on model architectures and task types, providing a concise overview of the field. This chapter builds on that foundation, exploring recent advancements in predictive modeling through a motivation-driven lens. Key motivations include improving data quality and preprocessing, addressing time-series challenges, modeling complex relationships, integrating domain knowledge for accuracy, using advanced optimization techniques, and prioritizing interpretability for ethical deployment.

To organize these efforts, we created a hierarchical framework that connects datasets to tasks, tasks to challenges, and challenges to methods. This motivation-driven perspective clarifies the alignment of methods with specific challenges, highlighting gaps and encouraging targeted innovation. Unlike method-driven reviews, which focus on technical details, or task-driven approaches, which isolate tasks, our framework emphasizes the motivations behind research, enhancing understanding of shared challenges like data heterogeneity, temporal irregularities, and bias mitigation. By linking challenges across domains, this approach fosters cross-disciplinary insights and connects theoretical advancements with practical needs. It also improves interpretability, providing a systematic roadmap for addressing complex problems and guiding researchers toward impactful real-world applications. This framework offers a comprehensive view of the research landscape, supporting the development of innovative, generalizable methodologies.

This integrated representation is visualized through a Sankey diagram (Fig. 9) and complemented by statistical summaries and classifications in Tables 3 and 4, providing a clear overview of the relationships among datasets, tasks, challenges, and methods in predictive modeling research.

3.1. Insights from the sankey diagram

The Sankey diagram reveals key insights into dataset-task alignments, methodological trends, and research priorities. Datasets like MIMIC (D1) and eICU (D4), rich in time-series data, are closely linked to short-term prediction tasks like diagnosis (T1) and medication prediction (T4). RNNs and Transformers (M1) are commonly used for these tasks, highlighting their importance in medical predictions. Chronic care tasks, such as medication prediction (T4) and readmission prediction (T6), align with datasets like IQVIA (D2) and Diabetes (D5), where domain knowledge (C4) is essential. Methods like GNNs and knowledge graphs (M4) are ideal for capturing patient-drug interactions and complex relationships. Versatile datasets, such as Physionet (D3), support multiple tasks like ECG and glucose-level predictions, leveraging techniques like meta-learning.

The analysis identifies several challenges and research priorities. Time-series datasets, such as MIMIC (D1) and Physionet (D3), emphasize data quality (C1) and sequence processing (C2), requiring RNNs and Transformers. Tasks involving hidden relationships (C3) and domain knowledge (C4) rely on methods like GNNs and knowledge graphs. Universal challenges, including model optimization (C5) and interpretability (C6), are prevalent across tasks, highlighting their importance for reliable healthcare solutions. These insights underscore the field's focus on both technical advancements and practical applications.

In summary, the Sankey diagram emphasizes the interplay between datasets, tasks, and methods in shaping research priorities. In our framework, *short-term predictions* refer to forecasting immediate outcomes over a limited time horizon — such as acute patient events — by leveraging time-series data and recurrent neural networks (RNNs). In contrast, *long-term predictions* encompass forecasting outcomes over extended periods, often within chronic care scenarios. These tasks typically require the incorporation of domain knowledge and the use of graph neural networks (GNNs) to capture complex, long-range interactions. Cross-cutting challenges such as model optimization and interpretability continue to drive ongoing efforts to enhance healthcare applications. The following subsections explore the innovations, practical relevance, and future directions of each research motivation,

Table 3
Summarization of recent work on predictive modeling in healthcare (Part 1). C1: Data Quality and Processing , C2: Time Series Data Processing, C3: Hidden or Higher Complex Relationship, C4: Domain Knowledge Utilization, C5: Model Optimization and Improvement, C6: Model Interpretability and Trustworthiness. M1: Recurrent Neural Networks (RNNs) and Variants, M2: Convolutional Neural Networks (CNN), M3: Attention Mechanisms and Transformer Models, M4: Knowledge Graph and Graph Neural Networks (GNN), M5: Generative Adversarial Networks (GANs), M6: Other Methods (*e.g.*, Hierarchical propagation, Memory network, Pre-training, Meta-Learning, Contrastive learning,). T1: Diagnosis Prediction, T2: Risk Prediction, T3: Mortality Prediction, T4: Medication Prediction, T5: Length-of-Stay Prediction, T6: Readmission Prediction.

Model	Venue	Challenge						Method						Task					
		C1	C2	C3	C4	C5	C6	M1	M2	M3	M4	M5	M6	T1	T2	T3	T4	T5	T6
Dipole	KDD		✓	✓				✓		✓				✓					
C-MemNNs	AAAI	✓			✓			✓					✓	✓					
MiME	NIPS	✓											✓	✓	✓				
KAME	CIKM	✓								✓	✓			✓					
RAIM	KDD	✓		✓						✓				✓				✓	
MCA-RNN	ICDM			✓									✓	✓					
CAMP	ICDM		✓	✓				✓		✓			✓	✓					
MNN	IJCAI	✓			✓			✓		✓				✓					
MMORE	IJCAI				✓					✓	✓			✓					
KSI	WWW				✓			✓		✓				✓					
HAP	KDD				✓					✓				✓					
MedPath	WWW						✓				✓			✓					
CGL	IJCAI				✓			✓			✓			✓					
RAPT	KDD					✓				✓			✓	✓					
SETOR	ICDM		✓	✓						✓	✓			✓					
Sherbet	Trans	✓					✓			✓	✓		✓	✓					
MetaCare++	SIGIR					✓							✓	✓					
Chet	AAAI		✓	✓							✓			✓					
MEGACare	InfFus	✓		✓	✓					✓	✓			✓					
KerPrint	AAAI						✓			✓	✓			✓					
SeqCare	WWW	✓									✓			✓					
RETAIN	NIPS		✓				✓	✓		✓				✓					
Flex-MoE	NIPS	✓		✓		✓				✓				✓					
Med-ST	ICML		✓	✓		✓				✓			✓	✓					
TRANS	IJCAI		✓	✓						✓			✓	✓					
MPRE	ICDM	✓	✓	✓		✓				✓			✓	✓					
GRAM	KDD	✓			✓		✓				✓			✓	✓				
GCT	AAAI			✓							✓			✓	✓				
PRIME	KDD				✓			✓							✓				
Health-ATM	SDM			✓				✓	✓						✓			✓	
DG-RNN	ICDM	✓	✓		✓		✓	✓			✓				✓				
KnowRisk	ICDM	✓			✓		✓			✓	✓				✓				
HiTANet	KDD		✓							✓					✓				
ConCare	AAAI		✓			✓				✓					✓				
AdaCare	AAAI		✓				✓	✓		✓					✓				
StageNet	WWW		✓					✓	✓						✓				
INPREM	KDD						✓			✓					✓				
LSAN	CIKM	✓	✓							✓			✓		✓				
CONAN	AAAI	✓										✓			✓				
MaskEHR	SDM	✓										✓			✓				
SCEHR	ICDM					✓							✓		✓				
UNITE	WWW					✓	✓						✓		✓				
MedRetriever	CIKM					✓	✓	✓		✓					✓				
LDAM	ICDM	✓					✓	✓					✓		✓				
MendMKG	ICDM	✓								✓	✓				✓				
INPLIM	Trans					✓	✓	✓		✓					✓				
T-ContextGAN	Trans	✓						✓		✓			✓		✓				
GRACE	KDD	✓								✓		✓			✓				
FlexCare	KDD		✓	✓	✓	✓				✓			✓		✓				

providing deeper insight into the evolving landscape of predictive modeling.

3.2. Data quality and processing

High-quality, integrated data are crucial for accurate healthcare predictions. However, EHR data often suffer from issues like insufficiency, noise, and heterogeneity, limiting model performance. Addressing these challenges requires strategies to improve data completeness, reduce noise, and integrate multimodal data effectively.

3.2.1. Data insufficiency and noisy data

Healthcare predictive models often struggle with insufficient and noisy data due to incomplete records, subjective reporting, system inconsistencies, and the impact of emerging diseases with limited historical data. Noise also arises from misdiagnoses, equipment errors, and inaccuracies. To tackle these issues, solutions like data cleaning, augmentation, and advanced preprocessing are employed.

GRAM [187] enhances sparse datasets by incorporating hierarchical medical ontologies and graph-based attention. MiME [188] improves small datasets by training on auxiliary tasks to generate robust embeddings. DistCare [189] employs knowledge distillation and transfer learning to align features across datasets. For further data enhancement, GRACE [190] combines GANs with contrastive learning to generate synthetic EHR data, and GNDP [72] integrates graph structures and external knowledge to boost accuracy. MendMKG [191] addresses high dimensionality and sparsity by leveraging medical knowledge graphs, GCNs, and RNNs. Additionally, models like DG-RNN [18] incorporate medical knowledge graphs and attention mechanisms to improve predictions in data-scarce situations, while KAME [155] and KnowRisk [192] enhance robustness and interpretability with attention mechanisms. SAFARI [163] tackles feature sparsity by using sparse priors and graph-based aggregation.

Table 4
Summarization of recent work on predictive modeling in healthcare (Part 2).

Model	Venue	Challenge						Method						Task					
		C1	C2	C3	C4	C5	C6	M1	M2	M3	M4	M5	M6	T1	T2	T3	T4	T5	T6
GRASP	AAAI	✓								✓						✓			
MIAM	JBHI			✓						✓						✓		✓	
SAFARI	Trans	✓			✓					✓						✓			
MUSE	ICLR	✓								✓						✓			
CTCL	Trans	✓						✓		✓						✓			
DNA-T	JBHI		✓	✓						✓						✓			
Warpformer	KDD		✓	✓		✓				✓			✓			✓		✓	
FuseMoE	NIPS	✓	✓	✓		✓					✓					✓		✓	
GAMENet	AAAI		✓		✓			✓			✓		✓				✓		
G-BERT	IJCAI	✓			✓			✓			✓		✓				✓		
CompNet	CIKM			✓	✓				✓	✓			✓				✓		
SafeDrug	IJCAI				✓			✓			✓						✓		
SARMR	IJCAI				✓								✓				✓		
MICRON	IJCAI		✓	✓				✓		✓							✓		
DrugRec	NIPS	✓	✓		✓					✓	✓						✓		
CSEDrug	CIKM				✓			✓		✓	✓						✓		
PREMIER	TOIS	✓		✓				✓		✓	✓						✓		
MoleRec	WWW				✓						✓						✓		
4SDrug	KDD	✓						✓									✓		
DRMP	JBHI					✓											✓		
COGNet	WWW		✓							✓			✓				✓		
REFINE	NIPS				✓						✓						✓		
MK-GNN	Trans				✓					✓	✓						✓		
OntoPath	Trans				✓						✓						✓		
Carmen	AAAI				✓						✓						✓		
SHAPE	JBHI			✓				✓		✓							✓		
VITA	AAAI			✓				✓		✓							✓		
DistCare	WWW	✓					✓						✓					✓	
SAnD	AAAI		✓				✓	✓		✓								✓	
DLFS	Trans			✓				✓		✓								✓	
DFL	JBHI	✓		✓				✓		✓								✓	
PTGHRA	JBHI	✓						✓		✓								✓	
DoctorAI	PMLR		✓					✓											✓
TaGiTeD	AAAI	✓		✓			✓						✓						✓
Patient2Vec	IEEE	✓		✓		✓		✓		✓									✓
CONTENT	SciRep		✓					✓		✓									✓
MedCSP	ACL	✓		✓		✓		✓		✓									✓

3.2.2. *Handling imbalanced data*

Data imbalance in healthcare datasets, exacerbated by long-tailed distributions, noise, and rare event detection challenges, leads to biased predictions and reduced model performance [193]. Solutions like oversampling, synthetic data generation, and advanced architectures address these issues.

Several models tackle these challenges. SeqCare [194] balances class representations using self supervised graph contrastive learning. StratMed [195] improves learning for underrepresented entities with relevance stratification and dual-property networks. SASMOTE [196] enhances synthetic sample quality by adapting neighbor selection and reducing uncertainty. NNBRM [197] generates realistic synthetic samples to improve sensitivity to rare cases via a back-propagation neural network (BPNN). GAN-based methods further improve data fidelity. MedGAN [198] generates synthetic medical images while preserving structural and textural integrity. SynTEG [199] maintains temporal correlations and visit patterns in synthetic EHRs. MedDiffusion [200] captures temporal dependencies in patient sequences using a Denoising Diffusion Probabilistic Model (DDPM) with step-wise attention. PromptEHR [201] generates realistic synthetic EHRs by conditioning on patient features and integrating privacy measures. These models not only address class imbalance but also tackle noise, long-tailed distributions, and privacy concerns. SeqCare refines graph structures to exclude irrelevant information, StratMed strengthens learning for sparse data, and SASMOTE and NNBRM ensure class balance without artifacts. GAN-based solutions guarantee structural and temporal coherence, while PromptEHR facilitates multimodal data generation.

3.2.3. *Multimodal and missing modality*

Integrating multimodal data is challenging due to differences in structure, distribution, and feature representations, with missing modalities adding further complexity. To address this, approaches like multimodal deep learning, multi-task learning, and imputation strategies are commonly used.

MNN [202] integrates clinical notes and medical codes using attention-based bidirectional RNNs, enhancing prediction accuracy. Health-ATM [203] combines RNNs, CNNs, and time-aware attention to process noisy, incomplete data. PREMIER [204] uses a graph-based framework and recommender system for safe, personalized medication recommendations, while 4SDrug [74] improves symptom-drug interaction safety through set-oriented representations. More advanced models focus on missing modalities and consistency. MUSE [205] applies graph contrastive learning to model patient-modality relationships, while DrFuse [206] employs disentangled representations and disease-specific attention to enhance robustness. M3Care [207] ensures task-specific imputation through latent space alignment and kernel similarity metrics. FlexCare [208], a Hierarchical Transformer-based model, leverages attention mechanisms and hierarchical propagation to capture complex temporal dependencies in irregular EHR time series data, effectively addressing time series processing challenges and enhancing clinical risk prediction. FuseMoE [209] employs a mixture-of-experts transformer architecture — with sparse MoE fusion layers and an innovative Laplace gating function — to effectively fuse FlexiModal data, thereby addressing challenges of missing, irregular, and heterogeneous modalities for improved clinical predictive performance. Flex-MoE [210] employs a flexible mixture-of-experts framework that integrates a learnable missing modality bank with a sparse MoE layer and Transformer-based architecture to effectively fuse arbitrary combinations of modalities, thereby addressing the challenge of missing and incomplete data in multimodal medical settings. In addition to handling missing data, these models address secondary challenges: Health-ATM captures temporal dynamics, PREMIER and 4SDrug prioritize safety, MUSE mitigates modality collapse, and DrFuse enhances consistency. Collectively, these methods exemplify significant progress in integrating multimodal data for healthcare predictive modeling.

3.3. Time series data processing

Processing time series data in EHRs is crucial for tracking patient health and aiding clinical decisions. However, challenges such as irregular visit intervals, sparse entries, and missing values must be addressed for reliable modeling and accurate predictions. Techniques like memory-aware models, attention mechanisms, and knowledge-infused frameworks have improved the robustness and scalability of patient trajectory modeling.

3.3.1. Patient trajectories modeling

Patient trajectory modeling in EHRs faces challenges such as irregular data intervals, sparse entries, and integrating heterogeneous data. Advanced models offer solutions, such as the Attentive State–Space model [211], which improves temporal modeling by handling sparse data with memoryful dynamics. CAMP [212] and ICU-Sim [213] utilize co-attention, memory networks, and graph-based learning to capture long-term dependencies. PROMISE [167] incorporates multimodal data and pre-training to manage noisy and incomplete data. DNA-T [214] is a novel deformable neighborhood attention transformer that dynamically adjusts its receptive field using missing patterns to flexibly aggregate local features from irregular medical time series, thereby improving the prediction of patient mortality risk. MPRE [215] utilizes wavelet-based frequency transformation to extract multi-scale trend and variation information from dynamic EHR features, and then employs a 2D multi-extraction network together with a first-order difference attention mechanism to capture the correlations between these signals. TRANS [216] presents a novel approach that models EHRs as a temporal heterogeneous graph and employs a temporal Graph Transformer to integrate temporal dynamics and structured relationships among medical events, thereby improving diagnosis prediction accuracy.

Despite improvements, challenges remain. Real-time analysis is crucial for clinical interventions, requiring scalable architectures like CAMP's temporal encoding [212] and PROMISE's modular integration [167]. Scalability is key for handling large EHR datasets, ensuring efficiency as data volumes grow. Addressing missing or noisy data remains a priority, with methods like robust data imputation (Attentive State–Space model [211]) and auxiliary data integration (PROMISE [167]) enhancing reliability.

3.3.2. Visit interval and data imputation

Irregular visit intervals and missing values are common in time series data, complicating predictive modeling. To address this, imputation methods — from basic interpolation (*e.g.*, linear, spline) to advanced machine learning techniques — are essential for improving data reliability. Models like ATTAIN and HiTANet tackle these challenges with time-aware mechanisms. ATTAIN [217] captures temporal dependencies by encoding time gaps between events and applying attention within its LSTM architecture to focus on key past events, enhancing accuracy and interpretability. HiTANet [218] enhances this approach with a hierarchical attention mechanism, using local-level Transformers and global key-query attention to adjust the influence of past events based on temporal relevance. Both models excel in predicting patient trajectories and clinical outcomes. Zhang et al. [174] tackles the irregularity in both multivariate time series and clinical note sequences by dynamically integrating handcrafted imputation embeddings with learned interpolation and fusing them with interleaved self- and cross-attention mechanisms, thereby enhancing predictions such as in-hospital mortality and phenotype classification. Warpformer [219] introduces a multi-scale modeling framework that leverages an adaptive warping module and a customized doubly self-attention mechanism to unify irregular clinical time series at various granularities, effectively addressing both intra-series irregularity and inter-series discrepancy for improved clinical predictions. These models also enhance robustness: ATTAIN uses advanced imputation to handle missing data, while HiTANet's attention mechanism captures event impacts across multiple time scales, providing a deeper understanding of disease progression. Together, they set new benchmarks for predictive modeling in clinical settings, addressing the complexities of EHR data.

3.4. Hidden or higher complex relationship

Medical data involve complex, nonlinear dependencies across temporal, relational, and heterogeneous dimensions, making them difficult to model. Advanced techniques like CNNs, RNNs, and Variational Autoencoders (VAEs) are effective at capturing these patterns. Attention Mechanisms and GNNs further enhance the ability to model structured dependencies and uncover hidden insights within EHRs. Addressing these complexities requires tailored strategies in four key areas: temporal dynamics, complex relationships, heterogeneous data integration, and high-dimensional feature extraction.

For temporal dynamics, time-aware models, T-LSTM [220] capture short-term and long-term dependencies, while RAIM [169] integrates continuous monitoring data with discrete clinical events through attention mechanisms. SETOR [221] incorporates medical ontologies and attention-based graph representations to improve sequential predictions. Capturing complex relationships is critical. GCT [222] combines transformers and graph convolutional networks, and Chet [223] models disease progression using dynamic disease graphs and neural ODEs to capture the correlation. For heterogeneous data integration, MCA-RNN [224] improves diagnosis accuracy by incorporating patient-specific information with CVAEs, while DASNet [225] captures intra- and inter-correlations in heterogeneous temporal sequences through attention mechanisms. In high-dimensional feature extraction, MEGACare [80] enhances features using multi-view hypergraphs and medical knowledge, while CTCL [164] integrates diverse data types through cross-modal contrastive learning and multi-view GCNs, employing cross-attention to improve outcome predictions.

3.5. Domain knowledge utilization

Integrating domain knowledge into healthcare models improves accuracy, interpretability, and reliability by guiding feature selection, model design, and interpretation. Pharmacological insights, for example, help focus on relevant features in drug prediction, enhancing the understanding of complex interactions. Additionally, structured representations like knowledge graphs and ontologies systematize medical knowledge, improving reasoning and prediction. This integration enables transparency, trust among clinicians, and better decision-making.

However, the challenge remains in effectively combining diverse data sources to capture complex relationships, reduce data sparsity, and maintain efficiency. Models such as GAMENet [73] and G-BERT [226] combine EHRs with graph-based embeddings using dual-RNNs and GNNs with transformers, respectively. MMORE [83] and UNITE [85] integrate multiple ontologies and data sources, while CompNet [227] uses reinforcement learning to predict drug combinations. SafeDrug [19] integrates molecular structures, and MK-GNN [228] and OntoPath [229] model dependencies hierarchically. MoleRec [82] and CSEDrug [230] focus on molecular substructures for drug recommendations, and C-MemNNs [231] and SATexMCE [232] enhance diagnostic inference and text integration. KSI [233] and DPGNN [75] use graph-based methods to integrate online knowledge for disease prediction.

Addressing biases in medical data integration is another significant challenge. DrugRec [234] applies causal inference and longitudinal data for personalized treatment, while Carmen [235] combines molecular and DDI graphs with patient history for context-aware recommendations. Fine-grained interaction modeling is tackled by REFINE [236], which uses deep learning for personalized drug interaction modeling, in contrast to Carmen's graph-based method. Temporal event prediction, as seen in CGL [76], integrates text and graph learning, while HAP [84] uses hierarchical attention for healthcare representation learning, offering a distinct approach from MMORE's ontological focus.

3.6. Model optimization and improvement

Model optimization is essential for overcoming challenges in EHR analysis, such as temporal dynamics, data imbalance, heterogeneous integration, and adapting predictions to new patient scenarios. This section discusses three main areas: pre-training and improved loss functions, addressing cold-start issues, and achieving personalization. These strategies enhance the robustness, adaptability, and accuracy of predictive models.

3.6.1. Pre-training and enhanced loss function

Pre-training and enhanced loss functions play a key role in addressing challenges in EHR analysis, such as temporal dynamics, data imbalance, heterogeneous data integration, and robustness. By leveraging large-scale datasets, pre-training helps initialize model parameters, enhancing generalization to downstream tasks. Meanwhile, enhanced loss functions improve stability and prediction accuracy by tackling noisy, imbalanced data through class weighting and contrastive learning.

For temporal dynamics, which often involve irregular visit intervals and short sequences, RAPT [237] uses a time-aware transformer with pre-training tasks like similarity and masked prediction to capture robust temporal representations. Its composite loss function improves generalization across various applications. In dealing with class imbalance, SCEHR [159] applies supervised contrastive learning to maximize inter-class separation and minimize intra-class variation, enhancing robustness and accuracy by weighting imbalanced classes. MedCSP [238] is a collaborative pre-training framework that employs modality-specific aggregation and cross-source contrastive learning to unify multimodal medical data, thereby alleviating data scarcity and enhancing the model's applicability across diverse downstream tasks. Med-ST [239] is a multimodal pre-training framework that employs a Mixture-of-View Experts (MoVE) architecture, modality-weighted local alignment, and bidirectional cycle consistency to capture fine-grained spatial and temporal features from chest radiographs and radiological reports, addressing the challenge of modeling both multi-view spatial details and temporal dynamics. To integrate structured and unstructured data, MedRetriever [77] uses an attention-based retrieval mechanism, aligning pre-training with specific prediction tasks. Its composite loss function optimizes both text retrieval and health risk prediction. Lastly, BEHRT [240] and Med-BERT [241] showcase the power of large-scale pre-training, with BEHRT modeling temporal sequences using positional encodings and self-attention, and Med-BERT leveraging 28.5 million patient records for robust contextual embeddings. Both models utilize composite loss functions to handle multi-task predictions, achieving strong performance even with limited fine-tuning data.

3.6.2. Cold-start problems

Cold-start problems occur when predicting outcomes for new patients with limited or no historical data, especially in rare diseases. To address this, transfer learning and multi-task learning have been widely adopted to improve model adaptability in data-scarce contexts.

GraphRet [242] and MetaCare [243] provide effective solutions. GraphRet utilizes inductive graph convolutional networks and heterogeneous graphs to model relationships between diverse entities, with meta-learning enabling rapid adaptation to new patients and handling sparse data. MetaCare, through hierarchical subtyping, clusters patients based on shared traits, facilitating knowledge transfer specific to subtypes, while fine-tuning adjusts predictions for individual patients even with minimal records. Both models go beyond cold-start issues. GraphRet improves prediction accuracy by integrating heterogeneous data via relational graphs, and MetaCare enhances robustness by leveraging shared patterns across subpopulations through hierarchical subtyping.

3.6.3. Personalization

Personalization in healthcare prediction tailors recommendations based on individual characteristics such as genetic data, lifestyle, and medical history, enhancing treatment planning and clinical decisions. This approach addresses the heterogeneity and dynamics of clinical data, leading to more accurate, context-aware predictions and improved patient outcomes. However, managing data heterogeneity and capturing dynamic, interdependent clinical features remain key challenges.

Models like ConCare [157], MedPath [244], and ProCare [245] address these challenges using complementary strategies. ConCare employs a time-aware attention mechanism and multi-channel GRU to manage temporal irregularities and feature correlation. MedPath incorporates medical knowledge graphs and path reasoning to contextualize electronic medical records (EMRs) and capture complex relationships. ProCare focuses on disease progression, using pathway analysis and dynamic embeddings to model disease trajectories over time. Despite their differences, these models share the common goal of personalizing predictions through attention mechanisms and external knowledge. ProCare's dynamic modeling complements MedPath's contextual reasoning, while ConCare's temporal embeddings address irregular time-series data. By capturing individual characteristics, these models enhance predictive accuracy. Future work should explore multimodal data integration and real-time adaptation to further advance personalized medicine.

3.7. Model interpretability and trustworthiness

Machine learning models in healthcare must be interpretable, trustworthy, and capable of estimating uncertainty to support reliable clinical decision-making. Transparent predictions help clinicians understand the rationale behind model outputs, while uncertainty estimation quantifies confidence in predictions, guiding diagnosis and treatment. Addressing these challenges requires methods that integrate domain knowledge, manage complex data, and provide robust, actionable insights.

3.7.1. Interpretability and trustworthiness

Interpretability and trustworthiness are key for clinical adoption. Transparent models align predictions with clinical reasoning, making decisions understandable for clinicians and patients. Models like RETAIN [246] and GRAM [187] use attention mechanisms and medical ontologies to highlight relevant features, while DG-RNN [18] integrates domain-specific knowledge graphs to improve interpretability.

To address temporal complexity and irregular data in EHRs, models like ATTAIN [217] and AdaCare [70] employ time-aware attention and dynamic recalibration, while KerPrint [247] incorporates knowledge graphs to capture temporal dynamics and reduce information decay. For heterogeneous data integration, Sherbet [248] and LDAM [249] combine unstructured text with time-series data, while KnowRisk [192] and MedRetriever [77] leverage domain knowledge and visualization tools to enhance interpretability. These models enhance transparency and robustness by integrating attention mechanisms, domain knowledge, and adaptive data handling, supporting clinically meaningful predictions.

Table 5
Summary of recent work on clinical language models. In the **Task** column, **NER** refers to Named Entity Recognition, **RE** refers to Heart Failure, **QA** refers to Question Answering, **DC** refers to Document Classification, and **Pre.** refers to Clinical Prediction.

Models	DataSet			Task						Unique innovations and limitations
	MIMIC	PubMed	Others	NER	RE	QA	DC	Pre	Others	
AlphaBERT			✓						✓	✓ Character-level tokens, optimized for low-resource environments ✗ Limited performance on complex tasks, high domain dependency
BERT-MIMIC	✓			✓						✓ Sentence-agnostic temporal relation extraction, silver data augmentation ✗ Focuses on temporal tasks, limited generalizability
BERT-XML			✓					✓		✓ Multi-label attention for ICD coding, domain-specific vocabulary ✗ Primarily for ICD coding, lacks task versatility
BioBERT									✓	✓ Biomedical pretraining, optimized for NER, RE, QA tasks ✗ Requires domain-specific corpora, limited generalization
BlueBERT	✓	✓		✓	✓		✓		✓	✓ Domain transfer learning on PubMed and MIMIC datasets ✗ Less effective for clinical-specific tasks
ClinicalBERT	✓							✓	✓	✓ Clinical note representation, focus on 30-day readmission prediction ✗ Relies on MIMIC dataset, high computational complexity
Clinical KB-ALBERT	✓	✓		✓					✓	✓ Joint training with UMLS knowledge, enhanced semantic understanding ✗ Complex training, dependent on UMLS knowledge base
Clinical-longformer	✓			✓		✓		✓		✓ Sparse attention for long sequences, efficient memory usage ✗ Limited improvements for short-text tasks, high training complexity
Clinical-T5	✓			✓		✓			✓	✓ Decoder architecture for clinical text generation tasks ✗ Restricted access due to privacy, resource-intensive
Clinical XLNet	✓							✓		✓ Temporal sequence modeling for clinical notes ✗ Narrow application scope, lacks versatility
CharacterBERT	✓	✓		✓					✓	✓ Character-level CNN for open vocabulary, robust word representation ✗ Increased complexity for character-level processing
CXR-BERT	✓	✓						✓	✓	✓ Self-supervised contrastive learning for vision–language tasks ✗ Limited to radiology benchmarks, data-intensive
EHR-BERT		✓	✓						✓	✓ Sequential masked token prediction for anomaly detection in EHRs ✗ Dependent on EHR quality, limited modality support
RadBERT			✓				✓			✓ Radiology-specific language adaptation for NLP tasks ✗ Domain-specific scope, requires substantial annotated data
UMLSBERT	✓	✓		✓				✓	✓	✓ UMLS knowledge augmentation for improved contextual embeddings ✗ Highly domain-dependent, complex training process

3.7.2. *Uncertainty estimate*

Uncertainty estimation quantifies prediction confidence, enhancing decision-making in ambiguous situations. Models like Bayesian neural networks and deep ensembles offer confidence intervals, supporting anomaly detection and robustness. However, challenges such as data imbalance, noisy inputs, and data heterogeneity remain. Advanced models, such as CONAN [250], UNITE [85], SeqCare [194], and TrustSleepNet [251], address these issues with innovative techniques.

CONAN improves rare disease detection by using GAN-generated balanced samples to enhance sensitivity. UNITE integrates multimodal data with stochastic variational inference for robust uncertainty estimation. SeqCare employs medical knowledge graphs and adaptive fusion to manage data variability, while TrustSleepNet combines CNNs and attention mechanisms to provide reliable predictions in noisy environments. These models share common strategies for handling medical prediction challenges. CONAN and UNITE integrate diverse data using hierarchical embeddings and multimodal fusion, while SeqCare and TrustSleepNet use graph-based frameworks and attention mechanisms to manage complex inputs. For interpretability, SeqCare and TrustSleepNet align predictions with clinically relevant features through knowledge graphs and attention mechanisms. Regarding noisy data, CONAN leverages GANs to generate robust representations, and TrustSleepNet's CNN-based approach ensures reliable feature extraction, maintaining accuracy despite irregularities. These techniques collectively improve the robustness, interpretability, and adaptability of healthcare models.

4. Multimodal and LLMs

The previous review highlighted key research directions in medical data processing, focusing on task categorization and motivational challenges. Task differences have led to diverse methodologies and applications, while motivational challenges have driven technological innovations. However, all tasks share a common foundation in EHR data, which plays a crucial role in improving clinical decision-making and enabling personalized treatment. As large-scale EHR datasets, especially when combined with multi-omics data, continue to grow, foundational models like Clinical Language Models and EHRs FMs are gaining prominence. These models are critical for precision medicine, as they help extract actionable insights from complex healthcare data, enhancing the accuracy and quality of medical decisions.

4.1. *EHR-centric clinical foundation models*

4.1.1. *Clinical foundational models*

Foundational Models (FM) represent a shift in machine learning from task-specific models to pre-trained models that can be adapted with minimal fine-tuning to a range of tasks [33,252]. These models, typically trained on large, unlabeled datasets, show strong potential in fields like natural language processing (NLP) and clinical data analysis [34,253]. A prominent example of FM is the LLM, such as ChatGPT [254,255]. Initially designed for word prediction, LLMs now perform complex tasks, including passing medical exams, generating

Table 6
Summary of Recent Work on Foundation Models. In the **Task** column, **Mor.** refers to Mortality Prediction, **HF.** refers to Heart Failure, **LOS** refers to Length of Stay, **Read.** refers to Readmission, **Diag.** refers to Diagnosis Prediction, and **Tre.** refers to Treatment Prediction.

Models	DataSet		Task							Unique innovations and limitations
	Public	Private	Mor.	HF.	LOS	Read.	Diag.	Tre.	Others	
BEHRT	✓						✓			✓ Multimodal input; personalized prediction; scalable ✗ Limited for long sequence modeling
Med-BERT	✓			✓					✓	✓ Contextual embeddings boost performance on small datasets ✗ Lacks support for unstructured EHR data
CEHR-BERT		✓	✓			✓			✓	✓ Incorporates temporal data with time/age embeddings ✗ Complex architecture; long training time
Hi-BEHRT	✓			✓					✓	✓ Hierarchical transformer for long sequence modeling ✗ High computational cost
AdaDiag	✓	✓		✓						✓ Adversarial domain adaptation for cross-domain generalization ✗ Increased training complexity due to adversarial adaptation
DOCTORAI		✓					✓	✓		✓ RNN-based, general-purpose temporal model ✗ Limited flexibility compared to modern architectures
RETAIN	✓	✓		✓			✓	✓		✓ Reverse attention for interpretability and clinical relevance ✗ Slightly lower accuracy compared to complex DL models
GRAM	✓	✓		✓			✓			✓ Medical ontology alignment for sparse data ✗ High dependence on medical ontologies
CLMBR		✓		✓	✓	✓			✓	✓ NLP-inspired patient representation schemes ✗ Limited support for multimodal data
Patient2Vec		✓							✓	✓ Personalized interpretable embeddings ✗ Complex data preprocessing
DescEmb	✓		✓		✓	✓	✓			✓ Text-based code embedding to unify heterogeneous EHR systems ✗ Limited applicability to structured data
MixEHR	✓	✓	✓				✓		✓	✓ Multi-view Bayesian topic model ✗ May struggle with highly noisy data
UniHPF	✓		✓		✓	✓	✓			✓ Zero domain knowledge requirement ✗ Generalized results may lack domain specificity
GenHPF	✓	✓	✓		✓	✓	✓	✓		✓ Effective multi-task, multi-source learning ✗ Pretraining-dependent; high computational needs
HAIM	✓		✓		✓					✓ Multimodal integration with Shapley value quantification ✗ Complex multimodal fusion methods
MedGPT	✓	✓					✓			✓ Utilizes free-text EHRs for detailed prediction ✗ Sensitive to noise in text data

radiology reports, and drafting research papers. While excelling in text processing, these models have been extended to structured EHR data, enabling clinical applications like predicting readmission risks, recommending treatments, and diagnosing rare diseases.

Clinical foundational models can be divided into two types: Clinical Language Models (CLMs) and Electronic Health Record Foundational Models (EHR-FMs). CLMs (as shown in Table 5), specialized LLMs trained on clinical texts such as physician notes and PubMed articles, excel in tasks like drug name extraction and clinical entity recognition. While CLMs outperform general-purpose LLMs in their domain, their reliance on limited datasets like MIMIC-III and PubMed poses challenges in knowledge coverage and updates. EHR-FMs (as shown in Table 6) transform longitudinal patient data into high-dimensional embeddings, supporting tasks such as readmission prediction and length-of-stay estimation. However, their reliance on structured data, like diagnostic codes, limits their generalizability across diverse healthcare systems.

4.1.2. Current challenges of clinical FMs

(I) Data dependence and coverage gaps. The performance of clinical foundational models, especially those based on EHR, depends on the quality and scope of training data. General models like BioBERT [99] and BlueBERT [256], pre-trained on large biomedical corpora such as PubMed and MIMIC-III, offer strong multi-task adaptability and are suitable for a broad range of biomedical NLP tasks. However, they face limitations in rare diseases and up-to-date medical knowledge, as MIMIC-III only includes data until 2012. In contrast, specialized models like ClinicalBERT [257] and RadBERT [258], trained on clinical notes and radiology reports, excel at specific tasks like readmission risk prediction and report summarization, but struggle to generalize across domains. Additionally, models like Clinical KB-ALBERT [259] and EHR-BERT [260], which integrate domain-specific knowledge, perform well in tasks such as named entity recognition and anomaly detection, but are limited to narrow applications. This highlights the challenge of balancing the versatility of general models with the precision of domain-specific ones.

(II) Task sparsity and lack of cross-task comparability. Models like Clinical KB-ALBERT [259] and EHR-BERT [260] excel in tasks such as named entity recognition (NER) and anomaly detection, but their performance is mainly assessed on isolated tasks, complicating cross-task comparisons. The absence of standardized evaluation benchmarks further limits comparisons. For example, Hi-BEHRT [261] uses hierarchical transformers for long time-series data in readmission prediction, whereas RETAIN [246] applies reverse-time attention mechanisms for interpretability. These methodological differences hinder direct comparisons and obscure the relative performance of models across clinical tasks.

(III) Interpretability and trust challenges. Despite the adoption of attention mechanisms and graph-based structures in models like RETAIN [246] and GRAM [187], many clinical models still lack interpretability, limiting their practical use, especially in critical areas like disease diagnosis where understanding model predictions is essential. For instance, EHR-BERT [260] excels in anomaly detection but provides no clear explanation for its predictions, making it difficult for clinicians to trust its recommendations. Although RETAIN [246] improves interpretability with reverse-time attention, its reliance on complex domain knowledge and architecture limits its flexibility and generalizability.

(IV) Lack of standardized evaluation benchmarks. The lack of standardized evaluation benchmarks is a key challenge in comparing the performance of clinical foundational models. Models like BEHRT [240] and Med-BERT [241] are often tested on customized tasks with varying

definitions, making cross-study comparisons difficult. For example, binary classification tasks like mortality prediction can differ significantly due to variations in sample selection and feature engineering. This inconsistency in task definitions and evaluation criteria complicates benchmarking. Standardized evaluation frameworks, including widely used public datasets like MIMIC-IV, are essential to improve reproducibility, comparability, and transparency in research.

(V) Resource efficiency and scalability issues. Resource efficiency is crucial for deploying clinical foundational models in low-resource healthcare settings. Models like Clinical-Longformer [262] use sparse attention mechanisms to reduce computational complexity but still demand significant resources, limiting their use in smaller healthcare institutions. Similarly, multimodal models such as HAIM [263] and GenHPF [264], which integrate text, tabular, and imaging data, offer higher performance but are resource-intensive. Optimizing models for efficiency and scalability is essential, especially in resource-limited environments.

4.2. Integrating EHRs and multi-omics data

The rise of EHRs and high-throughput technologies has paved the way for precision medicine. By combining EHR and multi-omics data, comprehensive analyses can optimize clinical practices with tools like polygenic risk scores and big data analytics, improving patient care and supporting personalized solutions. Advanced integration methods align clinical and omics data, revealing disease mechanisms and enabling evidence-based decisions. Biobanks are essential for integrating EHR and omics data, offering valuable datasets for rare diseases and longitudinal studies. For instance, the UK Biobank merges genomic, imaging, and EHR data to explore the links between genetic variants and clinical outcomes [265]. International initiatives like BBMRI-ERIC foster data sharing and collaboration, enhancing biobank utility for biomedical research [266,267].

The integration of EHRs with multi-omics data is crucial for advancing disease understanding and precision medicine. This integration is driven by two complementary strategies: the *Phenotype-First Strategy* and the *Genotype-First Strategy*.

The *Phenotype-First Strategy* begins with clinical disease traits to define study cohorts, which are then used to investigate genetic factors. This strategy excels in disease characterization and is particularly effective in leveraging EHR-linked biobanks. Ritchie et al. [268] used population data to reduce recruitment costs, demonstrating the power of biobank integration. Furthermore, Veturi et al. [269] utilized U.K. Biobank data to identify lipid-associated genes and their pleiotropic effects. The strategy is also effective for rare diseases, as shown by Pena et al. [270], who identified causal genes. Additionally, advanced methods like those discussed by Zhao et al. [271] enhance phenotype definition and improve disease prediction. This strategy also aids in therapeutic discovery, as demonstrated by Lessard et al. [272], who integrated variant annotations to prioritize targets, and Su et al. [273], who linked multi-omics and EHR data to uncover immune changes in COVID-19.

In contrast, the *Genotype-First Strategy* starts with genetic variants and links them to clinical phenotypes, uncovering pleiotropic effects. This approach provides deep insights into genetic mechanisms and is valuable for both rare and common diseases. Park et al. [274] used U.K. Biobank data to study rare LMNA gene variants, uncovering novel phenotype associations. Drivas et al. [275] explored the role of ciliopathy genes in common diseases, while Guo et al. [276] integrated genotype–phenotype data in small-sample multi-omics studies. Wu et al. [277] introduced an AI-driven framework to predict causal relationships, deepening our understanding of complex traits. These studies illustrate the value of the genotype-first strategy in advancing precision medicine, as highlighted by Park et al. [278]. Both strategies, supported by biobanks, hold great promise for personalized medicine. They enable precise analyses, improve patient care, and contribute to the development of innovative treatments in precision healthcare.

5. Challenges and future directions

Medical prediction holds great promise for improving patient outcomes and advancing precision medicine. Recent studies have explored methods to enhance prognostic, diagnostic, and therapeutic outcomes in clinical practice. However, challenges remain in their practical implementation. This section discusses the challenges and opportunities in integrating multimodal medical data and applying advanced computational methods for precision medicine.

5.1. Data quality and consistency

Current intelligent healthcare models are data-centric [42], but variations in record-keeping across institutions increase EHR data heterogeneity, making integration and standardization difficult, thus limiting model generalization [44,279]. These discrepancies lead to challenges such as missing values, noise, and inconsistent visit intervals [280,281], affecting model accuracy and reliability.

Future research should focus on improving data collection to create high-quality, diverse datasets that enable the development of flexible, context-adaptive models by integrating multimodal data. Efforts should be directed toward aligning data with real-world contexts, constructing user-oriented datasets, and ensuring fine-grained alignment with domain-specific knowledge. Additionally, advancements in data cleaning and missing value imputation using machine learning techniques are needed to automatically detect and correct data errors. Novel time series analysis and interpolation methods should address irregular visit intervals, while standardizing data formats and coding systems is crucial to enhance consistency and interoperability, enabling the development of more robust, widely applicable healthcare models.

5.2. Multimodal data integration and higher-order relationship capture

The integration of diverse modalities in EHRs is complex due to the differences between structured and unstructured data, requiring advanced processing techniques and model architectures to capture implicit relationships. Current methods mainly focus on physiological signals [282], radiological, and pathological images [283], but future approaches should also incorporate medical audio and video data for better diagnosis and intervention [284]. Combining EHR with genomic and omics data could reveal disease–gene associations and molecular mechanisms of comorbidities. Additionally, integrating medical knowledge graphs, like UMLS and PrimeKG [285], into models remains a key challenge.

To address these issues, future research should focus on developing multimodal data integration technologies and capturing higher-order relationships [286]. This includes creating new data fusion techniques for heterogeneous data and improving natural language processing (NLP) tools, especially large pretrained language models (PLMs) for medical texts. Knowledge-enhanced strategies can improve data consistency and the model's ability to process clinical information. Specialized algorithms are needed to integrate genomic data with EHR, addressing differences in data formats and accurately capturing their interrelationships. Current methods that convert knowledge graphs into token sequences often lose information, but learning universal semantic tokens [281,287] could facilitate more efficient knowledge fusion.

5.3. Integrating genomic data into clinical workflows

Developing advanced analytics to extract actionable insights from EHR and multi-omic datasets remains a formidable challenge. Although recent computational advances have enabled routine interrogation of genomic data, issues such as data volume, sparsity, and high dimensionality continue to hinder traditional statistical approaches [288,289]. Specifically, the curse of dimensionality — where the number of features far exceeds available patient samples — often results in ill-conditioned feature matrices, complicating the integration of multimodal data. Moreover, reliably linking genotype to phenotype remains elusive, impeding the translation of genomic discoveries into clinical practice [290].

Future research should focus on developing robust dimensionality reduction and feature selection techniques to filter out irrelevant variants effectively. The increasing availability of large datasets presents a promising avenue for constructing clinical decision support systems and personalizing treatment, thereby enhancing the establishment of genetic links [291]. Innovative methodologies, such as the Base Pair Network introduced by Avsec et al. [292], offer potential by improving model interpretability through tracing output signals back to input sequences to reveal specific sequence motifs. Advancements in computational strategies and data integration techniques are essential for bridging the gap between genomic discoveries and the practical implementation of precision medicine.

5.4. Data privacy and security

EHR data contains sensitive personal health information, necessitating robust privacy and security measures. However, stringent privacy regulations hinder data sharing, limiting cross-institutional collaboration and model accuracy. Federated Learning (FL) offers a solution by allowing medical institutions to train models locally and share only model parameters, rather than raw data [293]. A central server aggregates these updates, enhancing model generalization and reducing the risk of data leakage [294–296].

Future research should focus on optimizing FL methods to improve efficiency and security. This includes developing more efficient algorithms to reduce communication and computational costs [297], integrating differential privacy techniques to protect against data reconstruction, and standardizing data formats to improve interoperability while ensuring privacy.

5.5. Model generalization and robustness

Generalization and robustness are key challenges for EHR clinical prediction models. Models often perform well within specific hospitals or patient groups but struggle to maintain performance across diverse settings due to differences in overfitting and annotation paradigms [298]. EHR data may also suffer from selection bias, limiting coverage, especially for rare diseases, and hindering generalization. Additionally, model stability and reliability under varying data distributions or noise are critical for practical adaptability.

To improve generalization and robustness, unified models that integrate various medical modalities and data types are essential. Techniques like domain adaptation and domain generalization help enhance stability across diverse data conditions, while methods such as data augmentation and transfer learning improve performance, particularly in data-scarce scenarios. Foundational models hold significant untapped potential. Research has shown that slight changes in natural language prompts can notably affect model performance, leading to various prompt strategies like automatic, soft, and pre-trained prompts [299,300]. The Chain-of-Thought (CoT) method improves transparency and performance [301,302], while alternative methods like Tree-of-Thought (ToT) and Graph-of-Thought (GoT) offer new optimization approaches [303,304]. Retrieval-augmented methods, which introduce external knowledge, have also shown promise [305]. Integrating these techniques with medical domains could unlock the full potential of foundational models in clinical applications.

5.6. Real-time prediction and interpretability

Real-time prediction and model interpretability are critical challenges in clinical applications. Clinicians require real-time or near-real-time predictions to support decision-making, which demands high computational efficiency [306]. However, traditional deep learning and complex machine learning models often require significant computational resources and time, which may not be feasible in clinical settings. Additionally, model interpretability is a key issue, as clinicians must understand and trust the model's predictions for clinical decision-making [307]. Many advanced models, such as deep neural networks, are considered “black boxes”, with complex internal mechanisms that are difficult to interpret, limiting their clinical applicability.

To address these challenges, future research should focus on improving both real-time prediction efficiency and model interpretability. For real-time prediction, optimizing and accelerating existing models is essential. Techniques such as pruning, quantization, and knowledge distillation can reduce computational complexity and enhance inference speed [308,309]. Additionally, more efficient Transformer variants, such as Longformer and Sparse Transformer, utilize sparse attention mechanisms to reduce computation and memory usage, enabling faster processing of long sequences. Moreover, medical language models can decompose problems into understandable sub-tasks, improving interpretability. Integrating medical expertise further enhances model transparency, helping clinicians understand the reasoning behind model decisions and reducing reliance on “black-box” models.

5.7. Health assessment

Current evaluation methods for predictive models rely on prompt engineering and benchmark datasets, where variations in prompts can lead to significant result differences. Domain-specific datasets in expert systems can introduce uncertainty, as models may encounter evaluation data during training, causing bias and unfair comparisons [310,311]. In healthcare, inaccurate predictions can have serious consequences, highlighting the need for reliable models. However, assessing model trustworthiness remains a challenge.

To address these issues, future efforts should focus on four key areas. First, trustworthiness assessment is critical, requiring advanced tools to identify and correct errors, ensuring reliable model outputs and supporting clinicians with trust scores and alerts. Second, comprehensive, multitask evaluation should assess models across diagnostic, prognostic, and treatment tasks to provide a fuller understanding of their capabilities. Third, multidimensional evaluation should consider interpretability, robustness, hallucination rates, redundancy, and bias, with new metrics developed to reflect real-world clinical performance. Finally, fair and standardized protocols are necessary, as current metrics fail for open-ended tasks, and human evaluators may introduce bias. Standardized, AI-based automated frameworks are crucial for robust, objective assessments.

5.8. Ethical and legal issues

Algorithmic bias and legal responsibility are key challenges in healthcare AI. Bias in EHR data can worsen disparities, leading to unfair treatment, especially when certain patient groups are under- or overrepresented. Legal concerns also arise when AI recommendations cause harm, raising questions of accountability in automated decision-making. **Future research should focus on** using fairness algorithms, such as reweighting and resampling, to address data bias and ensure equity across patient groups, and establishing clear legal frameworks that define responsibilities for developers, healthcare institutions, and users, ensuring patient consent, data privacy, and model transparency and accountability.

6. Conclusion

In summary, the digitization of healthcare information through EHR systems has enabled significant advancements in predictive modeling within the medical field. Our review has highlighted the recent progress in applying deep learning techniques to EHR data, particularly in patient representation learning and clinical outcome prediction. We have categorized various deep learning models, emphasizing their advancements and applications while identifying key challenges such as data heterogeneity, model interpretability, and data imbalance. Despite these challenges, the potential for deep learning to transform healthcare is evident, with improvements in diagnosis accuracy and personalized treatment plans. Moving forward, it is essential to enhance model transparency, generalizability, and adaptability to fully leverage AI's benefits in healthcare. By addressing these challenges, we can significantly advance medical informatics and improve healthcare outcomes. Looking ahead, continued innovation and interdisciplinary collaboration will be crucial in overcoming existing hurdles and unlocking the full potential of deep learning in healthcare. This will pave the way for more accurate, efficient, and personalized medical care, ultimately improving patient outcomes and advancing the field of medical informatics.

CRediT authorship contribution statement

Jialun Wu: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Kai He:** Supervision, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Rui Mao:** Validation, Methodology, Investigation, Data curation, Conceptualization. **Xuequn Shang:** Writing – review & editing, Validation, Supervision, Funding acquisition, Conceptualization. **Erik Cambria:** Writing – review & editing, Supervision, Project administration, Methodology, Formal analysis.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Kai He reports financial support was provided by National Research Foundation Singapore under Al Singapore Programme. Kai He reports financial support was provided by RIE2025 Industry Alignment Fund. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This project was funded by the National Natural Science Foundation of China (Grant. 62433016); the Key Technologies R&D Program of Ningxia (Grant. 2022BEG02025 and 2023BEG02023).

Data availability

No data was used for the research described in the article.

References

[1] S.T. Mennemeyer, N. Menachemi, S. Rahurkar, E.W. Ford, Impact of the HITECH act on physicians' adoption of electronic health records, J. Am. Med. Informatics Assoc. 23 (2) (2016) 375–379.
[2] K.-i. Hashimoto, K. Tabata, Population aging, health care, and growth, J. Popul. Econ. 23 (2010) 571–593.
[3] X. Wu, C. Liu, L. Wang, M. Bilal, Internet of things-enabled real-time health monitoring system using deep learning, Neural Comput. Appl. (2023) 1–12.
[4] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. Van Der Laak, B. Van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.
[5] A.P. James, B.V. Dasarathy, Medical image fusion: A survey of the state of the art, Inf. Fusion 19 (2014) 4–19.
[6] D. Zhang, C. Yin, J. Zeng, X. Yuan, P. Zhang, Combining structured and unstructured data for predictive models: A deep learning approach, BMC Med. Inform. Decis. Mak. 20 (2020) 1–11.
[7] J.M. de Oliveira, C.A. da Costa, R.S. Antunes, Data structuring of electronic health records: A systematic review, Heal. Technol. (2021) 1–17.
[8] C.G. Arnold, B. Sonn, F.J. Meyers, A. Vest, R. Puls, E. Zirkler, M. Edelmann, I.M. Brooks, A.A. Monte, Accessing and utilizing clinical and genomic data from an electronic health record data warehouse, Transl. Med. Commun. 8 (1) (2023) 7.
[9] M. Sung, S. Hahn, C.H. Han, J.M. Lee, J. Lee, J. Yoo, J. Heo, Y.S. Kim, K.S. Chung, Event prediction model considering time and input error using electronic medical records in the intensive care unit: Retrospective study, JMIR Med. Inform. 9 (11) (2021) e26426.
[10] H. Duan, Z. Sun, W. Dong, K. He, Z. Huang, On clinical event prediction in patient treatment trajectory using longitudinal electronic health records, IEEE J. Biomed. Heal. Inform. 24 (7) (2019) 2053–2063.
[11] P. Singhal, A. Tan, T. Drivas, K. Johnson, M. Ritchie, B. Beaulieu-Jones, Opportunities and challenges for biomarker discovery using electronic health record data, Trends Mol. Med. (2023).
[12] Q.S. Wells, D.K. Gupta, J.G. Smith, S.P. Collins, A.B. Storrow, J. Ferguson, M.L. Smith, J.M. Pulley, S. Collier, X. Wang, et al., Accelerating biomarker discovery through electronic health records, automated biobanking, and proteomics, J. Am. Coll. Cardiol. 73 (17) (2019) 2195–2205.
[13] Y. Si, J. Du, Z. Li, X. Jiang, T. Miller, F. Wang, W.J. Zheng, K. Roberts, Deep representation learning of patient data from electronic health records (EHR): A systematic review, J. Biomed. Inform. 115 (2021) 103671.
[14] M. Sushil, S. Šuster, K. Luyckx, W. Daelemans, Patient representation learning and interpretable evaluation using clinical notes, J. Biomed. Inform. 84 (2018) 103–113.
[15] R.J. Chen, M.Y. Lu, T.Y. Chen, D.F. Williamson, F. Mahmood, Synthetic data in machine learning for medicine and healthcare, Nat. Biomed. Eng. 5 (6) (2021) 493–497.
[16] S. Han, R. Mao, E. Cambria, Hierarchical attention network for explainable depression detection on Twitter aided by metaphor concept mappings, in: Proceedings of the 29th International Conference on Computational Linguistics, COLING, International Committee on Computational Linguistics, Gyeongju, Republic of Korea, 2022, pp. 94–104.
[17] X. Mei, R. Mao, X. Cai, L. Yang, E. Cambria, Medical report generation via multimodal spatio-temporal fusion, in: Proceedings of the 32nd ACM International Conference on Multimedia (ACM MM), Melbourne, Australia, 2024, pp. 4699–4708.
[18] C. Yin, R. Zhao, B. Qian, X. Lv, P. Zhang, Domain knowledge guided deep learning with electronic health records, in: 2019 IEEE International Conference on Data Mining, ICDM, IEEE, 2019, pp. 738–747.
[19] C. Yang, C. Xiao, F. Ma, L. Glass, J. Sun, SafeDrug: Dual molecular graph encoders for recommending effective and safe drug combinations, in: 30th International Joint Conference on Artificial Intelligence, IJCAI 2021, 2021, pp. 3735–3741.
[20] J. Xie, B. Zhang, J. Ma, D. Zeng, J. Lo-Ciganic, Readmission prediction for patients with heterogeneous medical history: A trajectory-based deep learning approach, ACM Trans. Manag. Inf. Syst. (TMIS) 13 (2) (2021) 1–27.
[21] D.C. Montgomery, E.A. Peck, G.G. Vining, Introduction to Linear Regression Analysis, John Wiley & Sons, 2021.
[22] J. Wu, J. Roy, W.F. Stewart, Prediction modeling using EHR data: Challenges, strategies, and a comparison of machine learning approaches, Med. Care 48 (6) (2010) S106–S113.
[23] Y.-Y. Song, L. Ying, Decision tree methods: Applications for classification and prediction, Shanghai Arch. Psychiatry 27 (2) (2015) 130.
[24] J.L. Speiser, M.E. Miller, J. Tooze, E. Ip, A comparison of random forest variable selection methods for classification prediction modeling, Expert Syst. Appl. 134 (2019) 93–101.
[25] R. Miotto, F. Wang, S. Wang, X. Jiang, J.T. Dudley, Deep learning for healthcare: Review, opportunities and challenges, Brief. Bioinform. 19 (6) (2018) 1236–1246.
[26] E. Cambria, R. Mao, M. Chen, Z. Wang, S.-B. Ho, Seven pillars for the future of artificial intelligence, IEEE Intell. Syst. 38 (6) (2023) 62–69.
[27] L. Yue, D. Tian, W. Chen, X. Han, M. Yin, Deep learning for heterogeneous medical data analysis, World Wide Web 23 (2020) 2715–2737.
[28] K. Guo, T. Xu, X. Kui, R. Zhang, T. Chi, Ifusion: Towards efficient intelligence fusion for deep learning from real-time and heterogeneous data, Inf. Fusion 51 (2019) 215–223.
[29] X. Li, L. Zhang, Unbalanced data processing using deep sparse learning technique, Future Gener. Comput. Syst. 125 (2021) 480–484.
[30] A. Kumar, S. Goel, N. Sinha, A. Bhardwaj, A review on unbalanced data classification, in: Proceedings of International Joint Conference on Advances in Computational Intelligence, IJCACI 2021, Springer, 2022, pp. 197–208.
[31] X. Li, H. Xiong, X. Li, X. Wu, X. Zhang, J. Liu, J. Bian, D. Dou, Interpretable deep learning: Interpretation, interpretability, trustworthiness, and beyond, Knowl. Inf. Syst. 64 (12) (2022) 3197–3234.

[32] A.A. Ismail, M. Gunady, H. Corrada Bravo, S. Feizi, Benchmarking deep learning interpretability in time series predictions, Adv. Neural Inf. Process. Syst. 33 (2020) 6441–6452.

[33] T. Shaik, X. Tao, L. Li, H. Xie, J.D. Velásquez, A survey of multimodal information fusion for smart healthcare: Mapping the journey from data to wisdom, Inf. Fusion 102 (2024) 102040.

[34] X. Tao, J.D. Velasquez, Multi-source information fusion for smart health with artificial intelligence, 2022.

[35] J. Wang, J. Luo, M. Ye, X. Wang, Y. Zhong, A. Chang, G. Huang, Z. Yin, C. Xiao, J. Sun, et al., Recent advances in predictive modeling with electronic health records, in: Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, 2024, pp. 8272–8280.

[36] N.A. Nasarudin, F. Al Jasmi, R.O. Sinnott, N. Zaki, H. Al Ashwal, E.A. Mohamed, M.S. Mohamad, A review of deep learning models and online healthcare databases for electronic health records and their use for health prediction, Artif. Intell. Rev. 57 (9) (2024) 1–26.

[37] A. Amirahmadi, M. Ohlsson, K. Etminani, Deep learning prediction models based on EHR trajectories: A systematic review, J. Biomed. Inform. (2023) 104430.

[38] B. Shickel, P.J. Tighe, A. Bihorac, P. Rashidi, Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis, IEEE J. Biomed. Heal. Inform. 22 (5) (2017) 1589–1604.

[39] Y. Juhn, H. Liu, Artificial intelligence approaches using natural language processing to advance EHR-based clinical research, J. Allergy Clin. Immunol. 145 (2) (2020) 463–469.

[40] A. Dagliati, A. Malovini, V. Tibollo, R. Bellazzi, Health informatics and EHR to support clinical research in the COVID-19 pandemic: An overview, Brief. Bioinform. 22 (2) (2021) 812–822.

[41] T. Al-Quraishi, N. Al-Quraishi, H. AlNabulsi, H. AL-Qarishey, A.H. Ali, Big data predictive analytics for personalized medicine: Perspectives and challenges, Appl. Data Sci. Anal. 2024 (2024) 32–38.

[42] Y. Zhang, J. Gao, Z. Tan, L. Zhou, K. Ding, M. Zhou, S. Zhang, D. Wang, Data-centric foundation models in computational healthcare: A survey, 2024, arXiv preprint arXiv:2401.02458.

[43] Q. Lin, Y. Zhu, X. Mei, L. Huang, J. Ma, K. He, Z. Peng, E. Cambria, M. Feng, Has multimodal learning delivered universal intelligence in healthcare? a comprehensive survey, Inf. Fusion (2024) 102795.

[44] K. He, R. Mao, Q. Lin, Y. Ruan, X. Lan, M. Feng, E. Cambria, A survey of large language models for healthcare: From data, technology, and applications to accountability and ethics, Inf. Fusion (2025) 102963.

[45] X. Lan, F. Wu, K. He, Q. Zhao, S. Hong, M. Feng, Gem: Empowering mllm for grounded ecg understanding with time series and images, 2025, arXiv preprint arXiv:2503.06073.

[46] R. Mao, G. Chen, X. Zhang, F. Guerin, E. Cambria, GPTEval: A survey on assessments of ChatGPT and GPT-4, in: Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), ELRA and ICCL, Torino, Italia, 2024, pp. 7844–7866.

[47] K. He, B. Mao, X. Zhou, Y. Li, T. Gong, C. Li, J. Wu, Knowledge enhanced coreference resolution via gated attention, in: 2022 IEEE International Conference on Bioinformatics and Biomedicine, BIBM, IEEE, 2022, pp. 2287–2293.

[48] B. Mao, C. Jia, Y. Huang, K. He, J. Wu, T. Gong, C. Li, Uncertainty-guided mutual consistency training for semi-supervised biomedical relation extraction, in: 2022 IEEE International Conference on Bioinformatics and Biomedicine, BIBM, IEEE, 2022, pp. 2318–2325.

[49] Y. Li, X. Ma, X. Zhou, P. Cheng, K. He, T. Gong, C. Li, Integrating K+ entities into coreference resolution on biomedical texts, IEEE/ ACM Trans. Comput. Biol. Bioinform. (2024).

[50] L. Tong, W. Shi, M. Isgut, Y. Zhong, P. Lais, L. Gloster, J. Sun, A. Swain, F. Giuste, M.D. Wang, Integrating multi-omics data with EHR for precision medicine using advanced artificial intelligence, IEEE Rev. Biomed. Eng. (2023).

[51] M. Kang, E. Ko, T.B. Mersha, A roadmap for multi-omics data integration using deep learning, Brief. Bioinform. 23 (1) (2022) bbab454.

[52] A.S. Tang, S.R. Woldemariam, S. Miramontes, B. Norgeot, T.T. Oskotsky, M. Sirota, Harnessing EHR data for health research, Nature Med. 30 (7) (2024) 1847–1855.

[53] X. He, X. Liu, F. Zuo, H. Shi, J. Jing, Artificial intelligence-based multi-omics analysis fuels cancer precision medicine, in: Seminars in Cancer Biology, vol. 88, Elsevier, 2023, pp. 187–200.

[54] K. He, J. Wu, X. Ma, C. Zhang, M. Huang, C. Li, L. Yao, Extracting kinship from obituary to enhance electronic health records for genetic research, in: Proceedings of the Fourth Social Media Mining for Health Applications (#SMM4H) Workshop & Shared Task, 2019, pp. 1–10.

[55] K. He, L. Yao, J. Zhang, Y. Li, C. Li, Construction of genealogical knowledge graphs from obituaries: Multitask neural network extraction system, J. Med. Internet Res. 23 (8) (2021) e25670.

[56] M.J. Page, J.E. McKenzie, P.M. Bossuyt, I. Boutron, T.C. Hoffmann, C.D. Mulrow, L. Shamseer, J.M. Tetzlaff, E.A. Akl, S.E. Brennan, et al., The PRISMA 2020 statement: An updated guideline for reporting systematic reviews, Bmj 372 (2021).

[57] K. Häyrinen, K. Saranto, P. Nykänen, Definition, structure, content, use and impacts of electronic health records: A review of the research literature, Int. J. Med. Inform. 77 (5) (2008) 291–304.

[58] N.G. Weiskopf, C. Weng, Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research, J. Am. Med. Inform. Assoc. 20 (1) (2013) 144–151.

[59] S.M. Meystre, G.K. Savova, K.C. Kipper-Schuler, J.F. Hurdle, Extracting information from textual documents in the electronic health record: A review of recent research, Yearb. Med. Inform. 17 (01) (2008) 128–144.

[60] O. Bodenreider, R. Cornet, D.J. Vreeman, Recent developments in clinical terminologies—SNOMED CT, LOINC, and RxNorm, Yearb. Med. Inform. 27 (01) (2018) 129–139.

[61] R. Yan, F. Zhang, X. Rao, Z. Lv, J. Li, L. Zhang, S. Liang, Y. Li, F. Ren, C. Zheng, et al., Richer fusion network for breast cancer classification based on multimodal data, BMC Med. Inform. Decis. Mak. 21 (2021) 1–15.

[62] M. Xu, L. Ouyang, L. Han, K. Sun, T. Yu, Q. Li, H. Tian, L. Safarnejad, H. Zhang, Y. Gao, et al., Accurately differentiating between patients with COVID-19, patients with other viral infections, and healthy individuals: Multimodal late fusion learning approach, J. Med. Internet Res. 23 (1) (2021) e25535.

[63] C. Fang, S. Bai, Q. Chen, Y. Zhou, L. Xia, L. Qin, S. Gong, X. Xie, C. Zhou, D. Tu, et al., Deep learning for predicting COVID-19 malignant progression, Med. Image Anal. 72 (2021) 102096.

[64] S.-C. Huang, A. Pareek, R. Zamanian, I. Banerjee, M.P. Lungren, Multimodal fusion with deep neural networks for leveraging CT imaging and electronic health record: A case-study in pulmonary embolism detection, Sci. Rep. 10 (1) (2020) 22147.

[65] Y. Chai, Y. Bian, H. Liu, J. Li, J. Xu, Glaucoma diagnosis in the Chinese context: An uncertainty information-centric Bayesian deep learning model, Inf. Process. Manage. 58 (2) (2021) 102454.

[66] S. Ding, H. Huang, Z. Li, X. Liu, S. Yang, SCNET: A novel UGI cancer screening framework based on semantic-level multimodal data fusion, IEEE J. Biomed. Heal. Inform. 25 (1) (2020) 143–151.

[67] K. He, R. Mao, T. Gong, E. Cambria, C. Li, JCBIE: A joint continual learning neural network for biomedical information extraction, BMC Bioinform. 23 (1) (2022) 549.

[68] R. Deng, N. Shaikh, G. Shannon, Y. Nie, Cross-modality attention-based multimodal fusion for non-small cell lung cancer (NSCLC) patient survival prediction, in: Medical Imaging 2024: Digital and Computational Pathology, vol. 12933, SPIE, 2024, pp. 46–50.

[69] N. Braman, J.W. Gordon, E.T. Goossens, C. Willis, M.C. Stumpe, J. Venkataraman, Deep orthogonal fusion: multimodal prognostic biomarker discovery integrating radiology, pathology, genomic, and clinical data, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24, Springer, 2021, pp. 667–677.

[70] L. Ma, J. Gao, Y. Wang, C. Zhang, J. Wang, W. Ruan, W. Tang, X. Gao, X. Ma, Adacare: Explainable clinical health status representation learning via scale-adaptive feature extraction and recalibration, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, 2020, pp. 825–832, 01.

[71] X. Zhang, B. Qian, S. Cao, Y. Li, H. Chen, Y. Zheng, I. Davidson, INPREM: An interpretable and trustworthy predictive model for healthcare, in: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2020, pp. 450–460.

[72] Y. Li, B. Qian, X. Zhang, H. Liu, Graph neural network-based diagnosis prediction, Big Data 8 (5) (2020) 379–390.

[73] J. Shang, C. Xiao, T. Ma, H. Li, J. Sun, Gamenet: Graph augmented memory networks for recommending medication combination, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 1126–1133, 01.

[74] Y. Tan, C. Kong, L. Yu, P. Li, C. Chen, X. Zheng, V.S. Hertzberg, C. Yang, 4Sdrug: Symptom-based set-to-set small and safe drug recommendation, in: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022, pp. 3970–3980.

[75] Z. Sun, H. Yin, H. Chen, T. Chen, L. Cui, F. Yang, Disease prediction via graph neural networks, IEEE J. Biomed. Heal. Inform. 25 (3) (2020) 818–826.

[76] C. Lu, C.K. Reddy, P. Chakraborty, S. Kleinberg, Y. Ning, Collaborative graph learning with auxiliary text for temporal event prediction in healthcare, in: International Joint Conference on Artificial Intelligence, 2021.

[77] M. Ye, S. Cui, Y. Wang, J. Luo, C. Xiao, F. Ma, Medretriever: Target-driven interpretable health risk prediction via retrieving unstructured medical text, in: Proceedings of the 30th ACM International Conference on Information & Knowledge Management, 2021, pp. 2414–2423.

[78] C.W. Cheong, K. Yin, W.K. Cheung, B.C. Fung, J. Poon, Adaptive integration of categorical and multi-relational ontologies with EHR data for medical concept embedding, ACM Trans. Intell. Syst. Technol. 14 (6) (2023) 1–20.

[79] K. Wang, N. Chen, T. Chen, Joint medical ontology representation learning for healthcare predictions, in: 2020 International Joint Conference on Neural Networks, IJCNN, IEEE, 2020, pp. 1–7.

[80] J. Wu, K. He, R. Mao, C. Li, E. Cambria, MEGACare: Knowledge-guided multi-view hypergraph predictive framework for healthcare, Inf. Fusion 100 (2023) 101939.

[81] C. Knox, M. Wilson, C.M. Klinger, M. Franklin, E. Oler, A. Wilson, A. Pon, J. Cox, N.E. Chin, S.A. Strawbridge, et al., DrugBank 6.0: The DrugBank knowledgebase for 2024, Nucleic Acids Res. 52 (D1) (2024) D1265–D1275.

[82] N. Yang, K. Zeng, Q. Wu, J. Yan, Molerec: Combinatorial drug recommendation with substructure-aware molecular representation learning, in: Proceedings of the ACM Web Conference 2023, 2023, pp. 4075–4085.

[83] L. Song, C.W. Cheong, K. Yin, W.K. Cheung, B.C. Fung, J. Poon, Medical concept embedding with multiple ontological representations, in: IJCAI, vol. 19, 2019, pp. 4613–4619.

[84] M. Zhang, C.R. King, M. Avidan, Y. Chen, Hierarchical attention propagation for healthcare representation learning, in: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2020, pp. 249–256.

[85] C. Chen, J. Liang, F. Ma, L. Glass, J. Sun, C. Xiao, Unite: Uncertainty-based health risk prediction leveraging multi-sourced data, in: Proceedings of the Web Conference 2021, 2021, pp. 217–226.

[86] C. Sirocchi, A. Bogliolo, S. Montagna, Medical-informed machine learning: Integrating prior knowledge into medical decision systems, BMC Med. Inform. Decis. Mak. 24 (Suppl 4) (2024) 186.

[87] M. Khalid, R. Rahman, A. Abbas, S. Kumari, I. Wajahat, S.A.C. Bukhari, Accelerating medical knowledge discovery through automated knowledge graph generation and enrichment, in: International Knowledge Graph and Semantic Web Conference, Springer, 2024, pp. 62–77.

[88] A. Flanagin, G. Curfman, K. Bibbins-Domingo, Data sharing and the growth of medical knowledge, JAMA 328 (24) (2022) 2398–2399.

[89] A. Cuff, The evolution of digital health and its continuing challenges, BMC Digit. Heal. 1 (1) (2023) 3.

[90] S. Denaxas, C. Stoeckert, Data standards and terminology including biomedical ontologies, in: Clinical Applications of Artificial Intelligence in Real-World Data, Springer, 2023, pp. 37–49.

[91] R.J. Little, D.B. Rubin, Statistical Analysis with Missing Data, vol. 793, John Wiley & Sons, 2019.

[92] Z. Chen, S. Tan, U. Chajewska, C. Rudin, R. Caruna, Missing values and imputation in healthcare data: Can interpretable machine learning help? in: Conference on Health, Inference, and Learning, PMLR, 2023, pp. 86–99.

[93] J. Pathak, K.R. Bailey, C.E. Beebe, S. Bethard, D.S. Carrell, P.J. Chen, D. Dligach, C.M. Endle, L.A. Hart, P.J. Haug, et al., Normalization and standardization of electronic health records for high-throughput phenotyping: The SHARPn consortium, J. Am. Med. Inform. Assoc. 20 (e2) (2013) e341–e348.

[94] X.H. Cao, I. Stojkovic, Z. Obradovic, A robust data scaling algorithm to improve classification accuracies in biomedical data, BMC BioInform. 17 (2016) 1–10.

[95] M. Liu, S. Li, H. Yuan, M.E.H. Ong, Y. Ning, F. Xie, S.E. Saffari, Y. Shang, V. Volovici, B. Chakraborty, et al., Handling missing values in healthcare data: A systematic review of deep learning-based imputation techniques, Artif. Intell. Med. 142 (2023) 102587.

[96] K. Kreimeyer, M. Foster, A. Pandey, N. Arya, G. Halford, S.F. Jones, R. Forshee, M. Walderhaug, T. Botsis, Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review, J. Biomed. Inform. 73 (2017) 14–29.

[97] E. Alsentzer, J. Murphy, W. Boag, W.-H. Weng, D. Jindi, T. Naumann, M. McDermott, Publicly available clinical BERT embeddings, in: Proceedings of the 2nd Clinical Natural Language Processing Workshop, 2019, pp. 72–78.

[98] D. Demner-Fushman, W.W. Chapman, C.J. McDonald, What can natural language processing do for clinical decision support? J. Biomed. Inform. 42 (5) (2009) 760–772.

[99] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C.H. So, J. Kang, BioBERT: A pre-trained biomedical language representation model for biomedical text mining, BioInformatics 36 (4) (2020) 1234–1240.

[100] A. Kaur, G. Dong, A complete review on image denoising techniques for medical images, Neural Process. Lett. 55 (6) (2023) 7807–7850.

[101] H. Lee, H. Lee, H. Hong, GenMix: Combining generative and mixture data augmentation for medical image classification, 2024, arXiv preprint arXiv:2405.20650.

[102] X. Zeng, N. Abdullah, P. Sumari, Self-supervised learning framework application for medical image analysis: A review and summary, BioMed. Eng. OnLine 23 (1) (2024) 107.

[103] Y. Ding, X. Qin, M. Zhang, J. Geng, D. Chen, F. Deng, C. Song, RLSegNet: An medical image segmentation network based on reinforcement learning, IEEE/ACM Trans. Comput. Biol. Bioinform. 20 (4) (2022) 2565–2576.

[104] Ö. Kasim, M. Tosun, Effective removal of eye-blink artifacts in EEG signals with semantic segmentation, Signal, Image Video Process. (2022) 1–7.

[105] H.U. Amin, A.S. Malik, R.F. Ahmad, N. Badruddin, N. Kamel, M. Hussain, W.-T. Chooi, Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques, Australas. Phys. Eng. Sci. Med. 38 (2015) 139–149.

[106] M.F. Issa, Z. Juhasz, Improved EOG artifact removal using wavelet enhanced independent component analysis, Brain Sci. 9 (12) (2019) 355.

[107] A.K. Singh, S. Krishnan, Trends in EEG signal feature extraction applications, Front. Artif. Intell. 5 (2023) 1072801.

[108] M. Mielczarek, J. Szyda, Review of alignment and SNP calling algorithms for next-generation sequencing data, J. Appl. Genet. 57 (1) (2016) 71–79.

[109] J.S. Wekesa, M. Kimwele, A review of multi-omics data integration through deep learning approaches for disease diagnosis, prognosis, and treatment, Front. Genet. 14 (2023) 1199087.

[110] P.H.G. Sanches, N.C. de Melo, A.M. Porcari, L.M. de Carvalho, Integrating molecular perspectives: Strategies for comprehensive multi-omics integrative data analysis and machine learning applications in transcriptomics, proteomics, and metabolomics, Biology 13 (11) (2024) 848.

[111] M. Mahmoud, N. Gobet, D.I. Cruz-Dávalos, N. Mounier, C. Dessimoz, F.J. Sedlazeck, Structural variant calling: The long and the short of it, Genome Biology 20 (2019) 1–14.

[112] R. Couronné, P. Probst, A.-L. Boulesteix, Random forest versus logistic regression: A large-scale benchmark experiment, BMC BioInform. 19 (2018) 1–14.

[113] A.B. Musa, Comparative study on classification performance between support vector machine and logistic regression, Int. J. Mach. Learn. Cybern. 4 (2013) 13–24.

[114] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.

[115] W. Hong, X. Zhou, S. Jin, Y. Lu, J. Pan, Q. Lin, S. Yang, T. Xu, Z. Basharat, M. Zippi, et al., A comparison of xgboost, random forest, and nomograph for the prediction of disease severity in patients with COVID-19 pneumonia: Implications of cytokine and immune cell profile, Front. Cell. Infect. Microbiol. 12 (2022) 819267.

[116] K. Cho, B. van Merrienboer, C. Gulcehre, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation, in: Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, 2014.

[117] S. Hochreiter, Long short-term memory, Neural Comput. MIT-Press (1997).

[118] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[119] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, Springer, 2015, pp. 234–241.

[120] A. Vaswani, Attention is all you need, Adv. Neural Inf. Process. Syst. (2017).

[121] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., An image is worth 16x16 words: Transformers for image recognition at scale, in: International Conference on Learning Representations, 2020.

[122] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, Swin transformer: Hierarchical vision transformer using shifted windows, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 10012–10022.

[123] T.N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, in: International Conference on Learning Representations, 2017.

[124] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, Y. Bengio, et al., Graph attention networks, Stat 1050 (20) (2017) 10–48550.

[125] J. Gilmer, S.S. Schoenholz, P.F. Riley, O. Vinyals, G.E. Dahl, Neural message passing for quantum chemistry, in: International Conference on Machine Learning, PMLR, 2017, pp. 1263–1272.

[126] J.D.M.-W.C. Kenton, L.K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, in: Proceedings of NaacL-HLT, vol. 1, Minneapolis, Minnesota, 2019, 2.

[127] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, D. Krishnan, Supervised contrastive learning, Adv. Neural Inf. Process. Syst. 33 (2020) 18661–18673.

[128] T. Chen, S. Kornblith, M. Norouzi, G. Hinton, A simple framework for contrastive learning of visual representations, in: International Conference on Machine Learning, PMLR, 2020, pp. 1597–1607.

[129] S. Azizi, B. Mustafa, F. Ryan, Z. Beaver, J. Freyberg, J. Deaton, A. Loh, A. Karthikesalingam, S. Kornblith, T. Chen, et al., Big self-supervised models advance medical image classification, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 3478–3488.

[130] J. Chen, Y. Geng, Z. Chen, J.Z. Pan, Y. He, W. Zhang, I. Horrocks, H. Chen, Zero-shot and few-shot learning with knowledge graphs: A comprehensive survey, Proc. IEEE 111 (6) (2023) 653–685.

[131] Y. Wang, Q. Yao, J.T. Kwok, L.M. Ni, Generalizing from a few examples: A survey on few-shot learning, ACM Comput. Surv. (Csur) 53 (3) (2020) 1–34.

[132] T. Hospedales, A. Antoniou, P. Micaelli, A. Storkey, Meta-learning in neural networks: A survey, IEEE Trans. Pattern Anal. Mach. Intell. 44 (9) (2021) 5149–5169.

[133] S. Ruder, et al., Transfer learning-machine learning's next frontier, 2017, Accessed April.

[134] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, A comprehensive survey on transfer learning, Proc. IEEE 109 (1) (2020) 43–76.

[135] P. Kairouz, H.B. McMahan, B. Avent, A. Bellet, M. Bennis, A.N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, et al., Advances and open problems in federated learning, Found. Trends® Mach. Learn. 14 (1–2) (2021) 1–210.

[136] G.A. Kaissis, M.R. Makowski, D. Rückert, R.F. Braren, Secure, privacy-preserving and federated machine learning in medical imaging, Nat. Mach. Intell. 2 (6) (2020) 305–311.

[137] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial networks, Commun. ACM 63 (11) (2020) 139–144.

[138] X. Yi, E. Walia, P. Babyn, Generative adversarial network in medical imaging: A review, Med. Image Anal. 58 (2019) 101552.

[139] K. Arulkumaran, M.P. Deisenroth, M. Brundage, A.A. Bharath, Deep reinforcement learning: A brief survey, IEEE Signal Process. Mag. 34 (6) (2017) 26–38.

[140] M. Komorowski, L.A. Celi, O. Badawi, A.C. Gordon, A.A. Faisal, The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care, Nature Med. 24 (11) (2018) 1716–1720.

[141] L. Yao, Z. Chu, S. Li, Y. Li, J. Gao, A. Zhang, A survey on causal inference, ACM Trans. Knowl. Discov. from Data (TKDD) 15 (5) (2021) 1–46.

[142] X. Pei, K. Zuo, Y. Li, Z. Pang, A review of the application of multimodal deep learning in medicine: bibliometrics and future directions, Int. J. Comput. Intell. Syst. 16 (1) (2023) 44.

[143] J. Li, D. Li, C. Xiong, S. Hoi, Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation, in: International Conference on Machine Learning, PMLR, 2022, pp. 12888–12900.

[144] A. Radford, J.W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, et al., Learning transferable visual models from natural language supervision, in: International Conference on Machine Learning, PMLR, 2021, pp. 8748–8763.

[145] Z. Cui, T. Gao, K. Talamadupula, Q. Ji, Knowledge-augmented deep learning and its applications: A survey, IEEE Trans. Neural Netw. Learn. Syst. (2023).

[146] B. Lambert, F. Forbes, S. Doyle, H. Dehaene, M. Dojat, Trustworthy clinical AI solutions: A unified review of uncertainty quantification in deep learning models for medical image analysis, Artif. Intell. Med. (2024) 102830.

[147] K. Zou, Z. Chen, X. Yuan, X. Shen, M. Wang, H. Fu, A review of uncertainty estimation and its application in medical imaging, Meta- Radiol. (2023) 100003.

[148] W. Farhan, Z. Wang, Y. Huang, S. Wang, F. Wang, X. Jiang, et al., A predictive model for medical events based on contextual embedding of temporal sequences, JMIR Med. Inform. 4 (4) (2016) e5977.

[149] A. Manashty, J. Light, Life model: A novel representation of life-long temporal sequences in health predictive analytics, Future Gener. Comput. Syst. 92 (2019) 141–156.

[150] H. Guan, M. Liu, Domain adaptation for medical image analysis: A survey, IEEE Trans. Biomed. Eng. 69 (3) (2021) 1173–1185.

[151] Y. Zhang, Y. Wei, Q. Wu, P. Zhao, S. Niu, J. Huang, M. Tan, Collaborative unsupervised domain adaptation for medical image diagnosis, IEEE Trans. Image Process. 29 (2020) 7834–7844.

[152] S. Aminabee, The future of healthcare and patient-centric care: Digital innovations, trends, and predictions, in: Emerging Technologies for Health Literacy and Medical Practice, IGI Global, 2024, pp. 240–262.

[153] S. Mulukuntla, S.P. Venkata, AI-driven personalized medicine: Assessing the impact of federal policies on advancing patient-centric care, EPH- Int. J. Med. Heal. Sci. 6 (2) (2020) 20–26.

[154] F. Ma, R. Chitta, J. Zhou, Q. You, T. Sun, J. Gao, Dipole: Diagnosis prediction in healthcare via attention-based bidirectional recurrent neural networks, in: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017, pp. 1903–1911.

[155] F. Ma, Q. You, H. Xiao, R. Chitta, J. Zhou, J. Gao, Kame: Knowledge-based attention model for diagnosis prediction in healthcare, in: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018, pp. 743–752.

[156] F. Ma, Y. Wang, H. Xiao, Y. Yuan, R. Chitta, J. Zhou, J. Gao, A general framework for diagnosis prediction via incorporating medical code descriptions, in: 2018 IEEE International Conference on Bioinformatics and Biomedicine, BIBM, IEEE, 2018, pp. 1070–1075.

[157] L. Ma, C. Zhang, Y. Wang, W. Ruan, J. Wang, W. Tang, X. Ma, X. Gao, J. Gao, Concare: Personalized clinical feature embedding via capturing the healthcare context, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, 2020, pp. 833–840, 01.

[158] J. Gao, C. Xiao, Y. Wang, W. Tang, L.M. Glass, J. Sun, Stagenet: Stage-aware neural networks for health risk prediction, in: Proceedings of the Web Conference 2020, 2020, pp. 530–540.

[159] C. Zang, F. Wang, Scehr: Supervised contrastive learning for clinical risk prediction using electronic health records, in: Proceedings. IEEE International Conference on Data Mining, vol. 2021, NIH Public Access, 2021, p. 857.

[160] T. Al-Quraishi, L. Al-Omairi, R. Thakkar, C. Singh, J.I. Agbinya, O.A. Mahdi, B. Das, Analysis of breast cancer survivability using machine learning predictive technique for post-surgical patients, in: Proceedings of International Conference for ICT (ICICT)-Zambia, vol. 5, 2023, pp. 12–18, 1.

[161] T. Al-Quraishi, C.K. NG, O.A. Mahdi, A. Gyasi, N. Al-Quraishi, Advanced ensemble classifier techniques for predicting tumor viability in osteosarcoma histological slide images, Appl. Data Sci. Anal. 2024 (2024) 52–68.

[162] C. Zhang, X. Gao, L. Ma, Y. Wang, J. Wang, W. Tang, GRASP: generic framework for health status representation learning based on incorporating knowledge from similar patients, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, 2021, pp. 715–723, 1.

[163] X. Ma, Y. Wang, X. Chu, L. Ma, W. Tang, J. Zhao, Y. Yuan, G. Wang, Patient health representation learning via correlational sparse prior of medical features, IEEE Trans. Knowl. Data Eng. (2022).

[164] M. Sun, X. Yang, J. Niu, Y. Gu, C. Wang, W. Zhang, A cross-modal clinical prediction system for intensive care unit patient outcome, Knowl.-Based Syst. 283 (2024) 111160.

[165] C.-Y. Chi, S. Ao, A. Winkler, K.-C. Fu, J. Xu, Y.-L. Ho, C.-H. Huang, R. Soltani, Predicting the mortality and readmission of in-hospital cardiac arrest patients with electronic health records: A machine learning approach, J. Med. Internet Res. 23 (9) (2021) e27798.

[166] J. Wu, Y. Dong, Z. Gao, T. Gong, C. Li, Dual attention and patient similarity network for drug recommendation, BioInformatics 39 (1) (2023) btad003.

[167] J. Wu, X. Yu, K. He, Z. Gao, T. Gong, PROMISE: A pre-trained knowledge-infused multimodal representation learning framework for medication recommendation, Inf. Process. Manage. 61 (4) (2024) 103758.

[168] E. Choi, M.T. Bahadori, A. Schuetz, W.F. Stewart, J. Sun, Doctor ai: Predicting clinical events via recurrent neural networks, in: Machine Learning for Healthcare Conference, PMLR, 2016, pp. 301–318.

[169] Y. Xu, S. Biswal, S.R. Deshpande, K.O. Maher, J. Sun, Raim: Recurrent attentive and intensive model of multimodal patient monitoring data, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018, pp. 2565–2573.

[170] J. Chen, Y. Wen, M. Pokojovy, T.-L.B. Tseng, P. McCaffrey, A. Vo, E. Walser, S. Moen, Multi-modal learning for inpatient length of stay prediction, Comput. Biol. Med. 171 (2024) 108121.

[171] S. Pokharel, G. Zuccon, X. Li, C.P. Utomo, Y. Li, Temporal tree representation for similarity computation between medical patients, Artif. Intell. Med. 108 (2020) 101900.

[172] J.N. Acosta, G.J. Falcone, P. Rajpurkar, E.J. Topol, Multimodal biomedical AI, Nature Med. 28 (9) (2022) 1773–1784.

[173] O. Gottesman, H. Kuivaniemi, G. Tromp, W.A. Faucett, R. Li, T.A. Manolio, S.C. Sanderson, J. Kannry, R. Zinberg, M.A. Basford, et al., The electronic medical records and genomics (eMERGE) network: Past, present, and future, Genet. Med. 15 (10) (2013) 761–771.

[174] X. Zhang, S. Li, Z. Chen, X. Yan, L.R. Petzold, Improving medical predictions by irregular multimodal electronic health records modeling, in: International Conference on Machine Learning, PMLR, 2023, pp. 41300–41313.

[175] Y.C. Lee, J. Cha, I. Shim, W.-Y. Park, S.W. Kang, D.H. Lim, H.-H. Won, Multimodal deep learning of fundus abnormalities and traditional risk factors for cardiovascular risk prediction, Npj Digit. Med. 6 (1) (2023) 14.

[176] K. Zhang, H.-Y. Zhou, D.T. Baptista-Hon, Y. Gao, X. Liu, E. Oermann, S. Xu, S. Jin, J. Zhang, Z. Sun, et al., Concepts and applications of digital twins in healthcare and medicine, Patterns 5 (8) (2024).

[177] A.P. Bradley, The use of the area under the ROC curve in the evaluation of machine learning algorithms, Pattern Recognit. 30 (7) (1997) 1145–1159.

[178] K. Boyd, K.H. Eng, C.D. Page, Area under the precision-recall curve: point estimates and confidence intervals, in: Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2013, Prague, Czech Republic, September 23-27, 2013, Proceedings, Part III 13, Springer, 2013, pp. 451–466.

[179] Z. Wang, A.C. Bovik, Mean squared error: Love it or leave it? A new look at signal fidelity measures, IEEE Signal Process. Mag. 26 (1) (2009) 98–117.

[180] A. De Myttenaere, B. Golden, B. Le Grand, F. Rossi, Mean absolute percentage error for regression models, Neurocomputing 192 (2016) 38–48.

[181] S. Niwattanakul, J. Singthongchai, E. Naenudorn, S. Wanapu, Using of jaccard coefficient for keywords similarity, in: Proceedings of the International Multiconference of Engineers and Computer Scientists, vol. 1, 2013, pp. 380–384, 6.

[182] C. Yang, Z. Wu, P. Jiang, Z. Lin, J. Gao, B.P. Danek, J. Sun, Pyhealth: A deep learning toolkit for healthcare applications, in: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2023, pp. 5788–5789.

[183] Z. Kraljevic, T. Searle, A. Shek, L. Roguski, K. Noor, D. Bean, A. Mascio, L. Zhu, A.A. Folarin, A. Roberts, et al., Multi-domain clinical natural language processing with MedCAT: The medical concept annotation toolkit, Artif. Intell. Med. 117 (2021) 102083.

[184] M.J. Cardoso, W. Li, R. Brown, N. Ma, E. Kerfoot, Y. Wang, B. Murrey, A. Myronenko, C. Zhao, D. Yang, et al., Monai: An open-source framework for deep learning in healthcare, 2022, arXiv preprint arXiv:2211.02701.

[185] E. Gibson, W. Li, C. Sudre, L. Fidon, D.I. Shakir, G. Wang, Z. Eaton-Rosen, R. Gray, T. Doel, Y. Hu, et al., NiftyNet: A deep-learning platform for medical imaging, Comput. Methods Programs Biomed. 158 (2018) 113–122.

[186] L. Oneto, F. Bisio, E. Cambria, D. Anguita, Statistical learning theory and elm for big social data analysis, IEEE Comput. Intell. Mag. 11 (3) (2016) 45–55.

[187] E. Choi, M.T. Bahadori, L. Song, W.F. Stewart, J. Sun, GRAM: Graph-based attention model for healthcare representation learning, in: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017, pp. 787–795.

[188] E. Choi, C. Xiao, W. Stewart, J. Sun, Mime: Multilevel medical embedding of electronic health records for predictive healthcare, Adv. Neural Inf. Process. Syst. 31 (2018).

[189] L. Ma, X. Ma, J. Gao, X. Jiao, Z. Yu, C. Zhang, W. Ruan, Y. Wang, W. Tang, J. Wang, Distilling knowledge from publicly available online EMR data to emerging epidemic for prognosis, in: Proceedings of the Web Conference 2021, 2021, pp. 3558–3568.

[190] H. Ren, J. Wang, W.X. Zhao, Generative adversarial networks enhanced pre-training for insufficient electronic health records modeling, in: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022, pp. 3810–3818.

[191] X. Xu, X. Xu, Y. Sun, X. Liu, X. Li, G. Xie, F. Wang, Predictive modeling of clinical events with mutual enhancement between longitudinal patient records and medical knowledge graph, in: 2021 IEEE International Conference on Data Mining, ICDM, IEEE, 2021, pp. 777–786.

[192] X. Zhang, B. Qian, Y. Li, C. Yin, X. Wang, Q. Zheng, KnowRisk: An interpretable knowledge-guided model for disease risk prediction, in: 2019 IEEE International Conference on Data Mining, ICDM, IEEE, 2019, pp. 1492–1497.

[193] R. Mao, Q. Liu, K. He, W. Li, E. Cambria, The biases of pre-trained language models: An empirical study on prompt-based sentiment analysis and emotion detection, IEEE Trans. Affect. Comput. 14 (3) (2023) 1743–1753.

[194] Y. Xu, X. Chu, K. Yang, Z. Wang, P. Zou, H. Ding, J. Zhao, Y. Wang, B. Xie, Seqcare: Sequential training with external medical knowledge graph for diagnosis prediction in healthcare data, in: Proceedings of the ACM Web Conference 2023, 2023, pp. 2819–2830.

[195] X. Li, S. Liang, Y. Hou, T. Ma, StratMed: Relevance stratification between biomedical entities for sparsity on medication recommendation, Knowl.-Based Syst. 284 (2024) 111239.

[196] T. Kosolwattana, C. Liu, R. Hu, S. Han, H. Chen, Y. Lin, A self-inspected adaptive SMOTE algorithm (SASMOTE) for highly imbalanced data classification in healthcare, BioData Min. 16 (1) (2023) 15.

[197] L.-S. Chen, S.-J. Cai, Neural-network-based resampling method for detecting diabetes mellitus, J. Med. Biol. Eng. 35 (2015) 824–832.

[198] K. Armanious, C. Jiang, M. Fischer, T. Küstner, T. Hepp, K. Nikolaou, S. Gatidis, B. Yang, MedGAN: Medical image translation using GANs, Comput. Med. Imaging Graph. 79 (2020) 101684.

[199] Z. Zhang, C. Yan, T.A. Lasko, J. Sun, B.A. Malin, SynTEG: A framework for temporal structured electronic health data simulation, J. Am. Med. Inform. Assoc. 28 (3) (2021) 596–604.

[200] Y. Zhong, S. Cui, J. Wang, X. Wang, Z. Yin, Y. Wang, H. Xiao, M. Huai, T. Wang, F. Ma, Meddiffusion: Boosting health risk prediction via diffusion-based data augmentation, in: Proceedings of the 2024 SIAM International Conference on Data Mining, SDM, SIAM, 2024, pp. 499–507.

[201] Z. Wang, J. Sun, PromptEHR: Conditional electronic healthcare records generation with prompt learning, in: Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing, vol. 2022, 2022, p. 2873.

[202] Z. Qiao, X. Wu, S. Ge, W. Fan, MNN: Multimodal attentional neural networks for diagnosis prediction, in: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19, International Joint Conferences on Artificial Intelligence Organization, 2019, pp. 5937–5943.

[203] T. Ma, C. Xiao, F. Wang, Health-atm: A deep architecture for multifaceted patient health record representation and risk prediction, in: Proceedings of the 2018 SIAM International Conference on Data Mining, SIAM, 2018, pp. 261–269.

[204] S. Bhoi, M.L. Lee, W. Hsu, H.S.A. Fang, N.C. Tan, Personalizing medication recommendation with a graph-based approach, ACM Trans. Inf. Syst. (TOIS) 40 (3) (2021) 1–23.

[205] Z. Wu, A. Dadu, N. Tustison, B. Avants, M. Nalls, J. Sun, F. Faghri, Multimodal patient representation learning with missing modalities and labels, in: The Twelfth International Conference on Learning Representations, 2023.

[206] W. Yao, K. Yin, W.K. Cheung, J. Liu, J. Qin, Drfuse: Learning disentangled representation for clinical multi-modal fusion with missing modality and modal inconsistency, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 38, 2024, pp. 16416–16424, 15.

[207] C. Zhang, X. Chu, L. Ma, Y. Zhu, Y. Wang, J. Wang, J. Zhao, M3care: Learning with missing modalities in multimodal healthcare data, in: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022, pp. 2418–2428.

[208] M. Xu, Z. Zhu, Y. Li, S. Zheng, Y. Zhao, K. He, Y. Zhao, FlexCare: Leveraging cross-task synergy for flexible multimodal healthcare prediction, in: Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2024, pp. 3610–3620.

[209] X. Han, H. Nguyen, C.W. Harris, N. Ho, S. Saria, FuseMoE: Mixture-of-experts transformers for fleximodal fusion, in: The Thirty-Eighth Annual Conference on Neural Information Processing Systems, 2024.

[210] S. Yun, I. Choi, J. Peng, Y. Wu, J. Bao, Q. Zhang, J. Xin, Q. Long, T. Chen, Flex-moe: Modeling arbitrary modality combination via the flexible mixture-of-experts, 2024, arXiv preprint arXiv:2410.08245.

[211] A.M. Alaa, M. van der Schaar, Attentive state-space modeling of disease progression, Adv. Neural Inf. Process. Syst. 32 (2019).

[212] J. Gao, X. Wang, Y. Wang, Z. Yang, J. Gao, J. Wang, W. Tang, X. Xie, Camp: Co-attention memory networks for diagnosis prediction in healthcare, in: 2019 IEEE International Conference on Data Mining, ICDM, IEEE, 2019, pp. 1036–1041.

[213] X. Zhang, B. Qian, Y. Li, Z. Gao, C. Guan, R. Wang, Y. Zheng, H. Zheng, C. Li, Learning representations from local to global for fine-grained patient similarity measuring in intensive care unit, in: 2022 IEEE International Conference on Data Mining, ICDM, IEEE, 2022, pp. 713–722.

[214] J. Huang, B. Yang, K. Yin, J. Xu, Dna-t: Deformable neighborhood attention transformer for irregular medical time series, IEEE J. Biomed. Heal. Inform. (2024).

[215] Z. Yu, J. Wang, W. Luo, R. Tse, G. Pau, Mpre: Multi-perspective patient representation extractor for disease prediction, in: 2023 IEEE International Conference on Data Mining, ICDM, IEEE, 2023, pp. 758–767.

[216] J. Chen, C. Yin, Y. Wang, P. Zhang, Predictive modeling with temporal graphical representation on electronic health records, in: IJCAI: Proceedings of the Conference, vol. 2024, 2024, p. 5763.

[217] Y. Zhang, X. Yang, J. Ivy, M. Chi, ATTAIN: Attention-based time-aware LSTM networks for disease progression modeling, in: Proceedings of the 28th International Joint Conference on Artificial Intelligence, 2019, pp. 4369–4375.

[218] J. Luo, M. Ye, C. Xiao, F. Ma, Hitanet: Hierarchical time-aware attention networks for risk prediction on electronic health records, in: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2020, pp. 647–656.

[219] J. Zhang, S. Zheng, W. Cao, J. Bian, J. Li, Warpformer: A multi-scale modeling approach for irregular clinical time series, in: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2023, pp. 3273–3285.

[220] I.M. Baytas, C. Xiao, X. Zhang, F. Wang, A.K. Jain, J. Zhou, Patient subtyping via time-aware LSTM networks, in: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017, pp. 65–74.

[221] X. Peng, G. Long, T. Shen, S. Wang, J. Jiang, Sequential diagnosis prediction with transformer and ontological representation, in: 2021 IEEE International Conference on Data Mining, ICDM, IEEE, 2021, pp. 489–498.

[222] E. Choi, Z. Xu, Y. Li, M. Dusenberry, G. Flores, E. Xue, A. Dai, Learning the graphical structure of electronic health records with graph convolutional transformer, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, 2020, pp. 606–613, 01.

[223] C. Lu, T. Han, Y. Ning, Context-aware health event prediction via transition functions on dynamic disease graphs, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36, 2022, pp. 4567–4574, 4.

[224] W. Lee, S. Park, W. Joo, I.-C. Moon, Diagnosis prediction via medical context attention networks using deep generative modeling, in: 2018 IEEE International Conference on Data Mining, ICDM, IEEE, 2018, pp. 1104–1109.

[225] Y. An, L. Zhang, H. Yang, L. Sun, B. Jin, C. Liu, R. Yu, X. Wei, Prediction of treatment medicines with dual adaptive sequential networks, IEEE Trans. Knowl. Data Eng. 34 (11) (2021) 5496–5509.

[226] J. Shang, T. Ma, C. Xiao, J. Sun, Pre-training of graph augmented transformers for medication recommendation, in: 28th International Joint Conference on Artificial Intelligence, IJCAI 2019, 2019, pp. 5953–5959.

[227] S. Wang, P. Ren, Z. Chen, Z. Ren, J. Ma, M. de Rijke, Order-free medicine combination prediction with graph convolutional reinforcement learning, in: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 2019, pp. 1623–1632.

[228] C. Gao, S. Yin, H. Wang, Z. Wang, Z. Du, X. Li, Medical-knowledge-based graph neural network for medication combination prediction, IEEE Trans. Neural Netw. Learn. Syst. (2023).

[229] Z. Yao, B. Liu, F. Wang, D. Sow, Y. Li, Ontology-aware prescription recommendation in treatment pathways using multi-evidence healthcare data, ACM Trans. Inf. Syst. 41 (4) (2023) 1–29.

[230] J. Wu, B. Qian, Y. Li, Z. Gao, M. Ju, Y. Yang, Y. Zheng, T. Gong, C. Li, X. Zhang, Leveraging multiple types of domain knowledge for safe and effective drug recommendation, in: Proceedings of the 31st ACM International Conference on Information & Knowledge Management, 2022, pp. 2169–2178.

[231] A. Prakash, S. Zhao, S. Hasan, V. Datla, K. Lee, A. Qadir, J. Liu, O. Farri, Condensed memory networks for clinical diagnostic inferencing, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 31, 2017, 1.

[232] S.J. Im, Y. Xu, J. Watson, Learning medical concept representation based on semantic information in medical textural data, Expert Syst. Appl. 238 (2024) 122123.

[233] T. Bai, S. Vucetic, Improving medical code prediction from clinical text via incorporating online knowledge sources, in: The World Wide Web Conference, 2019, pp. 72–82.

[234] H. Sun, S. Xie, S. Li, Y. Chen, J.-R. Wen, R. Yan, Debiased, longitudinal and coordinated drug recommendation through multi-visit clinic records, Adv. Neural Inf. Process. Syst. 35 (2022) 27837–27849.

[235] Q. Chen, X. Li, K. Geng, M. Wang, Context-aware safe medication recommendations with molecular graph and DDI graph embedding, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, 2023, pp. 7053–7060, 6.

[236] S. Bhoi, M.L. Lee, W. Hsu, N.C. Tan, REFINE: A fine-grained medication recommendation system using deep learning and personalized drug interaction modeling, Adv. Neural Inf. Process. Syst. 36 (2024).

[237] H. Ren, J. Wang, W.X. Zhao, N. Wu, Rapt: Pre-training of time-aware transformer for learning robust healthcare representation, in: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2021, pp. 3503–3511.

[238] X. Wang, J. Luo, J. Wang, Y. Zhong, X. Zhang, Y. Wang, P. Bhatia, C. Xiao, F. Ma, Unity in diversity: Collaborative pre-training across multimodal medical sources, in: Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2024, pp. 3644–3656.

[239] J. Yang, B. Su, W.X. Zhao, J.-R. Wen, Unlocking the power of spatial and temporal information in medical multimodal pre-training, in: Proceedings of the 41st International Conference on Machine Learning, 2024, pp. 56382–56396.

[240] Y. Li, S. Rao, J.R.A. Solares, A. Hassaine, R. Ramakrishnan, D. Canoy, Y. Zhu, K. Rahimi, G. Salimi-Khorshidi, BEHRT: Transformer for electronic health records, Sci. Rep. 10 (1) (2020) 7155.

[241] L. Rasmy, Y. Xiang, Z. Xie, C. Tao, D. Zhi, Med-BERT: Pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction, NPJ Digit. Med. 4 (1) (2021) 86.

[242] Z. Wang, R. Wen, X. Chen, S. Cao, S.-L. Huang, B. Qian, Y. Zheng, Online disease diagnosis with inductive heterogeneous graph convolutional networks, in: Proceedings of the Web Conference 2021, 2021, pp. 3349–3358.

[243] Y. Tan, C. Yang, X. Wei, C. Chen, W. Liu, L. Li, J. Zhou, X. Zheng, Metacare++: Meta-learning with hierarchical subtyping for cold-start diagnosis prediction in healthcare data, in: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022, pp. 449–459.

[244] M. Ye, S. Cui, Y. Wang, J. Luo, C. Xiao, F. Ma, Medpath: Augmenting health risk prediction via medical knowledge paths, in: Proceedings of the Web Conference 2021, 2021, pp. 1397–1409.

[245] Y. Tan, Z. Zhou, L. Yu, W. Liu, C. Chen, G. Ma, X. Hu, V.S. Hertzberg, C. Yang, Enhancing personalized healthcare via capturing disease severity, interaction, and progression, in: 2023 IEEE International Conference on Data Mining, ICDM, IEEE, 2023, pp. 1349–1354.

[246] E. Choi, M.T. Bahadori, J. Sun, J. Kulas, A. Schuetz, W. Stewart, Retain: An interpretable predictive model for healthcare using reverse time attention mechanism, Adv. Neural Inf. Process. Syst. 29 (2016).

[247] K. Yang, Y. Xu, P. Zou, H. Ding, J. Zhao, Y. Wang, B. Xie, KerPrint: local-global knowledge graph enhanced diagnosis prediction for retrospective and prospective interpretations, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, 2023, pp. 5357–5365, 4.

[248] C. Lu, C.K. Reddy, Y. Ning, Self-supervised graph learning with hyperbolic embedding for temporal health event prediction, IEEE Trans. Cybern. 53 (4) (2021) 2124–2136.

[249] S. Niu, Q. Yin, Y. Song, Y. Guo, X. Yang, Label dependent attention model for disease risk prediction using multimodal electronic health records, in: 2021 IEEE International Conference on Data Mining, ICDM, IEEE, 2021, pp. 449–458.

[250] L. Cui, S. Biswal, L.M. Glass, G. Lever, J. Sun, C. Xiao, CONAN: complementary pattern augmentation for rare disease detection, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, 2020, pp. 614–621, 01.

[251] G. Huang, F. Ma, Trustsleepnet: A trustable deep multimodal network for sleep stage classification, in: 2022 IEEE-EMBS International Conference on Biomedical and Health Informatics, BHI, IEEE, 2022, pp. 01–04.

[252] J. Duan, J. Xiong, Y. Li, W. Ding, Deep learning based multimodal biomedical data fusion: An overview and comparative review, Inf. Fusion (2024) 102536.

[253] X. Chen, H. Xie, Z. Li, G. Cheng, M. Leng, F.L. Wang, Information fusion and artificial intelligence for smart healthcare: A bibliometric study, Inf. Process. Manage. 60 (1) (2023) 103113.

[254] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F.L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, et al., Gpt-4 technical report, 2023, arXiv preprint arXiv:2303.08774.

[255] H. Nori, N. King, S.M. McKinney, D. Carignan, E. Horvitz, Capabilities of gpt-4 on medical challenge problems, 2023, arXiv preprint arXiv:2303.13375.

[256] Y. Peng, S. Yan, Z. Lu, Transfer learning in biomedical natural language processing: An evaluation of BERT and ELMo on ten benchmarking datasets, in: Proceedings of the 18th BioNLP Workshop and Shared Task, 2019, pp. 58–65.

[257] K. Huang, J. Altosaar, R. Ranganath, ClinicalBERT: Modeling clinical notes and predicting hospital readmission, 2019, ArXiv:1904.05342.

[258] A. Yan, J. McAuley, X. Lu, J. Du, E.Y. Chang, A. Gentili, C.-N. Hsu, RadBERT: Adapting transformer-based language models to radiology, Radiol.: Artif. Intell. 4 (4) (2022) e210258.

[259] B. Hao, H. Zhu, I.C. Paschalidis, Enhancing clinical BERT embedding using a biomedical knowledge base, in: 28th International Conference on Computational Linguistics (Coling 2020), 2020.

[260] H. Niu, O.A. Omitaomu, M.A. Langston, M. Olama, O. Ozmen, H.B. Klasky, A. Laurio, M. Ward, J. Nebeker, EHR-BERT: A BERT-based model for effective anomaly detection in electronic health records, J. Biomed. Inform. 150 (2024) 104605.

[261] Y. Li, M. Mamouei, G. Salimi-Khorshidi, S. Rao, A. Hassaine, D. Canoy, T. Lukasiewicz, K. Rahimi, Hi-BEHRT: Hierarchical transformer-based model for accurate prediction of clinical events using multimodal longitudinal electronic health records, IEEE J. Biomed. Heal. Inform. 27 (2) (2022) 1106–1117.

[262] Y. Li, R.M. Wehbe, F.S. Ahmad, H. Wang, Y. Luo, Clinical-longformer and clinical-bigbird: Transformers for long clinical sequences, 2022, arXiv preprint arXiv:2201.11838.

[263] L.R. Soenksen, Y. Ma, C. Zeng, L. Boussioux, K. Villalobos Carballo, L. Na, H.M. Wiberg, M.L. Li, I. Fuentes, D. Bertsimas, Integrated multimodal artificial intelligence framework for healthcare applications, NPJ Digit. Med. 5 (1) (2022) 149.

[264] K. Hur, J. Oh, J. Kim, J. Kim, M.J. Lee, E. Cho, S.-E. Moon, Y.-H. Kim, L. Atallah, E. Choi, Genhpf: General healthcare predictive framework for multi-task multi-source learning, IEEE J. Biomed. Heal. Inform. (2023).

[265] N. Allen, C. Sudlow, P. Downey, T. Peakman, J. Danesh, P. Elliott, J. Gallacher, J. Green, P. Matthews, J. Pell, et al., UK biobank: Current status and what it means for epidemiology, Heal. Policy Technol. 1 (3) (2012) 123–126.

[266] S. Padmanabhan, Handbook of Pharmacogenomics and Stratified Medicine, Academic Press, 2014.

[267] E. Bowton, J.R. Field, S. Wang, J.S. Schildcrout, S.L. Van Driest, J.T. Delaney, J. Cowan, P. Weeke, J.D. Mosley, Q.S. Wells, et al., Biobanks and electronic medical records: enabling cost-effective research, Sci. Transl. Med. 6 (234) (2014) 234cm3–234cm3.

[268] M.D. Ritchie, J.C. Denny, R.L. Zuvich, D.C. Crawford, J.S. Schildcrout, L. Bastarache, A.H. Ramirez, J.D. Mosley, J.M. Pulley, M.A. Basford, et al., Genome-and phenome-wide analyses of cardiac conduction identifies markers of arrhythmia risk, Circulation 127 (13) (2013) 1377–1385.

[269] Y. Veturi, A. Lucas, Y. Bradford, D. Hui, S. Dudek, E. Theusch, A. Verma, J.E. Miller, I. Kullo, H. Hakonarson, et al., A unified framework identifies new links between plasma lipids and diseases from electronic medical records across large-scale cohorts, Nature Genet. 53 (7) (2021) 972–981.

[270] L.D. Pena, Y.-H. Jiang, K. Schoch, R.C. Spillmann, N. Walley, N. Stong, S. Rapisardo Horn, J.A. Sullivan, A. McConkie-Rosell, S. Kansagra, et al., Looking beyond the exome: A phenotype-first approach to molecular diagnostic resolution in rare and undiagnosed diseases, Genet. Med. 20 (4) (2018) 464–469.

[271] J. Zhao, Q. Feng, W.-Q. Wei, Integration of omics and phenotypic data for precision medicine, Syst. Med. (2022) 19–35.

[272] S. Lessard, M. Chao, K. Reis, M. Beauvais, D.K. Rajpal, J. Sloane, P. Palta, K. Klinger, E. de Rinaldis, K. Shameer, et al., Leveraging large-scale multi-omics evidences to identify therapeutic targets from genome-wide association studies, BMC Genomics 25 (2024) 1111.

[273] Y. Su, D. Chen, D. Yuan, C. Lausted, J. Choi, C.L. Dai, V. Voillet, V.R. Duvvuri, K. Scherler, P. Troisch, et al., Multi-omics resolves a sharp disease-state shift between mild and moderate COVID-19, Cell 183 (6) (2020) 1479–1495.

[274] J. Park, M.G. Levin, C.M. Haggerty, D.N. Hartzel, R. Judy, R.L. Kember, N. Reza, R.G. Center, M.D. Ritchie, A.T. Owens, et al., A genome-first approach to aggregating rare genetic variants in LMNA for association with electronic health record phenotypes, Genet. Med. 22 (1) (2020) 102–111.

[275] T.G. Drivas, A. Lucas, X. Zhang, M.D. Ritchie, Mendelian pathway analysis of laboratory traits reveals distinct roles for ciliary subcompartments in common disease pathogenesis, Am. J. Hum. Genet. 108 (3) (2021) 482–501.

[276] X. Guo, Y. Song, S. Liu, M. Gao, Y. Qi, X. Shang, Linking genotype to phenotype in multi-omics data of small sample, BMC Genomics 22 (2021) 1–11.

[277] Y. Wu, L. Xie, AI-driven multi-omics integration for multi-scale predictive modeling of genotype-environment-phenotype relationships, Comput. Struct. Biotechnol. J. (2025).

[278] J. Park, A.M. Lucas, X. Zhang, K. Chaudhary, J.H. Cho, G. Nadkarni, A. Dobbyn, G. Chittoor, N.S. Josyula, N. Katz, et al., Exome-wide evaluation of rare coding variants using electronic health records identifies new gene–phenotype associations, Nature Med. 27 (1) (2021) 66–72.

[279] J. Wu, W. Gan, Z. Chen, S. Wan, P.S. Yu, Multimodal large language models: A survey, in: 2023 IEEE International Conference on Big Data (BigData), IEEE, 2023, pp. 2247–2256.

[280] X. Hu, L. Gu, Q. An, M. Zhang, L. Liu, K. Kobayashi, T. Harada, R.M. Summers, Y. Zhu, Expert knowledge-aware image difference graph representation learning for difference-aware medical visual question answering, in: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2023, pp. 4156–4165.

[281] S. Bannur, S. Hyland, Q. Liu, F. Perez-Garcia, M. Ilse, D.C. Castro, B. Boecking, H. Sharma, K. Bouzid, A. Thieme, et al., Learning to exploit temporal structure for biomedical vision-language processing, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 15016–15027.

[282] J. Ma, Q. Lin, Z. Jia, M. Feng, ST-uSleepNet: A spatial-temporal coupling prominence network for multi-channel sleep staging, 2024, arXiv preprint arXiv:2408.11884.

[283] Y. Ye, Y. Xie, J. Zhang, Z. Chen, Q. Wu, Y. Xia, Continual self-supervised learning: Towards universal multi-modal medical data representation learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 11114–11124.

[284] X. Zhang, C. Wu, Y. Zhang, W. Xie, Y. Wang, Knowledge-enhanced visual-language pre-training on chest radiology images, Nat. Commun. 14 (1) (2023) 4542.

[285] P. Chandak, K. Huang, M. Zitnik, Building a knowledge graph to enable precision medicine, Sci. Data 10 (1) (2023) 67.

[286] Q. Lin, J. Liu, R. Mao, F. Xu, E. Cambria, TECHS: Temporal logical graph networks for explainable extrapolation reasoning, in: Proceedings of the 61th Annual Meeting of the Association for Computational Linguistics, ACL, vol. 1, Association for Computational Linguistics, Toronto, Canada, 2023, pp. 1281–1293.

[287] Z. Liu, Y. Shi, A. Zhang, E. Zhang, K. Kawaguchi, X. Wang, T.-S. Chua, Rethinking tokenizer and decoder in masked graph modeling for molecules, Adv. Neural Inf. Process. Syst. 36 (2024).

[288] W. Kopp, A. Akalin, U. Ohler, Simultaneous dimensionality reduction and integration for single-cell ATAC-seq data using deep learning, Nat. Mach. Intell. 4 (2) (2022) 162–168.

[289] A. Chattopadhyay, T.-P. Lu, Gene-gene interaction: The curse of dimensionality, Ann. Transl. Med. 7 (24) (2019) 813.

[290] H. Zhang, L. Klareskog, A. Matussek, S.M. Pfister, M. Benson, Translating genomic medicine to the clinic: Challenges and opportunities, Genome Med. 11 (2019) 1–3.

[291] F.S. Collins, H. Varmus, A new initiative on precision medicine, N. Engl. J. Med. 372 (9) (2015) 793–795.

[292] Ž. Avsec, M. Weilert, A. Shrikumar, S. Krueger, A. Alexandari, K. Dalal, R. Fropf, C. McAnany, J. Gagneur, A. Kundaje, et al., Base-resolution models of transcription-factor binding reveal soft motif syntax, Nature Genet. 53 (3) (2021) 354–366.

[293] L. Li, Y. Fan, M. Tse, K.-Y. Lin, A review of applications in federated learning, Comput. Ind. Eng. 149 (2020) 106854.

[294] T.S. Brisimi, R. Chen, T. Mela, A. Olshevsky, I.C. Paschalidis, W. Shi, Federated learning of predictive models from federated electronic health records, Int. J. Med. Inform. 112 (2018) 59–67.

[295] Y. Gu, Y. Cheng, C.P. Chen, X. Wang, Proximal policy optimization with policy feedback, IEEE Trans. Syst. Man, Cybern.: Syst. 52 (7) (2021) 4600–4610.

[296] R.S. Antunes, C. André da Costa, A. Küderle, I.A. Yari, B. Eskofier, Federated learning for healthcare: Systematic review and architecture proposal, ACM Trans. Intell. Syst. Technol. (TIST) 13 (4) (2022) 1–23.

[297] T.K. Dang, X. Lan, J. Weng, M. Feng, Federated learning for electronic health records, ACM Trans. Intell. Syst. Technol. (TIST) 13 (5) (2022) 1–17.

[298] X. Zhang, R. Mao, E. Cambria, SenticVec: Toward robust and human-centric neurosymbolic sentiment analysis, in: Findings of the Association for Computational Linguistics: ACL, Association for Computational Linguistics, Bangkok, Thailand, 2024, pp. 4851–4863.

[299] T. Shin, Y. Razeghi, R.L. Logan IV, E. Wallace, S. Singh, AutoPrompt: Eliciting knowledge from language models with automatically generated prompts, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP, 2020, pp. 4222–4235.

[300] K. He, Y. Huang, R. Mao, T. Gong, C. Li, E. Cambria, Virtual prompt pre-training for prototype-based few-shot relation extraction, Expert Syst. Appl. (ISSN: 0957-4174) 213 (2023) 118927.

[301] J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q.V. Le, D. Zhou, et al., Chain-of-thought prompting elicits reasoning in large language models, Adv. Neural Inf. Process. Syst. 35 (2022) 24824–24837.

[302] Z. Zhang, A. Zhang, M. Li, G. Karypis, A. Smola, et al., Multimodal chain-of-thought reasoning in language models, Trans. Mach. Learn. Res. (2023).

[303] S. Yao, D. Yu, J. Zhao, I. Shafran, T. Griffiths, Y. Cao, K. Narasimhan, Tree of thoughts: Deliberate problem solving with large language models, Adv. Neural Inf. Process. Syst. 36 (2024).

[304] M. Besta, N. Blach, A. Kubicek, R. Gerstenberger, M. Podstawski, L. Gianinazzi, J. Gajda, T. Lehmann, H. Niewiadomski, P. Nyczyk, et al., Graph of thoughts: Solving elaborate problems with large language models, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 38, 2024, pp. 17682–17690, 16.

[305] Y. Zhu, C. Ren, Z. Wang, X. Zheng, S. Xie, J. Feng, X. Zhu, Z. Li, L. Ma, C. Pan, EMERGE: Enhancing multimodal electronic health records predictive modeling with retrieval-augmented generation, in: Proceedings of the 33rd ACM International Conference on Information and Knowledge Management, 2024, pp. 3549–3559.

[306] L. Huang, T. Denoeux, P. Vera, S. Ruan, Evidence fusion with contextual discounting for multi-modality medical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2022, pp. 401–411.

[307] L. Huang, S. Ruan, Y. Xing, M. Feng, A review of uncertainty quantification in medical image analysis: probabilistic and non-probabilistic methods, Med. Image Anal. (2024) 103223.

[308] Z.M. Ibrahim, D. Bean, T. Searle, L. Qian, H. Wu, A. Shek, Z. Kraljevic, J. Galloway, S. Norton, J.T. Teo, et al., A knowledge distillation ensemble framework for predicting short-and long-term hospitalization outcomes from electronic health records data, IEEE J. Biomed. Heal. Inform. 26 (1) (2021) 423–435.

[309] Q. Lin, R. Mao, J. Liu, F. Xu, E. Cambria, Fusing topology contexts and logical rules in language models for knowledge graph completion, Inf. Fusion 90 (2023) 253–264.

[310] Z. Chen, M. Varma, J.-B. Delbrouck, M. Paschali, L. Blankemeier, D. Van Veen, J.M.J. Valanarasu, A. Youssef, J.P. Cohen, E.P. Reis, et al., CheXagent: Towards a foundation model for chest X-Ray interpretation, in: AAAI 2024 Spring Symposium on Clinical Foundation Models.

[311] F. Bai, Y. Du, T. Huang, M.Q.-H. Meng, B. Zhao, M3d: Advancing 3d medical image analysis with multi-modal large language models, 2024, arXiv preprint arXiv:2404.00578.