Deep Learning for Health Time Series: Physiological Signals, Clinical Records, and Epidemic Dynamics*

LIGE ZHANG† and ZIWEI HUANG*, Emory University, USA
XIAODA WANG, Emory University, USA
GELEI XU, University of Notre Dame, USA
XIAO HU, Emory University, USA
CARL YANG, Emory University, USA
WEI JIN‡, Emory University, USA

Health time series (HTS) provide a fundamental representation of dynamic health processes, spanning organ-level physiological signals, patient-level clinical trajectories, and population-level epidemic dynamics. Despite their shared temporal nature, existing surveys often study these categories separately, leaving their methodological connections underdeveloped. In this survey, we provide a unified review of deep learning for HTS, organized around a scarcity-driven methodological development trajectory. Specifically, we examine how label scarcity motivates task-specific supervised learning with domain-aware temporal modeling; how sample scarcity motivates foundation model development and generative augmentation; and how information scarcity motivates multi-modal learning by incorporating complementary clinical information. Finally, we summarize representative healthcare applications and open research directions. By connecting data categories, scarcity challenges, and methodological paradigms, this survey provides a structured roadmap for understanding and developing deep learning methods for HTS.

Additional Key Words and Phrases: Temporal Data Mining, Healthcare AI, Health Time Series, Deep Learning, Foundation Models, Generative Models, Multi-modal Learning, Physiological Signals, Clinical Records, Epidemic Dynamics

1 Introduction

Recent advances in deep learning and artificial intelligence (AI) have revolutionized healthcare systems in multiple aspects. Bio-sequence models [2, 21] have accelerated drug discovery; Electronic Health Records (EHR) models [194] and medical imaging models [156, 160] have improved disease diagnosis. However, many healthcare problems cannot be fully captured by static observations alone. Health processes unfold over time, from organ-level *physiological signals* [126] to patient-level *clinical records* [122] and population-level *epidemic dynamics* [116]. **Health time series** (HTS) capture this temporal layer across three scales and are becoming a central component of the healthcare AI landscape.

While physiological signals, clinical records, and epidemic dynamics are typically studied by separate research communities, they confront three structurally common forms of scarcity that shape the deep learning methodology across these domains. First, **label scarcity** arises because high-quality expert annotations or clinical outcomes are expensive, delayed, or noisy. This scarcity motivates task-specific supervised methods that use domain-aware temporal modeling to compensate for limited labels, such as physiological morphology [266], frequency patterns [49], irregular clinical observations [283], and spatially coupled epidemic dynamics [239]. Second, **sample scarcity** arises because

*Preprint, under review.
†Both authors contributed equally to this research.
‡Corresponding author.

Authors' Contact Information: Lige Zhang, lige.zhang@emory.edu; Ziwei Huang, ziwei.huang@emory.edu, Emory University, Atlanta, Georgia, USA; Xiaoda Wang, Emory University, Atlanta, Georgia, USA, xiaoda.wang@emory.edu; Gelei Xu, University of Notre Dame, Notre Dame, Indiana, USA, gxu4@nd.edu; Xiao Hu, Emory University, Atlanta, Georgia, USA, xiao.hu@emory.edu; Carl Yang, Emory University, Atlanta, Georgia, USA, j.carlyang@emory.edu; Wei Jin, Emory University, Atlanta, Georgia, USA, wei.jin@emory.edu.

large-scale health datasets are difficult to collect, share, and harmonize, especially for rare diseases, uncommon clinical events, and newly emerging epidemics [80, 125]. Such limitations make it difficult to train robust models from scratch in each target setting, thereby motivating methods that reuse or expand limited data, including foundation model pre-training, adaptation, and generative augmentation to improve cross-domain generalization. Third, **information scarcity** arises because a time series alone often provides only a partial view of the underlying health state [16, 55, 203]. This scarcity motivates multi-modal learning methods that incorporate complementary information such as clinical notes, demographic variables, spatial context, and other non-temporal knowledge. These challenges jointly shape a methodological trajectory in which each paradigm addresses a specific bottleneck.

Existing HTS surveys, as summarized in Table 2 in Section 7, have mainly been organized around specific data sources, application scenarios, or methodological families. Some surveys focus on particular types of health-related temporal data, such as physiological signals [6, 171], clinical record trajectories [7, 166, 223, 264], and epidemic surveillance data [149, 201]. Others focus on specific healthcare applications, such as emotion recognition [98, 184], anomaly detection [271], and epidemic outbreak detection and forecasting [15, 209]. Another line of surveys examines particular methodological directions, including self-supervised learning, contrastive learning [145, 259], few-shot learning [125], foundation models [77], large language models (LLMs) [144], and generative models [150]. Despite these valuable efforts, existing surveys remain **fragmented** in scope. Most of them examine HTS from a single perspective, making it difficult to compare recurring modeling challenges across HTS types or identify transferable ideas across research communities. Thus, a unified view is needed to reveal both the **shared modeling strategies** of HTS and the natural **methodological development trajectory** through which the field has responded to diverse scarcity challenges.

To address this fragmentation, we present, to the best of our knowledge, **the first unified survey** of deep learning for HTS, organized around a scarcity-driven methodological trajectory across physiological, clinical records, and epidemic HTS categories. Instead of reviewing the field separately by data type, application, or model family, we use label scarcity, sample scarcity, and information scarcity as the main spine to connect methodological developments across HTS. This perspective highlights how methods developed for one HTS may inspire solutions for another, while also clarifying the domain-specific adaptations required by different observational levels. Concretely, our contributions are: **(i)** a fine-grained taxonomy of HTS across three data categories and three core task families (diagnosis, prognosis, decision support); **(ii)** a unified *three-scarcity analysis* linking scarcity conditions to the modeling requirements of different HTS categories; and **(iii)** a unified *methodological review* that connects task-specific supervised learning, foundation-model (pre-training, adaptation, generative data synthesis), and multi-modal learning into a single development trajectory instead of three disconnected literatures.

The remainder of the survey is organized as follows. Section 2 introduces the taxonomy of HTS, including data categories, key task formulations, basic deep learning models for time series modeling, and detailed illustrations of the main challenges. Section 3 reviews task-specific supervised methods for HTS, highlighting how domain-specific temporal units are modeled and how shared temporal representation learning strategies are adapted across HTS. Section 4 reviews foundation models for HTS from a unified perspective, including pre-training, transfer adaptation, and the use of generative models to compensate sample scarcity. Section 5 reviews multi-modal learning methods that incorporate complementary information beyond temporal signals for context enrichment, and what specific modality information is utilized for each HTS category. Section 6 further discusses representative applications in these different data categories. Section 7 summarizes our survey construction process, including paper collection, filtering, and selection criteria. Section 8 discusses open challenges and future research directions. Finally, Section 9 concludes the survey.

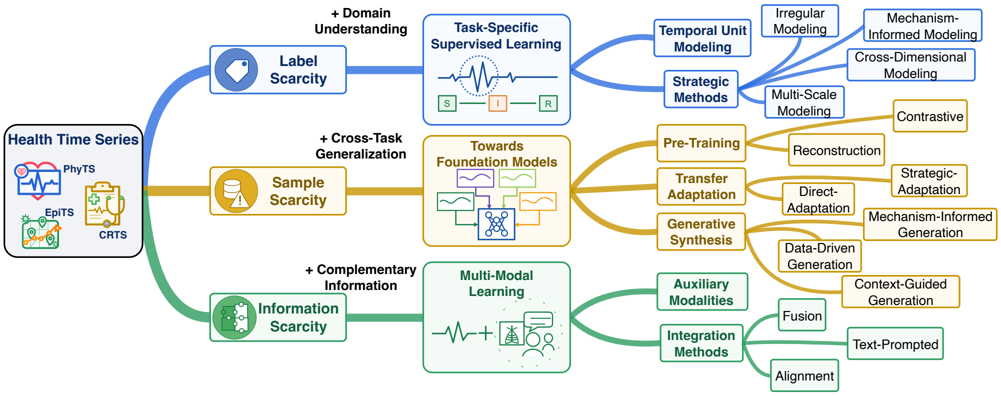

Fig. 1. Three scarcity challenges for HTS modeling and the induced learning paradigms.

2 Background

2.1 Health Time Series

A time series is a temporally ordered sequence of observations that records how one or more variables evolve over time. Formally, it can be represented as $X = \{(t_i, x_i)\}_{i=1}^{N}$, where t_i denotes the observation time, $x_i \in \mathbb{R}^d$ denotes a d-dimensional observation recorded at that time, and N is the number of observations. In regularly sampled settings, this reduces to the compact form $X = \{x_t\}_{t=1}^{T} \in \mathbb{R}^{T\times d}$, where each index t corresponds to a fixed sampling interval. In healthcare, such sequences may describe physiological processes, patient health states, clinical care trajectories, or population-level health dynamics, arising from settings such as wearable and bedside monitoring, structured EHR, and epidemic surveillance. Compared with generic time series, HTS are often noisy, partially observed, irregularly sampled, and heterogeneous across individuals, devices, institutions, and environments, while labels may be limited, delayed, or imperfect, and prediction errors may carry clinical or public-health consequences.

2.2 Data Categories

We group HTS data into three broad categories based on the level at which temporal observations are collected: *physiological time series* (PhyTS), *clinical record time series* (CRTS), and *epidemic time series* (EpiTS). Specifically, PhyTS describes organ-level biological activity, CRTS describes patient-level health states, and EpiTS describes population-level infectious disease dynamics. They typically exhibit distinct properties and hence require different modeling choices.

PhyTS. They are temporally ordered bio-signals that reflect the states of organisms over time, usually collected as raw or minimally processed signals by wearable sensors, diagnostic devices, or continuous monitoring systems [208, 238, 247]. Examples include ECG, EEG, PPG, EMG, EOG, and other continuous monitoring signals. In this category, each timestamp usually corresponds to a direct signal measurement, and much of the useful information lies in waveform morphology, oscillatory structure, and multi-scale temporal patterns [102, 252]. As a result, PhyTS are often characterized by high temporal resolution, strong local continuity, and rich short-range dynamics. From a data perspective, PhyTS often provide abundant raw recordings, but high-quality annotations remain much more limited [76, 145]. However, since these signals are frequently collected through wearable or non-invasive sensing devices, any motion artifacts or environmental interference can lead to measurement noise [16, 188, 297]. Their modeling is further complicated by cross-device discrepancies and inter-subject biological variability [59, 259, 299].

CRTS. They consist of temporally ordered patient-level records collected during clinical care to characterize patient states, disease progression, and care processes. Unlike PhyTS, which often reies on raw or high-resolution bio-signals, CRTS are constructed from structured clinical variables and care events documented in electronic health records or bedside charting systems. These include vital signs summaries, laboratory tests, medications, procedures, interventions, clinical scores, and other observations documented during hospitalization or follow-up care [95, 187]. These records are typically collected only when patients visit hospitals or undergo examinations, resulting in irregularly timed and incompletely observed patient trajectories [146, 218, 225, 286]. Together with heterogeneous variable types, privacy restrictions, and delayed or noisy outcome labels, these properties make CRTS challenging to model [258, 269].

EpiTS. They are population-level HTS that capture the temporal dynamics of infectious diseases across regions, communities, or entire populations. Each timestamped observation usually represents an aggregated regional indicator over a reporting period rather than an individual-level signal or event [54]. Typical examples include epidemic case counts, mortality trends, influenza-like illness rates, and surveillance indicators reported daily or weekly across locations [196]. These series are often coarse in temporal resolution and are strongly influenced by seasonality, policy shifts, reporting practices, mobility, and other exogenous factors, leading to substantial non-stationarity [116]. Furthermore, the amount of historical data is usually limited, especially for emerging outbreaks or newly identified diseases. The target variables to be predicted are often noisy, as reports can be delayed, later corrected, under-reported, or defined differently over time [17, 222, 235].

2.3 Core Task Formulations

Although the tasks for different HTS vary in their specific application contexts, they share common modeling objectives. We therefore organize HTS tasks into three broad categories: (a) **Diagnosis**, which infers the current health state from observed temporal evidence; (b) **Prognosis**, which predicts future outcomes, risks, or trajectories from historical observations; and (c) **Decision Support**, which evaluates, recommends, or optimizes actions that benefit healthcare. We provide a detailed description of each category below, and review detailed applications in Section 6.

Diagnosis. It aims to estimate the current health state represented by an observed time series window. Its concrete meaning varies across data categories: in PhyTS, it often involves identifying abnormal signal patterns or physiological states, such as arrhythmia diagnosis and sleep stage classification [80, 120]; in CRTS, it focuses on recognizing patient states from longitudinal clinical observations, such as acute care phenotyping and disease classification [41, 194, 218]; and in EpiTS, it corresponds to epidemic surveillance, where models track the current epidemic status and detect outbreak signals [57]. Across these scenarios, diagnostic tasks can be further categorized according to the structure of the target labels: *(i) multi-class diagnosis*, which distinguishes among mutually exclusive conditions or disease categories [80, 252]; *(ii) multi-label diagnosis*, which infers multiple co-occurring conditions simultaneously [218, 295]; and *(iii) sequential diagnosis*, which predicts diagnostic states for consecutive temporal segments [120, 167].

Prognosis. It aims to predict future health outcomes, risks, or trajectories based on historical time series data. In PhyTS it often involves predicting future physiological risks or signal trajectories, such as cardiovascular risk prediction and blood pressure prediction [16, 290]; in CRTS, it focuses on forecasting patient outcomes from longitudinal clinical observations, such as mortality prediction and survival analysis [202, 255]; and in EpiTS, it corresponds to epidemic forecasting, where models predict future epidemic trends [148, 239]. Across these scenarios, prognostic tasks can also be categorized into three types: *(i) Event risk prediction*, which estimates the probability of a future adverse event [16, 193]; *(ii) Trajectory forecasting*, which predicts the future temporal evolution of physiological, clinical, or epidemiological variables [239, 261]; and *(iii) Time-to-event prediction*, which estimates when a future event may occur [255].

Decision Support. It aims to support health decision processes by reasoning over observed time series and evaluating or recommending possible interventions. In PhyTS, it involves physiological reasoning that assists users in understanding bio-signals collected from monitoring devices [66, 162]; in CRTS, it focuses on longitudinal clinical reasoning and decision support [46]; and in EpiTS, it corresponds to public health policy support, such as intervention evaluation and contact tracing [67, 97, 219]. According to the modeling objective, this task can be further categorized into: *(i) Policy learning*, which directly optimizes action strategies, often through reinforcement learning [67, 219]; *(ii) Causal reasoning*, which estimates the potential effects of interventions or policies [97]; and *(iii) Language-based temporal Q&A*, which uses Large Language Models (LLMs) to summarize temporal evidence and answer user questions [256]. Compared to diagnosis and prognosis, this category remains less mature in HTS applications.

2.4 Deep Neural Architectures for Time Series Modeling

Deep learning models for HTS often require a temporal encoder that transforms raw temporal observations into latent representations. Despite their distinct temporal structures, PhyTS, CRTS, and EpiTS are often modeled by similar encoder families. Given a HTS $X = \{(t_i, x_i)\}_{i=1}^N$, a temporal encoder learns a mapping $H = f_\theta(X)$, where $H = \{h_i\}_{i=1}^N$ denotes time-indexed hidden representations, which are then used for diverse tasks with some adaptation modules. Different architecture families instantiate f_θ with different inductive biases. Feed-forward and MLP-style modules perform temporal, variable-wise, or patch-wise mixing through stacked nonlinear transformations [178, 241], providing efficient representation learning when observations can be arranged into fixed-length inputs. Recurrent models construct representations through sequential state transitions, making the hidden state explicitly dependent on the order of observations and enabling extensions that incorporate elapsed time, masking, or missingness [19, 33, 41, 245]. Convolutional models use local filtering and hierarchical composition to capture short-range motifs and multi-scale temporal patterns [80, 117, 182]. Transformer-based models use an attention mechanism to model long-range dependencies and interactions among time steps, variables, or channels [126, 183, 218, 252]. Neural-ODE models use continuous-time dynamics to model irregularly sampled sequences and obtain representations at arbitrary timestamps through numerical integration [111, 204, 239, 241].

2.5 The Scarcity-Driven Challenge in HTS Modeling

PhyTS, CRTS, and EpiTS differ substantially in their surface properties, such as temporal scale, sampling process, and application context. However, they are shaped by three common forms of scarcity: label scarcity, sample scarcity, and information scarcity. Unlike time series analysis in classical domains such as electricity load prediction [176], traffic flow modeling [89], and weather forecasting [114], these scarcity challenges are especially central to HTS modeling. We catalog these challenges below and use them as the organizing spine of the rest of the survey: each subsequent methodological section is framed as a response to one or more of these scarcity types. The resulting taxonomy and the organizational spine of this survey are summarized in Figure 1. To further illustrate the HTS development trajectory following the spine, we provide an additional timeline-style summary in Figure 2.

Label Scarcity. Label scarcity refers to the lack of high-quality annotations for observed HTS. For PhyTS, reliable labels often require expert interpretation from trained clinicians, such as cardiologists for ECG diagnosis or specialists for sleep staging, making annotation expensive and time-consuming [80, 120]. For CRTS, many labels are delayed, which requires long-term observation and careful follow-up [213]. For EpiTS, labels for outbreak status or disease burden can also be delayed or biased by reporting practices and testing availability. This scarcity, therefore, motivates

Table 1. Taxonomy of supervised modeling techniques and their relevance to different HTS categories.

Modeling Technique	PhyTS	CRTS	EpiTS
Basic temporal unit modeling	Physiological signal segments with informative local waveform morphology, rhythm, or spectral patterns [80, 117, 183, 199].	Heterogeneous clinical events, visits, and measurements ordered along patient histories [41, 133, 181, 194].	Epidemic surveillance observations, such as reported cases or disease activity over reporting windows [11, 129, 261].
Multi-scale modeling	Multi-timescale physiological structures, from cycle-level patterns to long-range rhythms [39, 242, 257].	Multi-timescale clinical evolution, from short-term vital changes to long-term disease progression [123, 153, 284].	Spatially multi-scale epidemic dynamics across hierarchical geographic units, such as states and countries [163, 227, 291].
Cross-dimensional interaction modeling	Multi-channel signal interactions, such as multi-lead ECG or multi-electrode EEG signals [9, 26, 91, 117, 131, 268].	Heterogeneous clinical variable interactions, such as vitals, laboratory tests, medications, and interventions [124, 134, 143, 152].	Spatial interactions among geographically or mobility-connected epidemic regions [50, 99, 239, 261].
Irregularity modeling	Less concerned, since signals are acquired at fixed sampling rates within fixed-length windows.	Irregular timestamps and incomplete clinical observations [33, 56, 126, 136, 181, 229].	Uneven surveillance records and irregular reporting intervals [111, 239, 244].
Mechanism-informed modeling	Morphological priors and mechanistic physiological ODEs [161, 211, 248, 289, 294].	Clinical-process priors, such as disease-stage and knowledge-guided structures [4, 36, 152].	Epidemic transmission processes guided by mechanistic epidemic ODEs [111, 200, 239, 241, 244].

the development of task-specific supervised methods that compensate for limited labels by incorporating domain knowledge, temporal inductive biases, and label-efficient learning strategies.

Sample Scarcity. Sample scarcity refers to the limited availability of representative samples for a target task or disease event. Even when labels are easily obtainable, the number of relevant target-domain examples may still be too small to train robust models from scratch. For PhyTS, sample scarcity may arise from rare diseases, underrepresented populations, or device-specific settings. For CRTS, large-scale datasets are difficult to collect and harmonize across institutions due to privacy constraints, cohort differences, and variable measurement practices [198, 223].For EpiTS, newly emerging diseases or local outbreaks often provide only short historical observation windows. Such sample scarcity has motivated two research directions: foundation models that pre-train on broader unlabeled data and transfer to sample-limited target domains, and generative models that synthesize samples for data augmentation.

Information Scarcity. Information scarcity refers to the fact that HTS alone often provides only a partial view of the underlying health state. For PhyTS, the same temporal pattern may correspond to different physiological conditions depending on patient context or clinical history, and thus may require further justification from reports about the patient's physical condition [138, 270]. For CRTS, pure longitudinal records may not fully capture the patient's disease status, since important information can be recorded in complementary sources such as medical images [82]. For EpiTS, observed case counts or incidence curves alone may be insufficient to explain disease spread, as epidemic dynamics are also shaped by mobility patterns, demographic structures, and policy interventions [55]. This scarcity motivates the development of multi-modal learning methods that integrate temporal signals with complementary information to obtain accurate estimation of underlying health state.

3 Task-Specific Supervised Learning

While end-to-end supervised learning is effective when labeled data is abundant, healthcare labels are often scarce and costly to obtain. Consequently, the effectiveness of HTS models depends heavily on the design of inductive biases that enable the extraction of informative patterns from limited data. Therefore, this section reviews task-specific supervised learning and the related key temporal modeling techniques, with a brief summary in Table 1.

3.1 Basic Temporal Unit Modeling

Learning informative patterns from temporal observations is the foundation of supervised HTS modeling. Different HTS are characterized by different temporal units and structures, which motivate distinct approaches to temporal encoding. For **PhyTS**, the basic units are dense signal segments, where informative structures are typically expressed as waveform morphology (e.g., QT prolongation and ST-segment elevation in ECG [266]), rhythm (e.g., cardiac-cycle-induced periodicity in ECG/PPG [248, 250, 262]), or spectral composition (e.g., spectral patterns associated with brain activity in EEG [117]). Accordingly, supervised PhyTS models often first extract local morphological or frequency-domain features [20, 80, 117, 199] and then model longer-range rhythm or state transitions [80, 131, 183, 199, 210]. In **CRTS**, the basic units are semantically heterogeneous clinical events or visits. Temporal encoding, therefore, focuses on representing these units and aggregating them along ordered patient histories to capture patient-state evolution. Early methods instantiate this paradigm with convolutional [169], recurrent [41], or attention-based encoders [218] over medical events and visits. Transformer-based methods further treat clinical events or visits as structured tokens, combining medical-code embeddings with positional, age, visit, or time-related embeddings to produce patient-level representations for downstream prediction [133, 181, 194]. In **EpiTS**, the basic units are surveillance observations indexed by region and reporting time, including reported cases, deaths, or influenza-like illness rates. Temporal encoding aggregates these units along historical reporting windows to capture epidemic trajectories. Early methods use recurrent, convolutional, or hybrid sequence encoders to model evolving epidemic patterns [11, 261], while transformer-based methods are also used to capture long-range temporal dependencies [129].

3.2 Strategic Modeling Methods

Multi-Scale Modeling. Multi-scale modeling extends basic temporal structure modeling by learning HTS at multiple resolutions. Here, “scale” may refer to temporal resolution, such as short-window patterns versus long-range patterns, or spatial resolution, such as local regions versus broader geographic regions. The main objective is to integrate fine-grained local signals with coarse-grained global patterns, which is especially important when health-related dynamics unfold hierarchically across time or space. In **PhyTS**, multi-scale modeling mainly addresses the hierarchical structure of physiological signals, where diverse resolutions of the waveform (e.g., local individual cardiac cycles and global cardiac rhythm in ECG) may provide complementary information. To capture such structures, models extend basic encoders into multi-resolution architectures that operate over multiple temporal windows or signal scales, as shown in representative works like PulseID [257], MSDCGTNet [39], and MELP [242]. **CRTS** also exhibit temporal dependencies at multiple temporal scales, ranging from rapid changes in vital signs and laboratory measurements to long-term disease progression across visits. Existing methods introduce multi-scale structures in different ways. First, progression-level abstraction represents longitudinal records through coarse disease stages or health states [68]. Second, temporal-feature-level modeling captures short- and long-term variations in clinical variables through scale-adaptive feature extraction [157] or local-global temporal attention [151]. Third, more recent methods construct explicit multi-resolution representations, using patch-based organization to group fine-grained observations into local temporal units and integrate them through inter-patch modeling [123, 153, 284]. In **EpiTS**, multi-scale modeling is more focused on a spatial level. Epidemic processes are organized over hierarchical geographic units, such as cities, states, and countries, where local outbreaks, regional spread, and national-level trends are dynamically coupled. Therefore, some methods incorporate multi-scale spatial context to model disease spread across different levels of granularity [227], while others explicitly encode geographical hierarchies to produce forecasts that remain coherent across spatial resolutions [163, 291].

Cross-Dimensional Interaction Modeling. Beyond the temporal structure within each sequence, HTS models often need to capture interactions among multiple coupled entities. The semantic meaning of these entities differs across HTS categories: signal channels in PhyTS, clinical measurements in CRTS, and geographic regions in EpiTS. Although variate-independent modeling can be effective for general time series [79, 135], HTS dimensions often correspond to meaningful clinical or spatial structures. Thus, modeling cross-dimensional interactions is essential for HTS, as it allows models to integrate information across variables. In **PhyTS**, this problem appears as multi-channel signal modeling. For example, 12-lead ECG records cardiac activity from different lead perspectives [199], while multi-electrode EEG captures brain activity across scalp regions [117]. Existing models capture channel interactions through spatial convolutions [117], graph-based message passing [9, 26, 91], cross-channel attention [131, 195, 268], or implicit channel mixing [25, 182]. In **CRTS**, the same modeling principle corresponds to dependency learning among heterogeneous clinical variables, such as vital signs, laboratory tests, and treatments. Graph-based models represent variables as nodes and learn sample-specific, adaptive, or knowledge-informed relations [134, 152, 288]; more recent methods further relax fixed variable-level interactions by modeling dependencies over aligned temporal patches [284], or higher-order structures [124, 143, 153] under asynchronous and irregular sampling. In **EpiTS**, the interacting dimensions are geographic regions. Disease trajectories in different regions may be coupled through population mobility or geographic proximity. Existing methods typically construct a graph over locations, where each location is represented as a node, and combine graph modules with temporal encoders to capture spatial-temporal propagation patterns [50, 99, 261].

Irregularity Modeling. Many HTS are rarely observed as perfectly regular sequences. Recordings may not be collected at even time intervals, and in some scenarios, only a subset of variables is measured at each time point. This issue is most prominent in EpiTS and CRTS, although it arises from different sources. For **PhyTS**, irregularity is usually less of a central concern because physiological signals are often acquired by devices at fixed sampling rates and fixed-length monitoring windows. For **EpiTS**, irregularity often arises from uneven surveillance records. For example, in the early stages of an outbreak, resource or infrastructure limitations may prevent some regions from conducting routine surveillance, resulting in irregular reporting intervals. Therefore, many EpiTS models [111, 239, 244] have tried to convert discrete and irregular surveillance observations into continuously evolving disease dynamics, mainly using neural differential equations [37, 104, 204]. We discuss ODE-based EpiTS modeling in more detail in the mechanism-informed modeling paragraph. For **CRTS**, irregularity also arises since clinical measurements are often recorded at uneven time intervals, and variables are measured only when clinically necessary or feasible. To address *irregular timing*, one line of work adds timestamp-derived quantities to input representations, such as discretized time-gap tokens in Deepr [169], timestamp-augmented visit representations in RETAIN [41], and time-, age-, interval-, or visit-aware token embeddings in EHR Transformers [132, 133, 181, 194, 218, 267, 272]. Another line injects time information into the model computation: GRU-D [33] and T-LSTM [19] use elapsed-time decay for stale observations or memory states, Phased LSTM [168] gates recurrent updates by timestamp, and attention-based models use observed times to define temporal alignment, warping, or timestamp-aware attention [126, 217, 283]. Continuous-time models address the same problem by evolving latent patient states between observations and updating them when new measurements arrive [37, 104, 126, 165, 204]. To address *incomplete observations*, early work incorporates missingness indicators as input features to improve recurrent prediction [136], while GRU-D [33] jointly uses masks and time gaps through learnable decay. Beyond using missingness as auxiliary input, imputation-based methods estimate missing values jointly with representation learning using recurrent [28, 278], attention-based [56], or generative architectures [48, 229].

Mechanism-Informed Modeling. Mechanism-informed modeling incorporates domain knowledge about the underlying health process into supervised HTS models. Compared with generic architectural priors, such knowledge may enter the model as explicit dynamic equations or more implicit inductive biases derived from known dynamic processes. For **PhyTS**, mechanism-informed modeling commonly appears in studies of physiological signal generation or cross-modal physiological transformation [248, 289, 294], where the modeled physiological process is often guided or constrained by corresponding mechanistic ODE systems [161]. A comprehensive discussion of this paradigm is provided in Section 4.3. In contrast, mechanism-informed approaches are less prevalent in supervised PhyTS modeling, where physiological knowledge is usually used implicitly as morphology-aware objectives rather than as explicit mechanistic constraints. For example, DMT [211] exploits waveform patterns associated with arterial stiffness and wave reflection to guide cuffless blood pressure estimation. For **CRTS**, patient trajectories are shaped by heterogeneous diseases, interventions, and care practices, making it difficult and uncommon to describe such dynamics with general mechanistic equations. As a result, CRTS models more often incorporate clinical knowledge as inductive biases in the model design. We summarize three representative forms of such clinical priors below. First, disease-progression priors are utilized to improve the latent-state transition: Attentive SS [4] introduces structured latent disease states with learned transition dynamics, while StageNet [68] augments recurrent updates with stage-aware memory and stage-adaptive convolution, so that hidden states evolve according to inferred disease stages rather than unconstrained recurrent dynamics. Second, event-process priors are utilized to improve the prediction head and training objective: Dynamic-DeepHit [118] predicts discrete-time event-risk distributions under competing risks, and SurvLatent ODE [165] couples continuous-time latent dynamics with survival modeling, replacing ordinary fixed-horizon classification with hazard- or event-time-based prediction. Third, medical-relation priors are utilized to improve the representation or interaction structure: GRAM [40] uses ontology-guided medical-code representations, KEDGN [152] constructs knowledge-guided variable graphs, and TRANS [36] builds temporal patient-event graphs. In these models, clinical knowledge is not merely used to interpret latent states after training, but is built into the transition, output, or relational structure of the model. For **EpiTS**, mechanism-informed modeling is the most mature and widely-used. Epidemic spread can be naturally described by compartmental transmission systems, such as SIR-style ODE models [74, 279]. Recent approaches integrate these mechanistic priors with neural differential equation frameworks [37, 104], allowing neural networks to learn flexible transmission dynamics while preserving continuous-time epidemiological structure. One line of work directly parameterizes mechanistic ODEs with neural networks so that transmission parameters can be learned from data while remaining constrained by epidemic dynamics [111]. These parameters can also be conditioned on spatial context: CausalGNN uses graph neural networks to generate ODE parameters from regional dependencies [244], while EARTH further models a continuously evolving disease transmission graph within a neural-ODE framework [239]. Another line of work uses mechanistic models as training-time guidance rather than as the final predictor. For example, neural predictors can be trained to imitate trajectories produced by ODE-based simulators [241], or to absorb latent epidemic dynamics from a source mechanistic model through gradient matching [200].

4 Towards Foundation Models

In practice, training a model from scratch for each target setting can be impractical due to the *sample scarcity* detailed in Section 2.5. HTS modeling has therefore moved toward foundation models, where a single model is pre-trained on large-scale unlabeled data aggregated across datasets or tasks and then reused for target downstream healthcare applications. Specifically, this section reviews three directions: foundation model pre-training for learning representations, transfer adaptation for reusing pre-trained models, and generative data synthesis for augmenting scarce data samples.

4.1 Foundation Model Pre-training

Foundation model pre-training aims to make use of large-scale unlabeled HTS to learn an encoder f_θ that produces general representations that can be transferred to unseen data or tasks. In this section, we focus on domain-specific foundation models pre-trained for each HTS, organized by pre-training paradigms.

4.1.1 Contrastive Pre-training. This paradigm learns representations by maximizing embedding similarity between related views of the same signal while pushing unrelated ones apart. Formally, given a sample x, two correlated views $x^{(1)}, x^{(2)} \sim \mathcal{T}(x)$ are generated via a transformation $\mathcal{T}$. Each view is then mapped to a latent representation by an encoder f_θ. The model is typically trained using a contrastive objective such as the one used in InfoNCE [177] to distinguish positive pairs from negatives. While this formulation is simple, its effectiveness for HTS foundation models largely depends on how these components are instantiated for specific signals and settings.

Sample Definition. A key design choice in contrastive pre-training is how to define the sample x. Unlike images or text, where a sample is often naturally defined, an HTS sample can be represented at multiple granularities. Existing methods have defined x as an entire recording [109], a temporal window or cropped segment [58, 233], a local patch or tokenized subsequence [92, 268, 284], or an irregular patient trajectory in clinical data [42, 107, 191]. This choice determines the level at which invariance is learned. Instance-level formulations encourage global semantic consistency, whereas segment- or patch-level formulations focus more on local temporal patterns and multi-scale structure. For irregular HTS, defining x as temporally aligned patches or density-aware segments can further facilitate modeling under heterogeneous and asynchronous sampling patterns [42, 284].

Positive Pair Construction. The construction of positive pairs is critical, as it determines the invariance learned by the model. Ideally, transformations used to generate positive pairs should preserve task-relevant semantic information while removing nuisance variations. Early approaches adopt augmentation-based methods such as temporal cropping, scaling, jittering, and masking to encourage robustness to signal perturbations [1, 58, 109, 273]. Beyond such generic transformations, recent works exploit the structural properties of time series to construct more semantically meaningful pairs, including local-global temporal views and temporal neighborhood consistency [58, 233], cross-channel or multi-lead views in PhyTS [25, 109], and multi-scale representations that capture both short- and long-term dynamics [268]. In CRTS, positive pairs are often defined at the intra-subject level based on temporal proximity or trajectory consistency, which helps contrastive learning handle irregular sampling and heterogeneous variables [8, 42, 101, 107, 191].

Negative Sampling and Contrastive Objectives. The definition of negative samples and the formulation of the training objective are also critical for contrastive pre-training. Most existing methods adopt an InfoNCE-style objective [177], which encourages representations of positive pairs to have higher similarity than those of negatives:

$$\mathcal{L}_{\text{con}}(i) = -\log \frac{\exp(\text{sim}(z_i^{(1)}, z_i^{(2)})/\tau)}{\sum_{k \in \mathcal{N}(i)} \exp(\text{sim}(z_i^{(1)}, z_k^{(2)})/\tau)},$$

where $(z_i^{(1)}, z_i^{(2)})$ denotes the positive pair corresponding to sample i, and $\mathcal{N}(i)$ denotes the contrastive candidate set (including the positive sample and a set of negative samples). The similarity function $\text{sim}(\cdot, \cdot)$ is commonly chosen as cosine similarity, and τ is a temperature parameter. In practice, $\mathcal{N}(i)$ is often approximated using samples from the same mini-batch, but more general constructions are also used depending on the task and data structure [58, 273]. However, in HTS modeling, naive negative sampling may introduce false negatives, where semantically similar signals (e.g., from the same subject or similar physiological states) are incorrectly treated as unrelated. To address this, recent work explores subject-aware sampling strategies and hard-negative selection to better reflect the underlying data structure [53, 101].

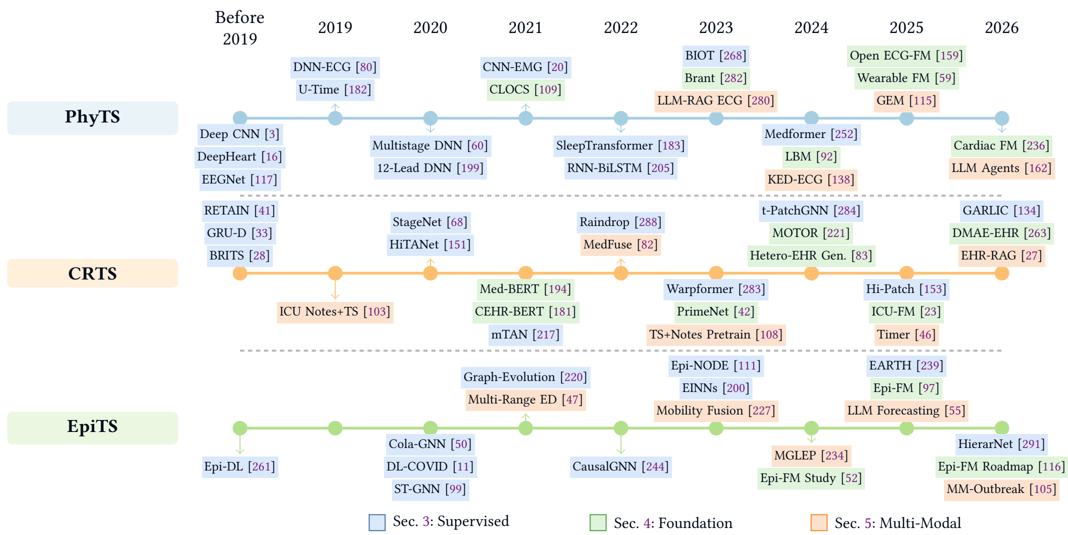

Fig. 2. Timeline of HTS methods across data types. Rows correspond to PhyTS, CRTS, and EpiTS. Citation-box colors indicate the corresponding methodological sections: task-specific supervised learning, foundation models, and multi-modal learning.

Additionally, the contrastive objective can be applied at different granularities, including instance-level, segment-level, and patch-level contrast, which closely interacts with the sample definition and positive pair construction [1, 268].

4.1.2 Reconstruction Pre-training. This paradigm learns representations by recovering masked or corrupted parts of the input signal. A general objective can be written as $\mathcal{L}_{\text{recon}} = \mathbb{E}_X \left[\text{DIST}\left(X_M,\ g_\psi(f_\theta(X_{\setminus M}))\right)\right]$, where M denotes the masked subset, X_M and $X_{\setminus M}$ denotes the reconstruction target signal and observed input after masking. The encoder f_θ maps the observed signal into a latent representation, and the decoder g_ψ reconstructs the masked components. The function $\text{DIST}(\cdot,\cdot)$ then measures the discrepancy between the original and reconstructed signals. In general, such methods encourage the representation to preserve information that is sufficient to recover the original signal.

Mask Unit. A key design choice in reconstruction pre-training is what unit to mask. In time series, masking can be applied at multiple granularities. *(i) Point-level masking* removes individual observed values, such as the measurement of a specific variable at a specific time, and encourages local interpolation and denoising. This strategy is commonly adopted in early CRTS pre-training [42, 181, 231, 267, 272], where sparse and irregular observations make point-level reconstruction a natural objective. *(ii) Segment-level masking* removes continuous regions of the temporal signal, forcing the model to exploit longer-range temporal dependencies and surrounding context. This strategy has been widely adopted in PhyTS modeling [159, 179, 251, 296, 297]. *(iii) Structured-unit masking* represents the time series as discrete or structured units, where each unit may correspond to a local temporal patch, a multi-channel patch, or a structured event. This formulation provides a flexible masking interface across different HTS types. In PhyTS, patch-based masking is used to model temporal dependencies and cross-channel structure [92, 179, 282]. In CRTS, masking can be applied to irregular events, visits, or grouped observations [23, 42, 221].

Mask Position. Beyond the masking unit, the placement of reconstruction targets also matters. Existing methods can be broadly grouped into in-context reconstruction and forecasting-style reconstruction. *(i) In-context reconstruction* follows a BERT-like masked modeling formulation [51, 85], where target units are selected within the observed sequence and reconstructed from the bidirectional context. The key design choice lies in how masking locations are sampled.

CEHR-BERT [181] and TransEHR [267] adopt token-level random masking over discretized clinical events, while PrimeNet [42] further improves masking by sampling fixed-duration temporal regions to mitigate bias from non-uniform sampling density. For PhyTS, masking is typically structured rather than independent: LaBraM [92] masks channel-wise EEG patches, Brant [282] removes contiguous signal patches, and REVE [179] samples spatial-temporal blocks to prevent trivial reconstruction from local correlations. *(ii) Forecasting-style reconstruction* follows a GPT-like next-token-prediction formulation [22], where targets are placed after the observed context and the model is trained to predict future values or events. This objective aligns naturally with prognostic tasks in HTS. TransformEHR [272] predicts next-visit clinical outcomes from historical visits, and MOTOR [221] formulates pretraining as time-to-event prediction over longitudinal records. MIRA [126] extends this paradigm to irregular time series by modeling continuous-time dynamics under heterogeneous sampling. Recent efforts also explore the combination of several masking strategies to improve generalization performance [23, 24, 231].

Reconstruction Objectives. Reconstruction-based methods also differ in the space where the target is defined. *(i) Observation-space reconstruction* directly predicts the masked or future observations, such as clinical events or waveform segments [179, 181, 221, 267, 272, 296]. This objective preserves the semantics of the original data and is often closely aligned with downstream tasks. *(ii) Representation-space reconstruction* predicts latent embeddings, patch representations, or quantized codes instead of raw observations. This strategy is common in models that first tokenize HTS into structured units. For example, LaBraM [92] predicts quantized neural representations for masked EEG patches, and Brant [282] reconstructs patch-level embeddings of neural signals. Compared with input-space reconstruction, this objective promotes more abstract, transferable representations with less reliance on low-level interpolation.

4.2 Transfer Adaptation of Foundation Models

After pre-training, foundation models are rarely deployed directly. Instead, they are adapted to downstream healthcare tasks through various transfer strategies that address domain shifts and task-specific requirements. This section reviews common adaptation approaches and their applications in PhyTS, CRTS, and EpiTS.

4.2.1 General Adaptation Strategies.

Direct Adaptation. Foundation models can be directly applied to downstream tasks with little or no parameter updates, serving as temporal encoders. This strategy relies on the assumption that large-scale pre-training has captured general temporal patterns that can transfer to unseen time series data [116]. The pre-trained models used for direct adaptation can be general time series foundation models trained in various domains, such as Time-MoE [212], Chronos [10], and Moirai-MoE [141]. Recent works have tried to directly adapt them for epidemic forecasting [52, 97]. The pre-trained models can also be health-specific foundation models such as ECG foundation models [59, 127, 159] and EEG foundation models [49, 179, 297]. Another possible direction is to adapt LLMs as time series encoders [75]. However, because LLMs are pre-trained on language tokens rather than temporal signals, their direct adaptation to HTS depends on effective input representation alignment (see Section 5 for further discussion), and its effectiveness for numerical temporal modeling remains debated [226].

Strategic Adaptation. Beyond direct adaptation, HTS foundation models can also be adapted in more strategic ways, which we summarize into four categories. *(i) Fine-tuning.* A widely adopted strategy involves updating the parameters of a foundation model [24, 127, 265] by training it on additional labeled data from the target domain. This approach directly tailors the pre-trained model to specific downstream task. Furthermore, for foundation models built upon MoE architectures, which have become increasingly prevalent recently, adaptation can also be performed at the expert level. For example, SpecMoE [49] introduces learnable masking matrices to selectively amplify or suppress expert activations, enabling task-specific routing while preserving pre-trained knowledge. *(ii) Repurposing*. It adapts a pre-trained foundation model to a different task type rather than merely fine-tuning it for the same objective. For example, Huang et al. [88] repurposes a forecasting-oriented time series foundation model for HTS classification by freezing the backbone encoder and substituting the original forecasting head with a lightweight classification module. In general, repurposing typically involves training only the newly introduced task-specific parameters on labeled target data, while keeping the pre-trained backbone frozen to preserve its ability for general temporal representations encoding. *(iii) Distillation.* It transfers knowledge from a large pre-trained model to a smaller or more efficient model, enabling downstream applications under resource constraints. For example, PPG-Distill [170] compresses a foundation model into lightweight models that are easier to deploy on wearable devices. *(iv) Integration.* Different foundation models may be expertized in capturing different temporal patterns. Integration-based methods aim to ensemble multiple pre-trained models instead of adapting a single model. For example, ECG-MoE [266] treats multiple foundation models as specialized experts and dynamically integrates their representations through task-conditioned routing.

4.2.2 The Need of Adaptation in Different Healthcare Scenarios.

PhyTS collected from devices of different manufacturers may exhibit significant distributional discrepancies due to variations in their signal processing methods [59]. Such domain shifts may degrade model performance when transferred across devices. In addition, substantial inter-subject variability inherent in biological signals further complicates generalization, as physiological patterns can differ significantly between individuals [259, 299]. These factors, *cross-device variation* and *cross-subject variation*, make effective transfer adaptation of physiological foundation models extremely challenging. **CRTS** varies substantially between institutions due to significant differences in patient populations, clinical practices, and data collection protocols [23]. As a result, models trained in one hospital often fail to generalize to another. This *cross-institution variation* poses significant challenges for transferring foundation models for the clinical record. Each **EpiTS** has their unique intrinsic dynamic, with transmission patterns evolving over time due to viral mutations, behavioral responses, and policy interventions [186, 222]. Consequently, models trained on past outbreaks may struggle to generalize to new events or even different phases of the same epidemic. These factors, *cross-event variation* and *temporal non-stationarity*, pose challenges for the transfer generalization of epidemic foundation models.

4.3 Synthetic Data via Generative Models

Generative models can also help address sample scarcity by synthesizing data for both task-specific supervised learning and large-scale pre-training. Existing work mainly focuses on PhyTS and CRTS, whereas EpiTS generation remains underexplored and is currently limited to data augmentation [245]. We therefore focus on PhyTS and CRTS generation and organize existing methods by the characteristic generative process: (a) **data-driven generation**, (b) **mechanism-informed generation**, and (c) **context-guided generation**.

Data-driven generation. This approach aims to synthesize HTS by directly learning the empirical distribution of observed data. In the reviewed HTS literature, this method is mainly used for CRTS generation, where models synthesize tabular and longitudinal EHR data. In particular, synthetic CRTS are expected to mimic key properties of real clinical data, such as heterogeneous feature types and incomplete observations, as discussed in Section 2. Recent works adopt deep generative frameworks, including GAN- and diffusion-based models, to synthesize heterogeneous clinical records with continuous and discrete variables [230, 277]. These methods aim to generate realistic patient records that preserve the statistical and temporal characteristics of real CRTS. Another line of work addresses partially observed heterogeneous records, where certain feature groups may be absent. For example, FLEXGEN-EHR [83] enables generation from incomplete records by flexibly conditioning on the available feature groups. In addition, some methods reformulate longitudinal tabular records into alternative data modalities to better leverage existing generative model families. For example, TimeHR [100] converts longitudinal EHR data into image-like structures, allowing well-established image generative models to be applied to clinical record synthesis.

Mechanism-Informed Generation. This approach aims to synthesize HTS that are consistent with underlying biological dynamics. In the reviewed HTS literature, this direction is mainly explored in PhyTS generation, where generated samples need to preserve waveform morphology, cardiac dynamics, and inter-channel consistency. Early deep generative approaches directly incorporate morphological priors into the synthesis of ECG. For example, Golany and Radinsky [71] imposes constraints by aligning the generated P–Q–R–S–T wave characteristics with those of real ECG signals. However, morphology-level priors do not explicitly constrain the underlying cardiac dynamics, which have been classically described using ODE systems [161]. Building on this insight, subsequent work integrates ODE-based mechanisms into deep generative models. Specifically, Golany et al. [72], Yehuda and Radinsky [275] enforces dynamical consistency through auxiliary losses based on ODE, while ODE-GAN [70] parameterizes the generator itself as a neural ODE, allowing signal generation according to the learned latent dynamics. Beyond temporal consistency with biological dynamics, physiological generation also seeks to synthesize multi-channel signals with coherent inter-channel relationships. Therefore, later approaches, such as ME-GAN [35], introduce channel-dependent generation to capture inter-channel dependencies. PhysioPDE-GAN [276] further incorporates PDE systems to jointly model temporal dynamics and spatial correlations across channels. Although many physiological synthesis methods are GAN-based, diffusion models have also been explored to improve generation quality and flexibility [5, 248].

Contextual-Aware Generation. This approach synthesizes HTS conditioned on additional contextual information. In the reviewed literature, this direction is mainly used in PhyTS generation, where one physiological signal, clinical text, or patient-specific context can be used to guide the synthesis of another physiological signal. A representative example is PPG-to-ECG generation, where adversarial frameworks such as CardioGAN [206] and P2E-WGAN [237] synthesize ECG waveforms from PPG inputs, and recent diffusion-based approaches further improve generation quality [64, 65, 216]. Similar ideas have been extended to ECG generation from millimeter-wave signals [147], EMG generation from kinematic inputs [246], and ECG generation conditioned on clinical reports or patient-specific information [113, 248]. These methods differ from mechanism-guided synthesis in that conditioning information is provided by external contextual signals, allowing the generated physiological signal to reflect shared physiological couplings across different recordings.

5 Multi-Modal Learning in Healthcare Time Series

Foundation models and generative synthesis mitigate label and sample scarcity by reusing pre-trained representations or expanding available data. However, HTS also suffers from *information scarcity*, where numerical temporal signals often provide only partial and indirect observations of the underlying health state. Multi-modal learning addresses this limitation by incorporating complementary information beyond the time series itself, such as patient metadata, clinical records, imaging, environmental factors, or other contextual signals. Unlike many general time series domains, where naturally aligned multi-modal data are scarce and auxiliary modalities may be only weakly related to the target signals [30, 34, 139], HTS are often accompanied by clinically or epidemiologically meaningful auxiliary information. Thus, multi-modal learning is particularly important for healthcare time series, as it provides critical context for interpreting physiological states, clinical trajectories, and epidemic transmission dynamics.

Formally, a general multi-modal HTS model can be written as $\hat{Y} = f(X^{ts}, \{X^m\}_{m=1}^M)$, where $X^{ts} = \{(t_i, x_i)\}_{i=1}^N$ denotes the time series, X^m denotes the m-th auxiliary modality, and $\hat{Y}$ can represent different healthcare outputs, such as a diagnostic result, risk score, generated report, reasoning response, or future trajectory. We organize this section from two perspectives. Section 5.1 reviews how multi-modal information is integrated with time series, and Section 5.2 discusses which auxiliary modalities are commonly used across different scenarios.

5.1 General Multi-Modal Integration Strategies

Fusion-Based Methods. Fusion-based methods explicitly combine modality-specific representations during the forward propagation process. In general, modality-specific encoders first generate latent representations:$z^{ts} = f_{ts}(X^{ts})$ and $z^m = f_m(X^m), m = 1, \dots, M$ for individual inputs; then a fusion operator combines these embeddings into a fused representation: $z = \text{Fuse}(z^{ts}, \{z^m\}_{m=1}^M)$. Depending on the position where the integration occurs, fusion strategies can be categorized as early, intermediate, or late fusion; the fusion operator can be instantiated in different forms, such as concatenation, projection, or cross-attention [285]. For modality fusion at the early stage, Lan et al. [115], Yang et al. [270] encodes ECG numerical sequences into temporal embeddings, inserts them into prompted text tokens, and then feeds the combined sequence into an LLM for physiological reasoning. For fusion at the intermediate stage, MEIT [240] encodes physiological signals with a temporal encoder and injects the projected embeddings into the Key/Value vectors inside the transformer layers of pre-trained LLMs, allowing language tokens to attend to physiological representations during self-attention. For fusion at the late stage, Kim et al. [105] separately encodes population mobility information as graph representations and social-media tweets as language representations, then uses cross-attention to integrate these high-level features and feeds them into a joint downstream prediction head for epidemic outbreak analysis.

Alignment-Based Methods. Alignment-based methods learn compatible representations across modalities by encouraging modality-specific encoders to capture shared semantic information. In multi-modal HTS, alignment is commonly formulated at the sample level and the token level. *(i) Sample-level alignment* typically follows a contrastive learning paradigm [190]. Modalities from the same sample are treated as positive pairs, while modalities from different samples are treated as negative pairs. For instance, a physiological signal and its paired textual report form a positive pair, while reports from other samples serve as negatives. Given matched representations (z_i^{ts}, z_i^m) and in-batch mismatched pairs (z_i^{ts}, z_k^m), the contrastive loss in Equation 4.1.1 encourages matched pairs to be close and mismatched pairs to be separated. This strategy has been widely used in multi-modal physiological modeling, where numerical physiological signals are aligned with textual reports or descriptions to support zero-shot classification and downstream transfer [128, 138, 242, 281]. It has also been applied to multi-modal clinical record modeling, where modality components associated with the same clinical event, such as an ICU stay or patient visit, are treated as complementary views and encouraged to share compatible representations [108, 249]. *(ii) Token-level alignment* focuses on the compatibility between temporal token embeddings and language token embeddings. It projects time-series tokens into the representation space of pre-trained LLMs, allowing temporally encoded information and related language tokens to be processed in a shared embedding space [93, 224]. In this setting, physiological signals can be mapped into LLM-compatible representation spaces to support multi-modal reasoning [29].

Text-Prompted Methods. Instead of performing multi-modal integration at the representation level, text-prompted methods use language as the primary interface for integrating heterogeneous modalities. These methods convert time series and auxiliary modalities into language-compatible descriptions, and then leverage LLMs as joint reasoning or prediction backbones. According to how the textual interface is constructed, existing methods can be broadly divided into direct transcription and semantic abstraction. *(i) Direct transcription* converts time series into ordered language-like sequences, where each textual unit corresponds to a timestamped observation. Given a HTS $X^{ts} = \{(t_i, x_i)\}_{i=1}^{N}$, direct transcription maps each timestamped observation into a textual unit $u_i^{ts} = \phi_{\text{ts}}(t_i, x_i)$, yielding a language-compatible sequence $s^{ts} = [u_1, \ldots, u_N]$. Similarly, each auxiliary modality is textualized as $s^m = \phi_m(X^m), m = 1, \ldots, M$. Such a method is especially useful for CRTS, since they are already structured as longitudinal measurements and care events. Existing studies serialize EHR trajectories into LLM-readable inputs for clinical modeling, covering representation learning and fine-tuning [61], ICU outcome prediction [44, 86], discharge prediction [287], and forecasting with clinical notes [172, 243]. However, point-wise textualization can generate long and redundant prompts, motivating compressed tokenization [298] and retrieval-based history selection [27]. *(ii) Semantic abstraction* converts time series into high-level textual descriptions instead of preserving every timestamped observation. Specifically, for a time series X^{ts}, it produces $s_{ts} = \phi_{ts}(X^{ts}) = [v_1, \ldots, v_K]$, where each v_k summarizes a domain-specific feature, and K is not necessarily determined by N. Here, ϕ_{ts} may be implemented through signal-processing modules [62, 137, 280], clinical templates [90, 121], retrieval systems [27, 280], or LLM-based summarization [46, 112]. Compared with direct transcription, this approach produces more compact and interpretable prompts and can naturally incorporate domain knowledge and heterogeneous context. However, it introduces an information bottleneck and may discard fine-grained numerical values and temporal details. Thus, text-prompted learning should be viewed as a complementary interface for numerical time series modeling [38, 69].

5.2 Multi-Modal Information in Different HTS Scenarios

In multi-modal HTS modeling, auxiliary modalities can be broadly categorized into (a) **self-derived views** and (b) **external complementary sources**. Self-derived views are constructed from the time series themselves, such as textual summaries or other representations that highlight temporal characteristics. External complementary sources provide information beyond temporal signals, such as clinical reports or other domain knowledge. The specific forms of these modalities vary across different HTS scenarios.

Multi-modal PhyTS. *(i) self-derived views* include textual descriptions of waveform morphology, physiological patterns, diagnostic cues [62, 137, 280], and image-like representations derived from raw signals [115]. *(ii) External sources* include diagnostic reports, clinical interpretations, and patient metadata, which provide semantic or contextual information beyond the numerical waveform [138, 270, 281]. In addition, multi-modality provides a language-based interface for question answering, interpretation, and personalized feedback [63, 106, 137, 142, 162], extending physiological time series modeling beyond classical prediction tasks toward interactive and personalized health support.

Multi-modal CRTS. *(i) self-derived views* are constructed from structured EHR trajectories themselves, such as serialized records, pseudo-notes, or compact textual summaries of measurements and care events. Recent studies have used these self-derived views for diverse tasks, including clinical prediction, temporal reasoning, and decision support [44, 46, 61, 86, 90, 112, 121, 287]. *(ii) External sources* mainly include clinical text and medical imaging. Clinical notes provide narrative information about symptoms, clinician assessments, treatment rationales, and care plans, complementing structured EHR trajectories in clinical prediction tasks [103, 108, 154, 249, 254]. Medical imaging provides additional visual and semantic evidence, especially when chest X-rays, radiology reports, discharge notes, and temporal records are also available [82, 96, 253].

Multi-modal EpiTS. Most studies here focus on integrating *external sources* into joint modeling, where these auxiliary modalities mainly include spatial signals and contextual public-health information. Specifically, spatial signals, such as human mobility data, complement epidemiological trajectories by approximating contact opportunities, population movement, and exposure risk [31, 47]. Contextual information captures external factors related to public behavior,

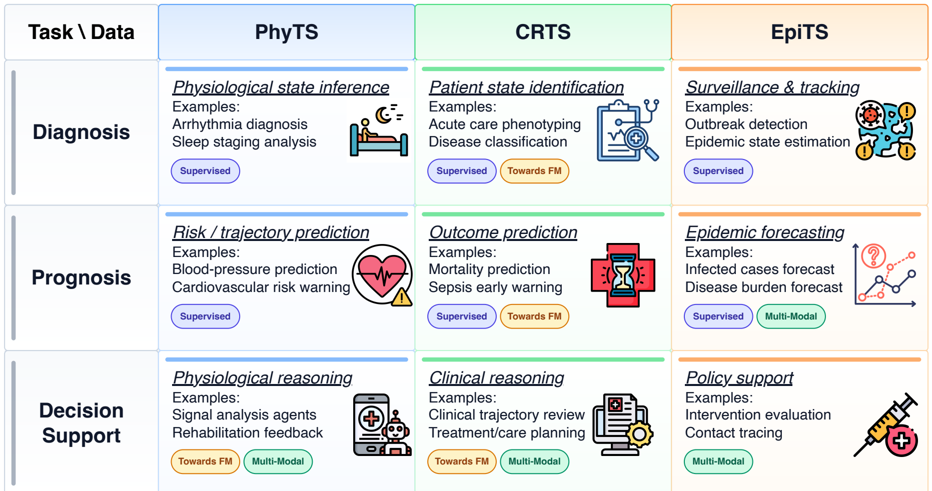

Fig. 3. HTS types × Healthcare tasks × (main) Modeling paradigms.

intervention status, and evolving pathogen characteristics. For example, Internet search trends and pandemic-related social media activity can provide early behavioral indicators of outbreak dynamics [57, 234]. In addition, recent work further incorporates policy changes and genomic surveillance signals for contextualized epidemic forecasting [55, 228].

6 Application Domains

In this section, we provide a review of the representative applications of HTS. In particular, the applications are organized according to the task taxonomy introduced in Section 2.3, and a brief summary is presented in Figure 3.

6.1 Diagnosis

Diagnostic applications infer the current health state from observed temporal evidence.

In **PhyTS**, diagnostic applications are well established in cardiovascular, neurological, and sleep-related settings. For example, ECG is widely used for cardiovascular diagnosis, with arrhythmia detection being a central application. Early work studied rhythm classification on single-lead ambulatory ECG [80], while later studies extended deep ECG interpretation to standard 12-lead recordings for multiple diagnostic abnormalities [199]. Beyond rhythm classification, ECG-based applications also support broader cardiac assessment, including cardiac dysfunction screening, identification of atrial fibrillation during sinus rhythm, and acute myocardial infarction detection [12, 13, 140]. Wearable physiological signals can also support state detection in daily-life settings, such as passive atrial fibrillation detection from smartwatch data [232]. In neurological and sleep-related applications, EEG and polysomnography support sleep staging, where consecutive temporal windows are assigned sleep-stage labels using EEG, EOG, EMG, and related channels [182, 183]. EEG also supports seizure detection, where models identify abnormal events in long recordings [3].

In **CRTS**, diagnosis corresponds to identifying a patient's current or clinically relevant state from structured clinical records. This includes diagnosis prediction, disease classification, and diagnosis-related clinical-event prediction from historical visits or hospital trajectories [132, 133, 155, 181, 194, 218]. Another diagnostic use case is patient phenotyping and subtyping, where longitudinal clinical observations are used to characterize patient groups or acute-care phenotypes [19, 81]. Monitoring can also be viewed as repeated diagnosis when the goal is to update the estimated patient state as new observations arrive, supporting bedside reassessment and online patient monitoring in ICU and inpatient settings [273]. Overall, CRTS diagnosis is less tied to a single disease-specific test than PhyTS diagnosis; it more often concerns patient-state identification across heterogeneous clinical histories.

In **EpiTS**, diagnosis appears as real-time surveillance and epidemic-state estimation, where the goal is to assess current disease activity rather than forecast future burden. Public-health surveillance commonly tracks indicators such as reported cases, incidence rates, influenza-like illness rates, hospitalizations, or mortality to support situational awareness across regions [45, 54, 158]. Epidemic tracking has also been studied by combining influenza-related surveillance signals with epidemic-state estimation, such as Google Flu Trends with state-space SEIR modeling [57]. When surveillance is organized across locations, these applications can support current hotspot identification and regional risk assessment, providing the diagnostic counterpart to future regional spread forecasting.

6.2 Prognosis

Prognostic applications predict future health outcomes, risks, or trajectories from historical temporal observations.

In **PhyTS**, wearable and remote monitoring have expanded physiological applications from episodic clinical testing to longitudinal risk assessment. Wearable signals such as PPG, heart-rate, and activity trajectories have been used for cardiovascular risk prediction from daily-life monitoring data [16]. ECG and PPG applications also support broader physiological risk assessment, including blood-pressure estimation and forecasting [60, 290]. These applications often aggregate noisy signals over longer horizons, making robustness to motion artifacts, device placement, user behavior, and sampling variability central to clinical utility.

In **CRTS**, a major prognostic application is acute care risk prediction, where patient trajectories are used to estimate near-term outcomes such as mortality, clinical deterioration, sepsis onset, and prolonged length of stay. In ICU settings, structured vital signs, laboratory measurements, and clinical observations have been used for mortality, decompensation, length-of-stay prediction, and phenotyping [81, 122, 218, 231]. In addition, early warning systems are also a particularly important prognostic setting because the prediction horizon directly affects clinical utility. For example, Wang et al. [254] studied sepsis prediction using early ICU measurements and clinical notes, illustrating the need to evaluate models by both accuracy and how early they identify at-risk patients. Bedside monitoring also belongs to prognosis when the goal is to update future-risk estimates as new measurements, notes, or imaging evidence become available. Prior work has used ICU time series and clinical notes for mortality, decompensation, and length-of-stay prediction [103], with related studies predicting in-hospital mortality from structured EHR trajectories, clinical notes, or LLM-generated summaries [18, 154]. When available, imaging evidence can further complement longitudinal clinical measurements for bedside outcome assessment [82, 96].

In **EpiTS**, prognostic applications mainly forecast future disease activity for public-health preparedness and planning. Short-term forecasting estimates near-future surveillance indicators from historical epidemic observations. Representative applications include influenza-like illness forecasting in the United States and Japan [261] and COVID-19 case prediction in India [11]. Operational forecasting settings also consider public-health targets such as hospitalizations and mortality, as reflected in COVID-19 forecast aggregation [45] and FluSight influenza forecasting evaluations [158]. EpiTS prognosis is especially important for emerging or rapidly evolving outbreaks, where local historical data may be limited. For example, pandemic forecasting has been studied across countries at different epidemic stages [180], and epidemic forecasting under sparse local surveillance has been studied using synthetic information [245]. Beyond single-region forecasting, regional spread forecasting estimates future disease activity across connected locations, supporting hotspot detection and regional risk assessment [50, 239, 244]. Mobility-aware forecasting further supports this regional perspective by relating population movement to spatial disease spread under changing intervention conditions [227]. In practice, EpiTS prognosis is complicated by reporting delays and data revisions, which can affect both real-time forecasts and retrospective evaluation [17, 222, 235].

6.3 Decision Support

Decision-support applications use HTS to inform actions rather than only estimate states or predict future outcomes. This task family remains less mature than diagnosis and prognosis, but it is important because real healthcare and public-health settings require decisions about when to intervene, which action to take, how strongly to respond, and how recommendations should be updated as new temporal observations arrive.

In **PhyTS**, decision support appears in rehabilitation, neuromuscular assessment, human–machine interaction, assistive technology, and personalized physiological feedback. For example, EMG and EOG signals are used in rehabilitation and assistive interaction [32, 130]. Specifically, EMG supports gesture recognition, prosthetic control, fatigue monitoring, and motor rehabilitation, while EOG captures eye-movement-related potentials for sleep analysis, fatigue detection, and assistive interaction. EEG-based brain–computer interfaces also support assistive interaction and neurotechnology applications [117]. In daily-life monitoring settings, wearable physiological signals can further support personalized health feedback and user-facing physiological insight generation [63, 66, 162]. Compared with diagnosis-oriented ECG or EEG interpretation, these applications are closer to action support because their outputs may guide device control, rehabilitation feedback, assistive interaction, or personalized self-management.

In **CRTS**, decision support is centered on care delivery, where temporal patient records are used to inform clinical actions, workflow decisions, or care planning. One representative setting is treatment strategy support in intensive care, especially for sepsis management, where longitudinal patient records can be used to study treatment decisions and their downstream outcomes [110]. Another setting is longitudinal clinical reasoning and summarization, where patient histories are organized to support review of disease progression, recent events, and possible care needs [27, 46, 112, 121]. Related workflow-oriented applications, such as delirium prediction, ICU readmission prediction, and discharge prediction, can support care coordination and clinical prioritization, although they are closer to prognostic support than direct intervention recommendation [44, 86, 287]. Overall, CRTS decision support remains an emerging area, with a clear gap between retrospective prediction and prospective, intervention-aware clinical action.

In **EpiTS**, decision support includes intervention evaluation, contact tracing, resource allocation, and policy planning. Forecasts of regional spread can support coordinated allocation of public-health resources, while mobility-aware and policy-aware studies examine how population movement and interventions shape transmission and public-health response [31, 97, 227]. Epidemic control and contact-tracing applications have also studied how to choose interventions under competing public-health, social, and economic objectives [67, 219]. Compared with pure forecasting, these applications require explicit consideration of how public-health actions alter future epidemic dynamics.

6.4 Datasets and Benchmarks

PhyTS research relies on benchmarks spanning different modalities, acquisition settings, and annotation levels. For ECG, rhythm-event detection is commonly evaluated on the MIT-BIH Arrhythmia Database [164] and the PhysioNet/CinC 2017 single-lead atrial fibrillation challenge dataset [43], while multi-label 12-lead diagnostic classification often uses PTB-XL [238], the Chapman-Shaoxing ECG database [293], and MIMIC-IV-ECG [73]. For sleep and neurological monitoring, Sleep-EDF [102] and SHHS [189] are widely used for polysomnography-based sleep analysis, while CHB-MIT [215] and the TUH EEG Corpus [174] support seizure detection and clinical EEG modeling. For wearable and

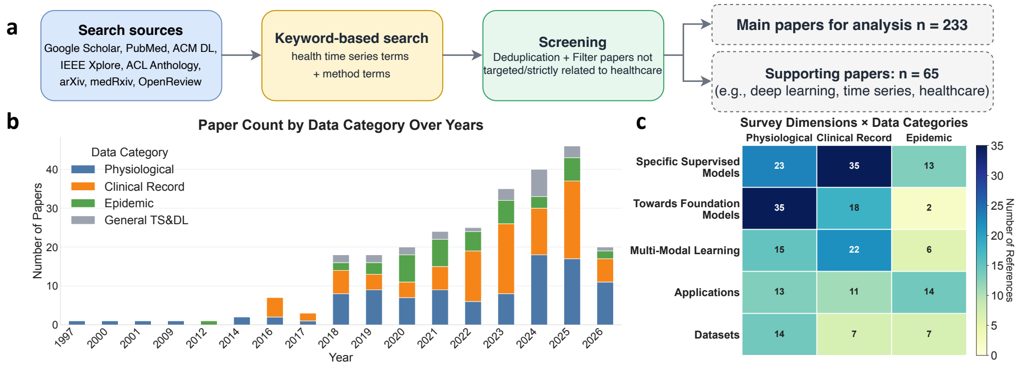

Fig. 4. Quantitative summary of the surveyed literature. (a) Paper searching and screening process. (b) Number of included papers by publication year and HTS data categories. (c) Distribution of reviewed papers across survey dimensions and HTS data categories.

rehabilitation settings, PPG-DaLiA [197], the IEEE Signal Processing Cup 2015 wrist-PPG dataset [292], BIDMC PPG and Respiration [185], WESAD [207], and NinaPro [14] cover heart-rate estimation, respiratory-rate estimation, stress detection, and EMG-based gesture or prosthetic-control tasks.

CRTS research is supported by publicly accessible critical care and longitudinal EHR databases. MIMIC-III [95] and MIMIC-IV [94] provide multi-year ICU records from a medical center, while eICU [187] offers multi-center coverage across over 200 U.S. hospitals. HiRID provides two-minute-resolution physiological recordings [274], and EHRSHOT [260] includes longitudinal outpatient and inpatient trajectories with 15 clinical prediction tasks designed for few-shot evaluation of foundation models. On the benchmark side, Harutyunyan et al. [81] established a widely adopted suite on MIMIC-III covering in-hospital mortality, decompensation, length-of-stay, and phenotyping, while the HiRID-ICU-Benchmark [274] extended this to high-resolution data with reproducible pipelines. More recently, MIMIC-IV-EXT [243] introduced a 22-million-event temporal stream supporting long-horizon modeling.

EpiTS research relies on publicly reported surveillance datasets that differ in temporal resolution, geographic scope, and reporting reliability. For seasonal influenza, the U.S. CDC ILI surveillance system has served as the primary domestic data source, and the FluSight challenge [196] has coordinated annual multi-team probabilistic forecasting evaluations since 2013, its most recent assessment [158] covers the 2021–23 seasons under laboratory-confirmed hospitalization as the forecast target. For COVID-19, the JHU CSSE dashboard [54] provided global case and mortality TS as ground truth, and the US COVID-19 Forecast Hub [45] aggregated forecasts from over 90 modeling teams across cases, hospitalizations, and deaths. A recurring difficulty across these resources is that ground truth is subject to reporting delays, revisions, and definitional changes [17, 222, 235], making retrospective evaluation sensitive to the data version used.

7 Survey Methodology

Search Strategy. We conducted a systematic literature review to synthesize recent research progress in deep learning for HTS. Unlike reviews focused on a single clinical question or one data modality, this survey aims to organize a broad and rapidly evolving literature according to data types, modeling challenges, and methodological paradigms. We mainly considered papers published from **2016** to **May 2026**, with the final search update completed in **May 2026**. The overall search and screening strategy is highlighted in Figure 4 (a). We searched major academic databases and publication platforms to capture both representative and recent works in artificial intelligence and healthcare, including Google Scholar, PubMed, ACM Digital Library, IEEE Xplore, ACL Anthology, arXiv, medRxiv, and OpenReview. The search was

Table 2. Comparison of representative HTS surveys. The symbols '✓", '△", and "–" denote substantial coverage, partial coverage, and no major focus, respectively. Methodological dimensions include task diversity (diagnosis, prognosis, decision support), emerging deep learning paradigms (foundation models, LLMs, generative AI), and multi-modal information.

Survey	Primary Scope	PhyTS	CRTS	EpiTS	Task Diversity	Emerging DL	Multi-Modality
Sun et al. [223]	Focuses on CRTS irregular modeling.	–	✓	–	–	–	–
Morid et al. [166]	Reviews DL methods for CRTS prediction.	–	✓	–	–	–	–
Xie et al. [264]	Reviews temporal representation learning for CRTS.	–	✓	–	△	–	–
Amirahmadi et al. [7]	Focuses on CRTS prediction and data-processing techniques.	–	✓	–	△	–	–
Ren et al. [198]	Covers EHR modeling from DL to LLMs, but not restricted to HTS.	–	△	–	✓	✓	✓
Alqudah and Moussavi [6]	Reviews general ML/DL methods for PhyTS.	✓	–	–	△	–	–
Liu et al. [145]	Focuses on contrastive self-supervised learning for PhyTS.	✓	–	–	–	△	–
Weng et al. [259]	Systematically reviews self-supervised learning for PhyTS.	✓	–	–	–	△	–
Kamble and Sengupta [98]	Reviews emotion recognition using PhyTS.	✓	–	–	–	–	–
Pillalamarri and Shanmugam [184]	Reviews multi-modal learning for emotion recognition using PhyTS.	✓	–	–	–	△	✓
Nie et al. [171]	Reviews DL methods for PhyTS, mostly about supervised methods.	✓	–	–	△	–	–
Yang et al. [271]	Focuses on PhyTS anomaly detection.	✓	–	–	–	–	–
Li et al. [125]	Reviews few-shot learning for PhyTS.	✓	–	–	–	–	–
Gu et al. [77]	Surveys foundation models for PhyTS.	✓	–	–	✓	✓	✓
Liu et al. [149]	Reviews GNNs for spatial-temporal epidemic modeling.	–	–	✓	✓	–	–
Rodriguez et al. [201]	Reviews general ML methods for epidemic forecasting.	–	–	✓	–	△	–
Sharma et al. [209]	Reviews DL methods for epidemic modeling and control.	–	–	✓	✓	△	✓
Babanejaddehaki et al. [15]	Reviews DL methods for epidemic outbreak detection and forecasting.	–	–	✓	△	–	–
Liu et al. [144]	Focuses on LLMs for HTS, with limited coverage of non-LLM methods.	✓	✓	–	–	△	✓
Loni et al. [150]	Reviews generative AI for synthetic medical text and HTS.	✓	✓	–	–	✓	–
Our survey	Reviews diverse HTS, organized by evolving trajectory and modeling paradigms.	✓	✓	✓	✓	✓	✓

conducted using three groups of keywords. The first group emphasized HTS data, covering both general terminologies (e.g., health time series, medical time series) and specific data modalities, ranging from bio-signals (e.g., ECG, EEG, PPG, EMG) to structured clinical records (e.g., ICU records, EHR trajectories), as well as specific infectious-disease surveillance data (e.g., COVID-19). The second group covered method-related terms, including neural architectures, learning paradigms, and emerging model families (e.g., RNNs, Transformers, contrastive learning, foundation models, generative models, multi-modal). The third group covered common modeling challenges in HTS (e.g., irregular sampling, missingness, data scarcity, distribution shift, measurement noise). Representative Boolean search strings included:

(i) (bio-signals OR longitudinal EHR or epidemic time series) AND (deep learning OR machine learning);

(ii) (medical time series OR health time series) AND (foundation model OR multi-modal OR generative model);

(iii) (medical time series OR health time series) AND (irregularity OR scarcity OR noise).

Screening and Eligibility Criteria. After the initial collection, papers were screened according to their relevance to the scope of this review. We included method-related papers if they addressed healthcare-related temporal modeling problems and proposed, adapted, or evaluated computational methods for HTS. Generic time series papers were retained only when they represented influential methodological developments with practical relevance to HTS modeling. In addition, relevant surveys, foundational machine learning works, dataset papers, and domain-specific studies were cited as contextual references to motivate key challenges, clarify background concepts, and position the reviewed methods.

Result. After deduplication and relevance filtering, the final reference set contained 298 papers, including 233 papers for the main analysis and 65 papers as supporting material. The quantitative summaries in Figure 4 (b,c) are based on the 233 main-analysis papers: Figure 4 (b) reports their distribution across publication years and HTS categories, while Figure 4 (c) summarizes the coverage of survey dimensions and data categories.

8 Future Directions

Despite rapid progress in HTS modeling, current methods remain closer to predictive models than deployable clinical or public-health decision-support systems. Future progress requires moving beyond predictive accuracy toward models that support temporally grounded decisions, provide faithful clinical evidence, reason over long and sparse trajectories, interact with external tools or specialized temporal models, adapt under deployment-time shifts, and satisfy safety, fairness, and privacy requirements. We summarize these challenges along six future directions.

From Prediction to Actionable Decision Support. Although recent HTS models have advanced diagnostic and prognostic modeling [97, 122, 173, 218, 231], most remain primarily predictive, while real clinical and public-health settings require decisions about when to intervene, which action to take, how strongly to respond, and how to update recommendations as new observations arrive. This is challenging because decisions are part of the temporal process: treatments, monitoring intensity, resource allocation, or vaccination policies can alter the future trajectories being predicted [110]. Future work should therefore develop treatment-aware, policy-aware, and counterfactual HTS models that estimate outcomes under alternative interventions.

Reliable and Clinically Grounded Interpretation. Interpretation is critical for HTS models because predictions must be assessed together with clinical evidence. Early work has explored interpretable temporal architectural designs [41, 68, 218], while recent language-based methods further translate time series into readable reports, pseudo-notes, or natural-language rationales [121, 137]. However, readability alone is insufficient: explanations must be grounded in relevant temporal evidence, calibrated with uncertainty, and consistent with clinical knowledge. Future work should develop faithful, uncertainty-aware, and clinically grounded interpretation frameworks that verify whether highlighted timestamps, variables, or generated rationales truly support model predictions.

Long-Horizon Health Dynamics Modeling. Many HTS tasks are framed as fixed-window prediction problems, leaving long-horizon health dynamics underexplored. Long-horizon modeling requires using extended patient histories to maintain evolving health-state representations and reason about future risks or trajectories over months or years. This is important because future health may depend on distant events, accumulated comorbidities, prior interventions, behavioral patterns, and delayed disease progression. For example, Shmatko et al. [214] models the progression and competing nature of over one thousand diseases from patient histories observed until age 60, and samples future trajectories up to 20 years. Future work should develop HTS models that connect long accumulated histories with broad future trajectories, integrate repeated measurements with interventions and delayed outcomes, quantify uncertainty over distant futures, and support monitoring, prevention, and intervention planning.

Integrating Agentic Systems with Specialized Temporal Models. Agentic systems offer a promising interface for HTS applications that require querying temporal evidence, retrieving clinical or public-health knowledge, comparing actions, and updating recommendations as new observations arrive [66, 87, 162, 256]. However, agents should not be viewed as replacements for specialized HTS models, since key HTS evidence is often encoded in numerical temporal structures that LLM-based agents may analyze inefficiently or unreliably. Future work should integrate agents with dedicated temporal models, for example by aligning temporal representations with language-token spaces or invoking temporal models as external analysis tools. Such integration may enable interactive and context-aware support while preserving the fidelity, efficiency, and domain grounding of HTS models.

Robust Deployment-Time Adaptation. Although representation learning and foundation-model pretraining improve HTS model transfer across datasets, tasks, and modalities, they do not guarantee robustness after deployment. Real deployments involve cross-environment shifts across hospitals, devices, populations, or measurement protocols, as well as temporal shifts within the same environment over calendar time. In healthcare, these shifts are often entangled with changes in treatment practice, missingness, sampling frequency, device noise, and population characteristics. Studies on temporal dataset shift show that adaptation can improve robustness [78, 84, 119], but naive adaptation may be unstable or harmful under realistic shifts. Future work should develop deployment-time adaptation frameworks that detect drift, adapt from limited or unlabeled target data, preserve calibration, and satisfy privacy and safety constraints.

Trustworthy Deployment. Strong benchmark performance alone is insufficient for HTS deployment, which requires evidence of reliability in the intended clinical or public-health setting. Three gaps are central. First, the *aggregate-to-subgroup gap* means average performance may hide poor accuracy for clinically important groups defined by sex, race, age, institution, or device type. Prior work on healthcare algorithmic bias shows that proxy outcomes and data-generating processes can reproduce existing disparities [175], motivating subgroup-specific performance and calibration reporting. Second, the *benchmark-to-setting gap* means performance on one dataset does not ensure transportability to the target hospital, device, population, or calendar period. External validation is therefore needed to assess generalizability [192], especially under temporal dataset shift [78]. Third, the *research-to-deployment gap* reflects that clinical deployment is a governed lifecycle involving ethics review, health-data privacy regulations such as HIPAA, data-use agreements, and regulatory pathways such as FDA authorization. Future work should develop deployment-oriented evaluation pipelines covering subgroup calibration, cross-setting transferability, privacy-preserving multi-institutional validation, human-in-the-loop assessment, and post-deployment monitoring.

9 Conclusion

This survey reviewed deep learning for HTS across physiological, clinical records, and epidemic time series. We organized the survey under a unified scarcity-driven perspective, showing how task-specific supervised learning, foundation models, and multi-modal learning have emerged to address label scarcity, sample scarcity, and information scarcity in healthcare data. We further summarized representative application domains and key research gaps, highlighting meaningful future directions. We hope this survey provides a unified map for understanding methodological connections across HTS domains and for developing accurate, trustworthy, generalizable, and clinically useful HTS models.

Ethics and Privacy Statement

This survey uses only published literature and collects no human-subject or personal health data. Nevertheless, real-world HTS applications require privacy protection, bias and safety assessment, and clinical validation.

References

[1] Salar Abbaspourazad, Oussama Elachqar, Andrew Miller, Saba Emrani, Udhyakumar Nallasamy, and Ian Shapiro. 2024. Large-scale Training of Foundation Models for Wearable Biosignals. In *The Twelfth International Conference on Learning Representations*.

[2] Josh Abramson, Jonas Adler, Jack Dunger, Richard Evans, Tim Green, Alexander Pritzel, Olaf Ronneberger, Lindsay Willmore, Andrew J Ballard, Joshua Bambrick, et al. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630, 8016 (2024), 493–500.

[3] U Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, and Hojjat Adeli. 2018. Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in biology and medicine* 100 (2018), 270–278.

[4] Ahmed M Alaa and Mihaela van der Schaar. 2019. Attentive state-space modeling of disease progression. *Advances in neural information processing systems* 32 (2019).

[5] Juan Miguel Lopez Alcaraz and Nils Strodthoff. 2023. Diffusion-based conditional ECG generation with structured state space models. *Computers in biology and medicine* 163 (2023), 107115.

[6] Ali Mohammad Alqudah and Zahra Moussavi. 2025. Bridging Signal Intelligence and Clinical Insight: A Comprehensive Review of Feature Engineering, Model Interpretability, and Machine Learning in Biomedical Signal Analysis. *Applied Sciences* 15, 22 (2025), 12036.

[7] Ali Amirahmadi, Mattias Ohlsson, and Kobra Etminani. 2023. Deep learning prediction models based on EHR trajectories: A systematic review. *Journal of biomedical informatics* 144 (2023), 104430.

[8] Ying An, Guanglei Cai, Xianlai Chen, and Lin Guo. 2023. PARSE: A personalized clinical time-series representation learning framework via abnormal offsets analysis. *Computer Methods and Programs in Biomedicine* 242 (2023), 107838.

[9] Bahare Andayeshgar, Fardin Abdali-Mohammadi, Majid Sepahvand, Alireza Daneshkhah, Afshin Almasi, and Nader Salari. 2022. Developing graph convolutional networks and mutual information for arrhythmic diagnosis based on multichannel ECG signals. *International Journal of Environmental Research and Public Health* 19, 17 (2022), 10707.

[10] Abdul Fatir Ansari, Lorenzo Stella, Ali Caner Turkmen, Xiyuan Zhang, Pedro Mercado, Huibin Shen, Oleksandr Shchur, Syama Sundar Rangapuram, Sebastian Pineda Arango, Shubham Kapoor, Jasper Zschiegner, Danielle C. Maddix, Hao Wang, Michael W. Mahoney, Kari Torkkola, Andrew Gordon Wilson, Michael Bohlke-Schneider, and Bernie Wang. 2024. Chronos: Learning the Language of Time Series. *Transactions on Machine Learning Research* (2024). Expert Certification.

[11] Parul Arora, Himanshu Kumar, and Bijaya Ketan Panigrahi. 2020. Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India. *Chaos, solitons & fractals* 139 (2020), 110017.

[12] Zachi I Attia, Suraj Kapa, Francisco Lopez-Jimenez, Paul M McKie, Dorothy J Ladewig, Gaurav Satam, Patricia A Pellikka, Maurice Enriquez-Sarano, Peter A Noseworthy, Thomas M Munger, et al. 2019. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. *Nature medicine* 25, 1 (2019), 70–74.

[13] Zachi I Attia, Peter A Noseworthy, Francisco Lopez-Jimenez, Samuel J Asirvatham, Abhishek J Deshmukh, Bernard J Gersh, Rickey E Carter, Xiaoxi Yao, Alejandro A Rabinstein, Brad J Erickson, et al. 2019. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *The Lancet* 394, 10201 (2019), 861–867.

[14] Manfredo Atzori, Arjan Gijsberts, Claudio Castellini, Barbara Caputo, Anne-Gabrielle Mittaz Hager, Simone Elsig, Giorgio Giatsidis, Franco Bassetto, and Henning Müller. 2014. Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Scientific data* 1, 1 (2014), 140053.

[15] Ghazaleh Babanejaddehaki, Aijun An, and Manos Papagelis. 2025. Disease outbreak detection and forecasting: A review of methods and data sources. *ACM Transactions on Computing for Healthcare* 6, 2 (2025), 1–40.

[16] Brandon Ballinger, Johnson Hsieh, Avesh Singh, Nimit Sohoni, Jack Wang, Geoffrey Tison, Gregory Marcus, Jose Sanchez, Carol Maguire, Jeffrey Olgin, and Mark Pletcher. 2018. DeepHeart: Semi-Supervised Sequence Learning for Cardiovascular Risk Prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.

[17] Leonardo S Bastos, Theodoros Economou, Marcelo FC Gomes, Daniel AM Villela, Flavio C Coelho, Oswaldo G Cruz, Oliver Stoner, Trevor Bailey, and Claudia T Codeço. 2019. A modelling approach for correcting reporting delays in disease surveillance data. *Statistics in medicine* 38, 22 (2019), 4363–4377.

[18] Harshavardhan Battula, Jiacheng Liu, and Jaideep Srivastava. 2024. Enhancing In-Hospital Mortality Prediction Using Multi-Representational Learning with LLM-Generated Expert Summaries. *arXiv preprint arXiv:2411.16818* (2024).

[19] Inci M Baytas, Cao Xiao, Xi Zhang, Fei Wang, Anil K Jain, and Jiayu Zhou. 2017. Patient subtyping via time-aware LSTM networks. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*. 65–74.

[20] Sami Briouza, Hassène Gritli, Nahla Khraief, Safya Belghith, and Dilbag Singh. 2021. A convolutional neural network-based architecture for EMG signal classification. In *2021 International conference on data analytics for business and industry (ICDABI)*. IEEE, 107–112.

[21] Garyk Brixi, Matthew G Durrant, Jerome Ku, Mohsen Naghipourfar, Michael Poli, Gwanggyu Sun, Greg Brockman, Daniel Chang, Alison Fanton, Gabriel A Gonzalez, et al. 2026. Genome modelling and design across all domains of life with Evo 2. *Nature* 652, 8112 (2026), 1349–1361.

[22] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.

[23] Manuel Burger, Daphné Chopard, Malte Londschien, Fedor Sergeev, Hugo Yèche, Rita Kuznetsova, Martin Faltys, Eike Gerdes, Polina Leshetkina, Peter Bühlmann, et al. 2025. A foundation model for intensive care: Unlocking generalization across tasks and domains at scale. *medRxiv* (2025), 2025–07.

[24] Manuel Burger, Fedor Sergeev, Malte Londschien, Daphné Chopard, Hugo Yèche, Eike Christian Gerdes, Polina Leshetkina, Alexander Morgenroth, Zeynep Babür, Jasmina Bogojeska, Martin Faltys, Rita Kuznetsova, and Gunnar Ratsch. 2024. Towards Foundation Models for Critical Care Time Series. In *Advancements In Medical Foundation Models: Explainability, Robustness, Security, and Beyond*.

[25] Donghong Cai, Junru Chen, Yang Yang, Teng Liu, and Yafeng Li. 2023. Mbrain: A multi-channel self-supervised learning framework for brain signals. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 130–141.

[26] Defu Cao, Yujing Wang, Juanyong Duan, Ce Zhang, Xia Zhu, Congrui Huang, Yunhai Tong, Bixiong Xu, Jing Bai, Jie Tong, and Qi Zhang. 2020. Spectral Temporal Graph Neural Network for Multivariate Time-series Forecasting. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 17766–17778.

[27] Lang Cao, Qingyu Chen, and Yue Guo. 2026. EHR-RAG: Bridging Long-Horizon Structured Electronic Health Records and Large Language Models via Enhanced Retrieval-Augmented Generation. *arXiv preprint arXiv:2601.21340* (2026).

[28] Wei Cao, Dong Wang, Jian Li, Hao Zhou, Lei Li, and Yitan Li. 2018. Brits: Bidirectional recurrent imputation for time series. *Advances in neural information processing systems* 31 (2018).

[29] Nimeesha Chan, Felix Parker, William Bennett, Tianyi Wu, Mung Yao Jia, James Fackler, and Kimia Ghobadi. 2024. Leveraging LLMs for Multimodal Medical Time Series Analysis. In *MLHC*.

[30] Ching Chang, Jeehyun Hwang, Yidan Shi, Haixin Wang, Wei Wang, Wen-Chih Peng, and Tien-Fu Chen. 2026. Time-imm: A dataset and benchmark for irregular multimodal multivariate time series. In *Advances in Neural Information Processing Systems*, Vol. 38.

[31] Serina Chang, Mandy L Wilson, Bryan Lewis, Zakaria Mehrab, Komal K Dudakiya, Emma Pierson, Pang Wei Koh, Jaline Gerardin, Beth Redbird, David Grusky, et al. 2021. Supporting covid-19 policy response with large-scale mobility-based modeling. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 2632–2642.

[32] Won-Du Chang. 2019. Electrooculograms for Human–Computer Interaction: A Review. *Sensors* 19, 12 (2019), 2690.

[33] Zhengping Che, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu. 2018. Recurrent neural networks for multivariate time series with missing values. *Scientific reports* 8, 1 (2018), 6085.

[34] Jialin Chen, Aosong Feng, Ziyu Zhao, Juan Garza, Gaukhar Nurbek, Cheng Qin, Ali Maatouk, Leandros Tassiulas, Yifeng Gao, and Rex Ying. 2025. Mtbench: A multimodal time series benchmark for temporal reasoning and question answering. *arXiv preprint arXiv:2503.16858* (2025).

[35] Jintai Chen, Kuanlun Liao, Kun Wei, Haochao Ying, Danny Z Chen, and Jian Wu. 2022. ME-GAN: Learning panoptic electrocardio representations for multi-view ECG synthesis conditioned on heart diseases. In *International Conference on Machine Learning*. PMLR, 3360–3370.

[36] Jiayuan Chen, Changchang Yin, Yuanlong Wang, and Ping Zhang. 2024. Predictive modeling with temporal graphical representation on electronic health records. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*. 5763–5771.

[37] Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. 2018. Neural ordinary differential equations. *Advances in neural information processing systems* 31 (2018).

[38] Tianyi Chen, Mingcheng Zhu, Zhiyao Luo, and Tingting Zhu. 2026. Cross-representation benchmarking in time-series electronic health records for clinical outcome prediction. In *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 7076–7080.

[39] Zhuoling Cheng, Xuekui Bu, Qingnan Wang, Tao Yang, and Jihui Tu. 2024. EEG-based emotion recognition using multi-scale dynamic CNN and gated transformer. *Scientific Reports* 14, 1 (2024), 31319.

[40] Edward Choi, Mohammad Taha Bahadori, Le Song, Walter F Stewart, and Jimeng Sun. 2017. GRAM: graph-based attention model for healthcare representation learning. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*. 787–795.

[41] Edward Choi, Mohammad Taha Bahadori, Jimeng Sun, Joshua Kulas, Andy Schuetz, and Walter Stewart. 2016. Retain: An interpretable predictive model for healthcare using reverse time attention mechanism. *Advances in neural information processing systems* 29 (2016).

[42] Ranak Roy Chowdhury, Jiacheng Li, Xiyuan Zhang, Dezhi Hong, Rajesh K Gupta, and Jingbo Shang. 2023. Primenet: Pre-training for irregular multivariate time series. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 7184–7192.

[43] Gari D Clifford, Chengyu Liu, Benjamin Moody, Li-wei H Lehman, Ikaro Silva, Qiao Li, Alistair E Johnson, and Roger G Mark. 2017. AF classification from a short single lead ECG recording: The PhysioNet/computing in cardiology challenge 2017. In *2017 computing in cardiology (CinC)*. IEEE, 1–4.

[44] Miguel Contreras, Sumit Kapoor, Jiaqing Zhang, Andrea Davidson, Yuanfang Ren, Ziyuan Guan, Tezcan Ozrazgat-Baslanti, Jessica Sena, Subhash Nerella, Azra Bihorac, et al. 2025. DeLLiriuM: A large language model for delirium prediction in the ICU using structured EHR. *Research Square* (2025), rs–3.

[45] Estee Y Cramer, Yuxin Huang, Yijin Wang, Evan L Ray, Matthew Cornell, Johannes Bracher, Andrea Brennen, Alvaro J Castro Rivadeneira, Aaron Gerding, Katie House, et al. 2022. The United States COVID-19 forecast hub dataset. *Scientific data* 9, 1 (2022), 462.

[46] Hejie Cui, Alyssa Unell, Bowen Chen, Jason Alan Fries, Emily Alsentzer, Sanmi Koyejo, and Nigam H Shah. 2025. Timer: Temporal instruction modeling and evaluation for longitudinal clinical records. *npj Digital Medicine* 8, 1 (2025), 577.

[47] Yue Cui, Chen Zhu, Guanyu Ye, Ziwei Wang, and Kai Zheng. 2021. Into the unobservables: A multi-range encoder-decoder framework for COVID-19 prediction. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 292–301.

[48] Zongyu Dai, Emily Getzen, and Qi Long. 2024. Sadi: Similarity-aware diffusion model-based imputation for incomplete temporal ehr data. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 4195–4203.

[49] Davy Darankoum, Chloé Habermacher, Julien Volle, and Sergei Grudinin. 2026. SpecMoE: Spectral Mixture-of-Experts Foundation Model for Cross-Species EEG Decoding. arXiv:2603.16739 [cs.LG]

[50] Songgaojun Deng, Shusen Wang, Huzefa Rangwala, Lijing Wang, and Yue Ning. 2020. Cola-GNN: Cross-location attention based graph neural networks for long-term ILI prediction. In *Proceedings of the 29th ACM international conference on information & knowledge management*. 245–254.

[51] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*. 4171–4186.

[52] Mrinmoy Dey, Aprameyo Chakrabartty, Dhruv Sarkar, and Tanujit Chakraborty. 2024. Do we really need Foundation Models for multi-step-ahead Epidemic Forecasting?. In *NeurIPS Workshop on Time Series in the Age of Large Models*.

[53] Cheng Ding, Zhicheng Guo, Zhaoliang Chen, Randall J Lee, Cynthia Rudin, and Xiao Hu. 2024. SiamQuality: A ConvNet-based foundation model for imperfect physiological signals. *arXiv preprint arXiv:2404.17667* (2024).

[54] Ensheng Dong, Hongru Du, and Lauren Gardner. 2020. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases* 20, 5 (2020), 533–534.

[55] Hongru Du, Yang Zhao, Jianan Zhao, Shaochong Xu, Xihong Lin, Yiran Chen, Lauren M Gardner, and Hao 'Frank' Yang. 2025. Advancing real-time infectious disease forecasting using large language models. *Nature Computational Science* 5, 6 (2025), 467–480.

[56] Wenjie Du, David Côté, and Yan Liu. 2023. Saits: Self-attention-based imputation for time series. *Expert Systems with Applications* 219 (2023), 119619.

[57] Vanja Dukic, Hedibert F Lopes, and Nicholas G Polson. 2012. Tracking epidemics with Google flu trends data and a state-space SEIR model. *J. Amer. Statist. Assoc.* 107, 500 (2012), 1410–1426.

[58] Emadeldeen Eldele, Mohamed Ragab, Zhenghua Chen, Min Wu, Chee Keong Kwoh, Xiaoli Li, and Cuntai Guan. 2021. Time-Series Representation Learning via Temporal and Contextual Contrasting. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*. 2352–2359.

[59] Eray Erturk, Fahad Kamran, Salar Abbaspourazad, Sean Jewell, Harsh Sharma, Yujie Li, Sinead Williamson, Nicholas J Foti, and Joseph Futoma. 2025. Beyond Sensor Data: Foundation Models of Behavioral Data from Wearables Improve Health Predictions. In *Forty-second International Conference on Machine Learning*.

[60] Jamal Esmaelpoor, Mohammad Hassan Moradi, and Abdolrahim Kadkhodamohammadi. 2020. A multistage deep neural network model for blood pressure estimation using photoplethysmogram signals. *Computers in Biology and Medicine* 120 (2020), 103719.

[61] Daniel Fadlon, David Dov, Aviya Bennett, Daphna Heller-Miron, Gad Levy, Kfir Bar, and Ahuva Weiss-Meilik. 2025. Decode Like a Clinician: Enhancing LLM Fine-Tuning with Temporal Structured Data Representation. In *Proceedings of the 14th International Joint Conference on Natural Language Processing and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics*. 1906–1922.

[62] Panpan Fan, Xiaodong Yue, Yufei Chen, Zhipeng Wei, Jie Shi, and Zhikang Xu. 2025. ECG Report Generation with Diagnostic Knowledge Enhanced Prompt Learning. In *Workshop on Clinical Image-Based Procedures*. Springer, 11–20.

[63] Cathy Mengying Fang, Valdemar Danry, Nathan Whitmore, Andria Bao, Andrew Hutchison, Cayden Pierce, and Pattie Maes. 2024. Physiollm: Supporting personalized health insights with wearables and large language models. In *2024 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*. IEEE, 1–8.

[64] Xiaocheng Fang, Zhengyao Ding, Jieyi Cai, Yujie Xiao, Bo Liu, Jiarui Jin, Haoyu Wang, Guangkun Nie, Shun Huang, Ting Chen, et al. 2026. ECGFlowCMR: Pretraining with ECG-Generated Cine CMR Improves Cardiac Disease Classification and Phenotype Prediction. *arXiv preprint arXiv:2601.20904* (2026).

[65] Xiaocheng Fang, Jiarui Jin, Haoyu Wang, Che Liu, Jieyi Cai, Yujie Xiao, Guangkun Nie, Bo Liu, Shun Huang, Hongyan Li, et al. 2025. PPGFlowECG: Latent Rectified Flow with Cross-Modal Encoding for PPG-Guided ECG Generation and Cardiovascular Disease Detection. *arXiv preprint arXiv:2509.19774* (2025).

[66] Mohammad Feli, Iman Azimi, Pasi Liljeberg, and Amir M Rahmani. 2025. An llm-powered agent for physiological data analysis: A case study on ppg-based heart rate estimation. In *2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 1–7.

[67] Tao Feng, Sirui Song, Tong Xia, and Yong Li. 2023. Contact tracing and epidemic intervention via deep reinforcement learning. *ACM Transactions on Knowledge Discovery from Data* 17, 3 (2023), 1–24.

[68] Junyi Gao, Cao Xiao, Yasha Wang, Wen Tang, Lucas M Glass, and Jimeng Sun. 2020. Stagenet: Stage-aware neural networks for health risk prediction. In *Proceedings of the web conference 2020*. 530–540.

[69] Yanjun Gao, Skatje Myers, Shan Chen, Dmitriy Dligach, Timothy A Miller, Danielle Bitterman, Matthew Churpek, and Majid Afshar. 2024. When Raw Data Prevails: Are Large Language Model Embeddings Effective in Numerical Data Representation for Medical Machine Learning Applications?. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen (Eds.). Association for Computational Linguistics, Miami, Florida, USA, 5414–5428.

[70] Tomer Golany, Daniel Freedman, and Kira Radinsky. 2021. Ecg ode-gan: Learning ordinary differential equations of ecg dynamics via generative adversarial learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 134–141.

[71] Tomer Golany and Kira Radinsky. 2019. Pgans: Personalized generative adversarial networks for ecg synthesis to improve patient-specific deep ecg classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 557–564.

[72] Tomer Golany, Kira Radinsky, and Daniel Freedman. 2020. SimGANs: Simulator-based generative adversarial networks for ECG synthesis to improve deep ECG classification. In *International Conference on Machine Learning*. PMLR, 3597–3606.

[73] Brian Gow, Tom Pollard, Larry A Nathanson, Alistair Johnson, Benjamin Moody, Chrystinne Fernandes, Nathaniel Greenbaum, Jonathan W Waks, Parastou Eslami, Tanner Carbonati, et al. 2023. MIMIC-IV-ECG: Diagnostic Electrocardiogram Matched Subset. *Type: dataset* (2023).

[74] Nicholas C Grassly and Christophe Fraser. 2008. Mathematical models of infectious disease transmission. *Nature Reviews Microbiology* 6, 6 (2008), 477–487.

[75] Nate Gruver, Marc Finzi, Shikai Qiu, and Andrew G Wilson. 2023. Large Language Models Are Zero-Shot Time Series Forecasters. In *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Eds.), Vol. 36. Curran Associates, Inc., 19622–19635.

[76] Xiao Gu, Zhangdaihong Liu, Jinpei Han, Jianing Qiu, Wenfei Fang, Lei Lu, Lei Clifton, Yuan-Ting Zhang, and David A Clifton. 2025. Transforming label-efficient decoding of healthcare wearables with self-supervised learning and "embedded" medical domain expertise. *Communications Engineering* 4, 1 (2025), 135.

[77] Xiao Gu, Yuxuan Shu, Jinpei Han, Yuxuan Liu, Zhangdaihong Liu, James Anibal, Veer Sangha, Edward Phillips, Bradley Segal, Hang Yuan, et al. 2025. Foundation models for biosignals: A survey. *Authorea Preprints* (2025).

[78] Lin Lawrence Guo, Stephen R Pfohl, Jason Fries, Alistair EW Johnson, Jose Posada, Catherine Aftandilian, Nigam Shah, and Lillian Sung. 2022. Evaluation of domain generalization and adaptation on improving model robustness to temporal dataset shift in clinical medicine. *Scientific reports* 12, 1 (2022), 2726.

[79] Lu Han, Han-Jia Ye, and De-Chuan Zhan. 2024. The capacity and robustness trade-off: Revisiting the channel independent strategy for multivariate time series forecasting. *IEEE Transactions on Knowledge and Data Engineering* 36, 11 (2024), 7129–7142.

[80] Awni Y Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H Tison, Codie Bourn, Mintu P Turakhia, and Andrew Y Ng. 2019. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine* 25, 1 (2019), 65–69.

[81] Hrayr Harutyunyan, Hrant Khachatrian, David C Kale, Greg Ver Steeg, and Aram Galstyan. 2019. Multitask learning and benchmarking with clinical time series data. *Scientific data* 6, 1 (2019), 96.

[82] Nasir Hayat, Krzysztof J. Geras, and Farah E. Shamout. 2022. MedFuse: Multi-modal fusion with clinical time-series data and chest X-ray images. In *Proceedings of the 7th Machine Learning for Healthcare Conference (Proceedings of Machine Learning Research, Vol. 182)*, Zachary Lipton, Rajesh Ranganath, Mark Sendak, Michael Sjoding, and Serena Yeung (Eds.). PMLR, 479–503.

[83] Huan He, William hao, Yuanzhe Xi, Yong Chen, Bradley Malin, and Joyce Ho. 2024. A Flexible Generative Model for Heterogeneous Tabular EHR with Missing Modality. In *The Twelfth International Conference on Learning Representations*.

[84] Huan He, Owen Queen, Teddy Koker, Consuelo Cuevas, Theodoros Tsiligkaridis, and Marinka Zitnik. 2023. Domain adaptation for time series under feature and label shifts. In *International conference on machine learning*. PMLR, 12746–12774.

[85] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. 2022. Masked autoencoders are scalable vision learners. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 16000–16009.

[86] Hoda Helmy, Ahmed Ibrahim, Maryam Arabi, Aamenah Sattar, and Ahmed Serag. 2025. Leveraging large language models to predict unplanned ICU readmissions from electronic health records. *Natural Language Processing Journal* (2025), 100182.

[87] A Ali Heydari, Ken Gu, Vidya Srinivas, Hong Yu, Zhihan Zhang, Yuwei Zhang, Akshay Paruchuri, Qian He, Hamid Palangi, Nova Hammerquist, et al. 2025. The anatomy of a personal health agent. *arXiv preprint arXiv:2508.20148* (2025).

[88] Nan Huang, Haishuai Wang, Zihuai He, Marinka Zitnik, and Xiang Zhang. 2024. Repurposing foundation model for generalizable medical time series classification. *arXiv preprint arXiv:2410.03794* (2024).

[89] Jiahao Ji, Jingyuan Wang, Chao Huang, Junjie Wu, Boren Xu, Zhenhe Wu, Junbo Zhang, and Yu Zheng. 2023. Spatio-temporal self-supervised learning for traffic flow prediction. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 37. 4356–4364.

[90] Zongliang Ji, Yifei Sun, Andre Amaral, Anna Goldenberg, and Rahul G Krishnan. 2026. Can we generate portable representations for clinical time series data using LLMs? *arXiv preprint arXiv:2603.23987* (2026).

[91] Ziyu Jia, Haichao Wang, Yucheng Liu, and Tianzi Jiang. 2024. Mutual distillation extracting spatial-temporal knowledge for lightweight multi-channel sleep stage classification. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1279–1289.

[92] Wei-Bang Jiang, Li-Ming Zhao, and Bao-Liang Lu. 2024. Large brain model for learning generic representations with tremendous EEG data in BCI. *arXiv preprint arXiv:2405.18765* (2024).

[93] Ming Jin, Shiyu Wang, Lintao Ma, Zhixuan Chu, James Zhang, Xiaoming Shi, Pin-Yu Chen, Yuxuan Liang, Yuan-Fang Li, Shirui Pan, et al. 2024. Time-llm: Time series forecasting by reprogramming large language models. In *International conference on learning representations*, Vol. 2024. 23857–23880.

[94] Alistair EW Johnson, Lucas Bulgarelli, Lu Shen, Alvin Gayles, Ayad Shammout, Steven Horng, Tom J Pollard, Sicheng Hao, Benjamin Moody, Brian Gow, et al. 2023. MIMIC-IV, a freely accessible electronic health record dataset. *Scientific data* 10, 1 (2023), 1.

[95] Alistair EW Johnson, Tom J Pollard, Lu Shen, Li-wei H Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G Mark. 2016. MIMIC-III, a freely accessible critical care database. *Scientific data* 3, 1 (2016), 1–9.

[96] Baraa Al Jorf and Farah Shamout. 2025. MedPatch: Confidence-Guided Multi-Stage Fusion for Multimodal Clinical Data. *arXiv preprint arXiv:2508.09182* (2025).

[97] Suprabhath Kalahasti, Benjamin Faucher, Boxuan Wang, Claudio Ascione, Ricardo Carbajal, Maxime Enault, Christophe Vincent Cassis, Titouan Launay, Caroline Guerrisi, Pierre-Yves Boëlle, et al. 2026. Foundation models for time series forecasting and policy evaluation in infectious disease epidemics. *Epidemics* (2026), 100916.

[98] Kranti Kamble and Joydeep Sengupta. 2023. A comprehensive survey on emotion recognition based on electroencephalograph (EEG) signals. *Multimedia Tools and Applications* 82, 18 (2023), 27269–27304.

[99] Amol Kapoor, Xue Ben, Luyang Liu, Bryan Perozzi, Matt Barnes, Martin Blais, and Shawn O'Banion. 2020. Examining covid-19 forecasting using spatio-temporal graph neural networks. *arXiv preprint arXiv:2007.03113* (2020).

[100] Hojjat Karami, Mary-Anne Hartley, David Atienza, and Anisoara Ionescu. 2025. Timehr: Image-based time series generation for electronic health records. *IEEE Journal of Biomedical and Health Informatics* (2025).

[101] Antonis Karantonis, Konstantinos Barmpas, Dimitrios Adamos, Nikolaos Laskaris, Stefanos Zafeiriou, and Yannis Panagakis. 2025. Subject-Aware Contrastive Learning for EEG Foundation Models. In *NeurIPS 2025 Workshop on Learning from Time Series for Health*.

[102] Bob Kemp, Aeilko H Zwinderman, Bert Tuk, Hilbert AC Kamphuisen, and Josefien JL Oberye. 2000. Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering* 47, 9 (2000), 1185–1194.

[103] Swaraj Khadanga, Karan Aggarwal, Shafiq Joty, and Jaideep Srivastava. 2019. Using clinical notes with time series data for ICU management. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. 6432–6437.

[104] Patrick Kidger, James Morrill, James Foster, and Terry Lyons. 2020. Neural controlled differential equations for irregular time series. *Advances in neural information processing systems* 33 (2020), 6696–6707.

[105] Jiwon Kim, Hyolim Jeon, Minhan Cho, Jiwon Kang, Shibo He, and Jinyoung Han. 2026. Multimodal learning for early prediction of COVID-19 outbreaks. *Information Processing & Management* 63, 3 (2026), 104562.

[106] Yubin Kim, Xuhai Xu, Daniel McDuff, Cynthia Breazeal, and Hae Won Park. 2024. Health-llm: Large language models for health prediction via wearable sensor data. *arXiv preprint arXiv:2401.06866* (2024).

[107] Ryan King, Shivesh Kodali, Conrad Krueger, Tianbao Yang, and Bobak J Mortazavi. 2024. An efficient contrastive unimodal pretraining method for ehr time series data. In *2024 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*. IEEE, 1–8.

[108] Ryan King, Tianbao Yang, and Bobak J Mortazavi. 2023. Multimodal pretraining of medical time series and notes. In *Machine Learning for Health (ML4H)*. PMLR, 244–255.

[109] Dani Kiyasseh, Tingting Zhu, and David A Clifton. 2021. Clocs: Contrastive learning of cardiac signals across space, time, and patients. In *International conference on machine learning*. PMLR, 5606–5615.

[110] Matthieu Komorowski, Leo A Celi, Omar Badawi, Anthony C Gordon, and A Aldo Faisal. 2018. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nature medicine* 24, 11 (2018), 1716–1720.

[111] Chrysoula Kosma, Giannis Nikolentzos, George Panagopoulos, Jean-Marc Steyaert, and Michalis Vazirgiannis. 2023. Neural ordinary differential equations for modeling epidemic spreading. *Transactions on Machine Learning Research* (2023).

[112] Maya Kruse, Shiyue Hu, Nicholas Derby, Yifu Wu, Samantha Stonbraker, Bingsheng Yao, Dakuo Wang, Elizabeth Goldberg, and Yanjun Gao. 2025. Large language models with temporal reasoning for longitudinal clinical summarization and prediction. In *Findings of ACL. EMNLP. Conference on Empirical Methods in Natural Language Processing*, Vol. 2025. 20715–20735.

[113] Yongfan Lai, Jiabo Chen, Qinghao Zhao, Deyun Zhang, Yue Wang, Shijia Geng, Hongyan Li, and Shenda Hong. 2025. Diffusets: 12-lead ecg generation conditioned on clinical text reports and patient-specific information. *Patterns* 6, 10 (2025).

[114] Remi Lam, Alvaro Sanchez-Gonzalez, Matthew Willson, Peter Wirnsberger, Meire Fortunato, Ferran Alet, Suman Ravuri, Timo Ewalds, Zach Eaton-Rosen, Weihua Hu, et al. 2023. Learning skillful medium-range global weather forecasting. *Science* 382, 6677 (2023), 1416–1421.

[115] Xiang Lan, Feng Wu, Kai He, Qinghao Zhao, Shenda Hong, and Mengling Feng. 2025. GEM: Empowering MLLM for Grounded ECG Understanding with Time Series and Images. In *Advances in Neural Information Processing Systems*, D. Belgrave, C. Zhang, H. Lin, R. Pascanu, P. Koniusz, M. Ghassemi, and N. Chen (Eds.), Vol. 38. Curran Associates, Inc., 94421–94455.

[116] Max SY Lau, C Jessica E Metcalf, Zewen Liu, Bryan T Grenfell, and Wei Jin. 2026. Toward AI foundation models for epidemics: Promise, challenges, and paths forward. *Proceedings of the National Academy of Sciences* 123, 13 (2026), e2526192123.

[117] Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. 2018. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of neural engineering* 15, 5 (2018), 056013.

[118] Changhee Lee, Jinsung Yoon, and Mihaela Van Der Schaar. 2019. Dynamic-deephit: A deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. *IEEE Transactions on Biomedical Engineering* 67, 1 (2019), 122–133.

[119] Gabriel Jason Lee, Jathurshan Pradeepkumar, and Jimeng Sun. 2026. Test-Time Adaptation for EEG Foundation Models: A Systematic Study under Real-World Distribution Shifts. *arXiv preprint arXiv:2604.16926* (2026).

[120] Hyojin Lee, You Rim Choi, Hyun Kyung Lee, Jaemin Jeong, Joopyo Hong, Hyun-Woo Shin, and Hyung-Sin Kim. 2025. Explainable vision transformer for automatic visual sleep staging on multimodal PSG signals. *npj Digital Medicine* 8, 1 (2025), 55.

[121] Simon A Lee, Sujay Jain, Alex Chen, Kyoka Ono, Arabdha Biswas, Ákos Rudas, Jennifer Fang, and Jeffrey N Chiang. 2025. Clinical decision support using pseudo-notes from multiple streams of EHR data. *npj Digital Medicine* 8, 1 (2025), 394.

[122] Yurim Lee, Eunji Jun, Jaehun Choi, and Heung-Il Suk. 2022. Multi-view integrative attention-based deep representation learning for irregular clinical time-series data. *IEEE Journal of Biomedical and Health Informatics* 26, 8 (2022), 4270–4280.

[123] Boyuan Li, Zhen Liu, Yicheng Luo, and Qianli Ma. 2026. Learning Recursive Multi-Scale Representations for Irregular Multivariate Time Series Forecasting. In *The Fourteenth International Conference on Learning Representations*.

[124] Boyuan Li, Yicheng Luo, Zhen Liu, Junhao Zheng, Jianming Lv, and Qianli Ma. 2025. HyperIMTS: Hypergraph Neural Network for Irregular Multivariate Time Series Forecasting. In *Forty-second International Conference on Machine Learning*.

[125] Chenqi Li, Timothy Denison, and Tingting Zhu. 2024. A survey of few-shot learning for biomedical time series. *IEEE Reviews in Biomedical Engineering* 18 (2024), 192–210.

[126] Hao Li, Bowen Deng, Chang Xu, ZhiYuan Feng, Viktor Schlegel, Yu-Hao Huang, Yizheng Sun, Jingyuan Sun, Kailai Yang, Yiyao Yu, et al. 2025. MIRA: Medical Time Series Foundation Model for Real-World Health Data. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

[127] Jun Li, Aaron D Aguirre, Valdery Moura Junior, Jiarui Jin, Che Liu, Lanhai Zhong, Chenxi Sun, Gari Clifford, M Brandon Westover, and Shenda Hong. 2025. An electrocardiogram foundation model built on over 10 million recordings. *Nejm ai* 2, 7 (2025), AIoa2401033.

[128] Jun Li, Che Liu, Sibo Cheng, Rossella Arcucci, and Shenda Hong. 2024. Frozen Language Model Helps ECG Zero-Shot Learning. In *Medical Imaging with Deep Learning (Proceedings of Machine Learning Research, Vol. 227)*, Ipek Oguz, Jack Noble, Xiaoxiao Li, Martin Styner, Christian Baumgartner, Mirabela Rusu, Tobias Heinmann, Despina Kontos, Bennett Landman, and Benoit Dawant (Eds.). PMLR, 402–415.

[129] Liang Li, Yuewen Jiang, and Biqing Huang. 2021. Long-term prediction for temporal propagation of seasonal influenza using transformer-based model. *Journal of biomedical informatics* 122 (2021).

[130] Wei Li, Ping Shi, and Hongliu Yu. 2021. Gesture recognition using surface electromyography and deep learning for prostheses hand: state-of-the-art, challenges, and future. *Frontiers in neuroscience* 15 (2021), 621885.

[131] Yamin Li, Ange Lou, Ziyuan Xu, Shengchao Zhang, Shiyu Wang, Dario J Englot, Soheil Kolouri, Daniel Moyer, Roza G Bayrak, and Catie Chang. 2024. Neurobolt: Resting-state eeg-to-fmri synthesis with multi-dimensional feature mapping. In *Advances in neural information processing systems*, Vol. 37. 23378–23405.

[132] Yikuan Li, Mohammad Mamouei, Gholamreza Salimi-Khorshidi, Shishir Rao, Abdelaali Hassaine, Dexter Canoy, Thomas Lukasiewicz, and Kazem Rahimi. 2022. Hi-BEHRT: hierarchical transformer-based model for accurate prediction of clinical events using multimodal longitudinal electronic health records. *IEEE journal of biomedical and health informatics* 27, 2 (2022), 1106–1117.

[133] Yikuan Li, Shishir Rao, José Roberto Ayala Solares, Abdelaali Hassaine, Rema Ramakrishnan, Dexter Canoy, Yajie Zhu, Kazem Rahimi, and Gholamreza Salimi-Khorshidi. 2020. BEHRT: transformer for electronic health records. *Scientific reports* 10, 1 (2020), 7155.

[134] Yanke Li, Ruirui Wang, Manuel Günther, and Diego Paez-Granados. 2026. GARLIC: Graph Attention-based Relational Learning of Multivariate Time Series in Intensive Care. In *The Fourteenth International Conference on Learning Representations*.

[135] Zhe Li, Shiyi Qi, Yiduo Li, and Zenglin Xu. 2023. Revisiting long-term time series forecasting: An investigation on linear mapping. *arXiv preprint arXiv:2305.10721* (2023).

[136] Zachary C Lipton, David C Kale, Randall Wetzel, et al. 2016. Modeling missing data in clinical time series with rnns. *Machine Learning for Healthcare* 56, 56 (2016), 253–270.

[137] Chunyu Liu, Yongpei Ma, Kavitha Kothur, Armin Nikpour, and Omid Kavehei. 2023. BioSignal Copilot: Leveraging the power of LLMs in drafting reports for biomedical signals. *medRxiv* (2023), 2023–06.

[138] Che Liu, Zhongwei Wan, Cheng Ouyang, Anand Shah, Wenjia Bai, and Rossella Arcucci. 2024. Zero-Shot ECG Classification with Multimodal Learning and Test-time Clinical Knowledge Enhancement. In *Forty-first International Conference on Machine Learning*.

[139] Haoxin Liu, Shangqing Xu, Zhiyuan Zhao, Lingkai Kong, Harshavardhan Kamarthi, Aditya B. Sasanur, Megha Sharma, Jiaming Cui, Qingsong Wen, Chao Zhang, and B. Aditya Prakash. 2024. Time-MMD: Multi-Domain Multimodal Dataset for Time Series Analysis. In *The Thirty-eight Conference on Neural Information Processing Systems Datasets and Benchmarks Track*.

[140] Wen-Cheng Liu, Chin-Sheng Lin, Chien-Sung Tsai, Tien-Ping Tsao, Cheng-Chung Cheng, Jun-Ting Liou, Wei-Shiang Lin, Shu-Meng Cheng, Yu-Sheng Lou, Chia-Cheng Lee, et al. 2021. A deep learning algorithm for detecting acute myocardial infarction: Deep learning model to detect AMI. *EuroIntervention* 17, 9 (2021), 765.

[141] Xu Liu, Juncheng Liu, Gerald Woo, Taha Aksu, Yuxuan Liang, Roger Zimmermann, Chenghao Liu, Junnan Li, Silvio Savarese, Caiming Xiong, and Doyen Sahoo. 2025. Moirai-MoE: Empowering Time Series Foundation Models with Sparse Mixture of Experts. In *Proceedings of the 42nd International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 267)*, Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (Eds.). PMLR, 38940–38962.

[142] Xin Liu, Daniel McDuff, Geza Kovacs, Isaac Galatzer-Levy, Jacob Sunshine, Jiening Zhan, Ming-Zher Poh, Shun Liao, Paolo Di Achille, and Shwetak Patel. 2023. Large language models are few-shot health learners. *arXiv preprint arXiv:2305.15525* (2023).

[143] Xvyuan Liu, Xiangfei Qiu, Hanyin Cheng, Xingjian Wu, Chenjuan Guo, Bin Yang, and Jilin Hu. 2026. ASTGI: Adaptive Spatio-Temporal Graph Interactions for Irregular Multivariate Time Series Forecasting. In *The Fourteenth International Conference on Learning Representations*.

[144] Xingyue Liu, Feizhong Zhou, Hanguang Xiao, Zhipeng Li, Shuai Liu, and Lingling Qian. 2026. A Survey on Large Language Models for Medical Time Series. *Expert Systems with Applications* (2026), 131364.

[145] Ziyu Liu, Azadeh Alavi, Minyi Li, and Xiang Zhang. 2023. Self-supervised contrastive learning for medical time series: A systematic review. *Sensors* 23, 9 (2023), 4221.

[146] Zitao Liu and Milos Hauskrecht. 2016. Learning adaptive forecasting models from irregularly sampled multivariate clinical data. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 30.

[147] Ziyang Liu, Siyuan He, Feng Liang, Chang Huang, Shuxin Zhong, and Kaishun Wu. 2026. mmJEPA-ECG: Cross-Posture Robust Contactless Electrocardiogram Monitoring via Millimeter Wave Radar Sensing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 40. 38962–38970.

[148] Zewen Liu, Juntong Ni, Bohan Wang, Max SY Lau, and Wei Jin. 2025. Pre-training Epidemic Time Series Forecasters with Compartmental Prototypes. *arXiv preprint arXiv:2502.03393* (2025).

[149] Zewen Liu, Guancheng Wan, B Aditya Prakash, Max SY Lau, and Wei Jin. 2024. A review of graph neural networks in epidemic modeling. In *Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining*. 6577–6587.

[150] Mohammad Loni, Fatemeh Poursalim, Mehdi Asadi, and Arash Gharehbaghi. 2025. A review on generative AI models for synthetic medical text, time series, and longitudinal data. *npj Digital Medicine* 8, 1 (2025), 281.

[151] Junyu Luo, Muchao Ye, Cao Xiao, and Fenglong Ma. 2020. 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*. 647–656.

[152] Yicheng Luo, Zhen Liu, Linghao Wang, Binquan Wu, Junhao Zheng, and Qianli Ma. 2024. Knowledge-Empowered Dynamic Graph Network for Irregularly Sampled Medical Time Series. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.

[153] Yicheng Luo, Bowen Zhang, Zhen Liu, and Qianli Ma. 2025. Hi-Patch: Hierarchical Patch GNN for Irregular Multivariate Time Series. In *Forty-second International Conference on Machine Learning*.

[154] Weimin Lyu, Xinyu Dong, Rachel Wong, Songzhu Zheng, Kayley Abell-Hart, Fusheng Wang, and Chao Chen. 2023. A multimodal transformer: Fusing clinical notes with structured ehr data for interpretable in-hospital mortality prediction. In *AMIA Annual Symposium Proceedings*, Vol. 2022. 719.

[155] Fenglong Ma, Radha Chitta, Jing Zhou, Quanzeng You, Tong Sun, and Jing Gao. 2017. 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*. 1903–1911.

[156] Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. 2024. Segment anything in medical images. *Nature communications* 15, 1 (2024), 654.

[157] Liantao Ma, Junyi Gao, Yasha Wang, Chaohe Zhang, Jiangtao Wang, Wenjie Ruan, Wen Tang, Xin Gao, and Xinyu Ma. 2020. 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. 825–832.

[158] Sarabeth M Mathis, Alexander E Webber, Tomás M León, Erin L Murray, Monica Sun, Lauren A White, Logan C Brooks, Alden Green, Addison J Hu, Roni Rosenfeld, et al. 2024. Evaluation of FluSight influenza forecasting in the 2021–22 and 2022–23 seasons with a new target laboratory-confirmed influenza hospitalizations. *Nature communications* 15, 1 (2024), 6289.

[159] Kaden McKeen, Sameer Masood, Augustin Toma, Barry Rubin, and Bo Wang. 2025. Ecg-fm: An open electrocardiogram foundation model. *Jamia Open* 8, 5 (2025), ooaf122.

[160] Scott Mayer McKinney, Marcin Sieniek, Varun Godbole, Jonathan Godwin, Natasha Antropova, Hutan Ashrafian, Trevor Back, Mary Chesus, Greg S Corrado, Ara Darzi, et al. 2020. International evaluation of an AI system for breast cancer screening. *Nature* 577, 7788 (2020), 89–94.

[161] Patrick E McSharry, Gari D Clifford, Lionel Tarassenko, and Leonard A Smith. 2003. A dynamical model for generating synthetic electrocardiogram signals. *IEEE transactions on biomedical engineering* 50, 3 (2003), 289–294.

[162] Mike A Merrill, Akshay Paruchuri, Naghmeh Rezaei, Geza Kovacs, Javier Perez, Yun Liu, Erik Schenck, Nova Hammerquist, Jake Sunshine, Shyam Tailor, et al. 2026. Transforming wearable data into personal health insights using large language model agents. *Nature Communications* (2026).

[163] Sonaxy Mohanty, Airi Shimamura, Charles D Nicholson, Andrés D González, and Talayeh Razzaghi. 2025. Hierarchical Time Series Forecasting of COVID-19 Cases Using County-Level Clustering Data. In *Operations Research Forum*, Vol. 6. Springer, 28.

[164] George B Moody and Roger G Mark. 2001. The impact of the MIT-BIH arrhythmia database. *IEEE engineering in medicine and biology magazine* 20, 3 (2001), 45–50.

[165] Intae Moon, Stefan Groha, and Alexander Gusev. 2022. SurvLatent ODE: A Neural ODE based time-to-event model with competing risks for longitudinal data improves cancer-associated Venous Thromboembolism (VTE) prediction. In *Machine Learning for Healthcare Conference*. PMLR, 800–827.

[166] Mohammad Amin Morid, Olivia R Liu Sheng, and Joseph Dunbar. 2023. Time series prediction using deep learning methods in healthcare. *ACM Transactions on Management Information Systems* 14, 1 (2023), 1–29.

[167] Seiichi Morokuma, Toshinari Hayashi, Masatomo Kanegae, Yoshihiko Mizukami, Shinji Asano, Ichiro Kimura, Yuji Tateizumi, Hitoshi Ueno, Subaru Ikeda, and Kyuichi Niizeki. 2023. Deep learning-based sleep stage classification with cardiorespiratory and body movement activities in individuals with suspected sleep disorders. *Scientific reports* 13, 1 (2023), 17730.

[168] Daniel Neil, Michael Pfeiffer, and Shih-Chii Liu. 2016. Phased lstm: Accelerating recurrent network training for long or event-based sequences. *Advances in neural information processing systems* 29 (2016).

[169] Phuoc Nguyen, Truyen Tran, Nilmini Wickramasinghe, and Svetha Venkatesh. 2016. Deepr: a convolutional net for medical records. *IEEE journal of biomedical and health informatics* 21, 1 (2016), 22–30.

[170] Juntong Ni, Saurabh Kataria, Shengpu Tang, Carl Yang, Xiao Hu, and Wei Jin. 2025. PPG-Distill: Efficient Photoplethysmography Signals Analysis via Foundation Model Distillation. In *NeurIPS 2025 Workshop on Learning from Time Series for Health*.

[171] Guangkun Nie, Jiabao Zhu, Gongzheng Tang, Deyun Zhang, Shijia Geng, Qinghao Zhao, and Shenda Hong. 2024. A review of deep learning methods for photoplethysmography data. *arXiv preprint arXiv:2401.12783* (2024).

[172] Shahriar Noroozizadeh, Sayantan Kumar, and Jeremy C Weiss. 2025. Forecasting from clinical textual time series: Adaptations of the encoder and decoder language model families. *ArXiv* (2025), arXiv–2504.

[173] Shahriar Noroozizadeh, Jeremy C Weiss, and George H Chen. 2023. Temporal supervised contrastive learning for modeling patient risk progression. In *Machine Learning for Health (ML4H)*. PMLR, 403–427.

[174] Iyad Obeid and Joseph Picone. 2016. The temple university hospital EEG data corpus. *Frontiers in neuroscience* 10 (2016), 196.

[175] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366, 6464 (2019), 447–453.

[176] David Obst, Joseph De Vilmarest, and Yannig Goude. 2021. Adaptive methods for short-term electricity load forecasting during COVID-19 lockdown in France. *IEEE transactions on power systems* 36, 5 (2021), 4754–4763.

[177] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. 2018. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748* (2018).

[178] Boris N Oreshkin, Dmitri Carpov, Nicolas Chapados, and Yoshua Bengio. 2019. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. *arXiv preprint arXiv:1905.10437* (2019).

[179] Yassine El Ouahidi, Jonathan Lys, Philipp Thölke, Nicolas Farrugia, Bastien Pasdeloup, Vincent Gripon, Karim Jerbi, and Giulia Lioi. 2025. REVE: A Foundation Model for EEG - Adapting to Any Setup with Large-Scale Pretraining on 25,000 Subjects. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

[180] George Panagopoulos, Giannis Nikolentzos, and Michalis Vazirgiannis. 2021. Transfer graph neural networks for pandemic forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 35. 4838–4845.

[181] Chao Pang, Xinzhuo Jiang, Krishna S Kalluri, Matthew Spotnitz, RuiJun Chen, Adler Perotte, and Karthik Natarajan. 2021. CEHR-BERT: Incorporating temporal information from structured EHR data to improve prediction tasks. In *Machine learning for health*. PMLR, 239–260.

[182] Mathias Perslev, Michael Jensen, Sune Darkner, Poul Jørgen Jennum, and Christian Igel. 2019. U-Time: A Fully Convolutional Network for Time Series Segmentation Applied to Sleep Staging. In *Advances in Neural Information Processing Systems*, Vol. 32. 4415–4426.

[183] Huy Phan, Kaare Mikkelsen, Oliver Y Chén, Philipp Koch, Alfred Mertins, and Maarten De Vos. 2022. Sleeptransformer: Automatic sleep staging with interpretability and uncertainty quantification. *IEEE Transactions on Biomedical Engineering* 69, 8 (2022), 2456–2467.

[184] Rajasekhar Pillalamarri and Udhayakumar Shanmugam. 2025. A review on EEG-based multimodal learning for emotion recognition. *Artificial Intelligence Review* 58, 5 (2025), 131.

[185] Marco AF Pimentel, Alistair EW Johnson, Peter H Charlton, Drew Birrenkott, Peter J Watkinson, Lionel Tarassenko, and David A Clifton. 2016. Toward a robust estimation of respiratory rate from pulse oximeters. *IEEE Transactions on Biomedical Engineering* 64, 8 (2016), 1914–1923.

[186] Virginia E Pitzer, Cécile Viboud, Lone Simonsen, Claudia Steiner, Catherine A Panozzo, Wladimir J Alonso, Mark A Miller, Roger I Glass, John W Glasser, Umesh D Parashar, et al. 2009. Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. *Science* 325, 5938 (2009), 290–294.

[187] Tom J Pollard, Alistair EW Johnson, Jesse D Raffa, Leo A Celi, Roger G Mark, and Omar Badawi. 2018. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Scientific data* 5, 1 (2018), 180178.

[188] David Pollreisz and Nima TaheriNejad. 2022. Detection and removal of motion artifacts in PPG signals. *Mobile Networks and Applications* 27, 2 (2022), 728–738.

[189] Stuart F Quan, Barbara V Howard, Conrad Iber, James P Kiley, F Javier Nieto, George T O'Connor, David M Rapoport, Susan Redline, John Robbins, Jonathan M Samet, et al. 1997. The sleep heart health study: design, rationale, and methods. *Sleep* 20, 12 (1997), 1077–1085.

[190] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning Transferable Visual Models From Natural Language Supervision. In *Proceedings of the 38th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 139)*, Marina Meila and Tong Zhang (Eds.). PMLR, 8748–8763.

[191] Aniruddh Raghu, Payal Chandak, Ridwan Alam, John Guttag, and Collin Stultz. 2022. Contrastive Pre-Training for Multimodal Medical Time Series. In *NeurIPS 2022 Workshop on Learning from Time Series for Health*.

[192] Chava L Ramspek, Kitty J Jager, Friedo W Dekker, Carmine Zoccali, and Merel van Diepen. 2021. External validation of prognostic models: what, why, how, when and where? *Clinical kidney journal* 14, 1 (2021), 49–58.

[193] Shishir Rao, Nouman Ahmed, Gholamreza Salimi-Khorshidi, Christopher Yau, Huimin Su, Nathalie Conrad, Folkert W Asselbergs, Mark Woodward, Rod Jackson, John GF Cleland, et al. 2025. A Transformer-based survival model for prediction of all-cause mortality in heart failure patients: a multi-cohort study. *arXiv preprint arXiv:2503.12317* (2025).

[194] Laila Rasmy, Yang Xiang, Ziqian Xie, Cui Tao, and Degui Zhi. 2021. Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ digital medicine* 4, 1 (2021), 86.

[195] Likith Reddy, Vivek Talwar, Shanmukh Alle, Raju S Bapi, and U Deva Priyakumar. 2021. Imle-net: An interpretable multi-level multi-channel model for ecg classification. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 1068–1074.

[196] Nicholas G Reich, Logan C Brooks, Spencer J Fox, Sasikiran Kandula, Craig J McGowan, Evan Moore, Dave Osthus, Evan L Ray, Abhinav Tushar, Teresa K Yamana, et al. 2019. A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States. *Proceedings of the National Academy of Sciences* 116, 8 (2019), 3146–3154.

[197] Attila Reiss, Ina Indlekofer, Philip Schmidt, and Kristof Van Laerhoven. 2019. Deep PPG: Large-scale heart rate estimation with convolutional neural networks. *Sensors* 19, 14 (2019), 3079.

[198] Weijieying Ren, Jingxi Zhu, Zehao Liu, Tianxiang Zhao, and Vasant Honavar. 2025. A comprehensive survey of electronic health record modeling: From deep learning approaches to large language models. *arXiv preprint arXiv:2507.12774* (2025).

[199] Antônio H Ribeiro, Manoel Horta Ribeiro, Gabriela MM Paixão, Derick M Oliveira, Paulo R Gomes, Jéssica A Canazart, Milton PS Ferreira, Carl R Andersson, Peter W Macfarlane, Wagner Meira Jr, et al. 2020. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nature communications* 11, 1 (2020), 1760.

[200] Alexander Rodríguez, Jiaming Cui, Naren Ramakrishnan, Bijaya Adhikari, and B Aditya Prakash. 2023. Einns: epidemiologically-informed neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 37. 14453–14460.

[201] Alexander Rodriguez, Harshavardhan Kamarthi, Pulak Agarwal, Javen Ho, Mira Patel, Suchet Sapre, and B Aditya Prakash. 2024. Machine learning for data-centric epidemic forecasting. *Nature Machine Intelligence* 6, 10 (2024), 1122–1131.

[202] Ruichen Rong, Zifan Gu, Hongyin Lai, Tanna L Nelson, Tony Keller, Clark Walker, Kevin W Jin, Catherine Chen, Ann Marie Navar, Ferdinand Velasco, et al. 2025. A deep learning model for clinical outcome prediction using longitudinal inpatient electronic health records. *JAMIA open* 8, 2 (2025), ooaf026.

[203] Roni Rosenfeld and Ryan J Tibshirani. 2021. Epidemic tracking and forecasting: Lessons learned from a tumultuous year. *Proceedings of the National Academy of Sciences* 118, 51 (2021), e2111456118.

[204] Yulia Rubanova, Ricky TQ Chen, and David K Duvenaud. 2019. Latent ordinary differential equations for irregularly-sampled time series. *Advances in neural information processing systems* 32 (2019).

[205] Nagwan Abdel Samee, Noha F Mahmoud, Eman A Aldhahri, Ahsan Rafiq, Mohammed Saleh Ali Muthanna, and Ijaz Ahmad. 2022. RNN and BiLSTM fusion for accurate automatic epileptic seizure diagnosis using EEG signals. *Life* 12, 12 (2022), 1946.

[206] Pritam Sarkar and Ali Etemad. 2021. Cardiogan: Attentive generative adversarial network with dual discriminators for synthesis of ecg from ppg. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 488–496.

[207] Philip Schmidt, Attila Reiss, Robert Duerichen, Claus Marberger, and Kristof Van Laerhoven. 2018. Introducing wesad, a multimodal dataset for wearable stress and affect detection. In *Proceedings of the 20th ACM international conference on multimodal interaction*. 400–408.

[208] Vinit Shah, Eva Von Weltin, Silvia Lopez, James Riley McHugh, Lillian Veloso, Meysam Golmohammadi, Iyad Obeid, and Joseph Picone. 2018. The temple university hospital seizure detection corpus. *Frontiers in neuroinformatics* 12 (2018), 83.

[209] Abhinav Sharma, Ankur Kumar, Sanjay Dhanka, Surita Maini, Kamakshi Rautela, Surya Kant, Amanpreet Kaur, and Aastha Palta. 2026. Artificial Intelligence in Epidemic Modeling and Pandemic Preparedness: A Comprehensive Review. *Archives of Computational Methods in Engineering* (2026), 1–31.

[210] Supreeth P Shashikumar, Amit J Shah, Gari D Clifford, and Shamim Nemati. 2018. Detection of paroxysmal atrial fibrillation using attention-based bidirectional recurrent neural networks. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 715–723.

[211] Yidan Shen, Neville Mathew, Maham Rahimi, Deependra Dhakal, George Zouridakis, Xin Fu, and Renjie Hu. 2026. DMT: Demographic Conditioning, Morphology-Enhanced Transformer for Cuffless Blood Pressure Estimation from PPG Signals. *arXiv preprint arXiv:2606.11125* (2026).

[212] Xiaoming Shi, Shiyu Wang, Yuqi Nie, Dianqi Li, Zhou Ye, Qingsong Wen, and Ming Jin. 2025. Time-MoE: Billion-Scale Time Series Foundation Models with Mixture of Experts. In *The Thirteenth International Conference on Learning Representations*.

[213] Benjamin Shickel, Patrick James Tighe, Azra Bihorac, and Parisa Rashidi. 2017. Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE journal of biomedical and health informatics* 22, 5 (2017), 1589–1604.

[214] Artem Shmatko, Alexander Wolfgang Jung, Kumar Gaurav, Søren Brunak, Laust Hvas Mortensen, Ewan Birney, Tom Fitzgerald, and Moritz Gerstung. 2025. Learning the natural history of human disease with generative transformers. *Nature* 647, 8088 (2025), 248–256.

[215] Ali Hossam Shoeb. 2009. *Application of machine learning to epileptic seizure onset detection and treatment.* Ph. D. Dissertation. Massachusetts Institute of Technology.

[216] Debaditya Shome, Pritam Sarkar, and Ali Etemad. 2024. Region-disentangled diffusion model for high-fidelity ppg-to-ecg translation. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 38. 15009–15019.

[217] Satya Narayan Shukla and Benjamin Marlin. 2021. Multi-Time Attention Networks for Irregularly Sampled Time Series. In *International Conference on Learning Representations*.

[218] Huan Song, Deepta Rajan, Jayaraman Thiagarajan, and Andreas Spanias. 2018. Attend and diagnose: Clinical time series analysis using attention models. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 32.

[219] Sirui Song, Zefang Zong, Yong Li, Xue Liu, and Yang Yu. 2020. Reinforced epidemic control: Saving both lives and economy. *arXiv preprint arXiv:2008.01257* (2020).

[220] Gabriel Spadon, Shenda Hong, Bruno Brandoli, Stan Matwin, Jose F Rodrigues-Jr, and Jimeng Sun. 2021. Pay attention to evolution: Time series forecasting with deep graph-evolution learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 9 (2021), 5368–5384.

[221] Ethan Steinberg, Jason Alan Fries, Yizhe Xu, and Nigam Shah. 2024. MOTOR: A Time-to-Event Foundation Model For Structured Medical Records. In *The Twelfth International Conference on Learning Representations*.

[222] Michael A Stoto, Abbey Woolverton, John Kraemer, Pepita Barlow, and Michael Clarke. 2022. COVID-19 data are messy: analytic methods for rigorous impact analyses with imperfect data. *Globalization and Health* 18, 1 (2022), 2.

[223] Chenxi Sun, Shenda Hong, Moxian Song, and Hongyan Li. 2020. A review of deep learning methods for irregularly sampled medical time series data. *arXiv preprint arXiv:2010.12493* (2020).

[224] Chenxi Sun, Hongyan Li, Yaliang Li, and Shenda Hong. 2024. TEST: Text Prototype Aligned Embedding to Activate LLM's Ability for Time Series. In *The Twelfth International Conference on Learning Representations*.

[225] Amelia LM Tan, Emily J Getzen, Meghan R Hutch, Zachary H Strasser, Alba Gutiérrez-Sacristán, Trang T Le, Arianna Dagliati, Michele Morris, David A Hanauer, Bertrand Moal, et al. 2023. Informative missingness: What can we learn from patterns in missing laboratory data in the electronic health record? *Journal of biomedical informatics* 139 (2023), 104306.

[226] Mingtian Tan, Mike A Merrill, Vinayak Gupta, Tim Althoff, and Thomas Hartvigsen. 2024. Are language models actually useful for time series forecasting?. In *Advances in Neural Information Processing Systems*, Vol. 37. 60162–60191.

[227] Yinzhou Tang, Huandong Wang, and Yong Li. 2023. Enhancing spatial spread prediction of infectious diseases through integrating multi-scale human mobility dynamics. In *Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems*. 1–12.

[228] Zhengxu Tang, Bohan Wang, Yiming Lu, Zewen Liu, Max SY Lau, and Wei Jin. 2026. Large Language Models at Population Scale: A Survey and Taxonomy of Public Health Applications. (2026).

[229] Yusuke Tashiro, Jiaming Song, Yang Song, and Stefano Ermon. 2021. Csdi: Conditional score-based diffusion models for probabilistic time series imputation. *Advances in neural information processing systems* 34 (2021), 24804–24816.

[230] Muhang Tian, Bernie Chen, Allan Guo, Shiyi Jiang, and Anru R Zhang. 2024. Reliable generation of privacy-preserving synthetic electronic health record time series via diffusion models. *Journal of the American Medical Informatics Association* 31, 11 (2024), 2529–2539.

[231] Sindhu Tipirneni and Chandan K Reddy. 2022. Self-supervised transformer for sparse and irregularly sampled multivariate clinical time-series. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 16, 6 (2022), 1–17.

[232] Geoffrey H Tison, José M Sanchez, Brandon Ballinger, Avesh Singh, Jeffrey E Olgin, Mark J Pletcher, Eric Vittinghoff, Emily S Lee, Shannon M Fan, Rachel A Gladstone, et al. 2018. Passive detection of atrial fibrillation using a commercially available smartwatch. *JAMA cardiology* 3, 5 (2018), 23–30.

[233] Sana Tonekaboni, Danny Eytan, and Anna Goldenberg. 2021. Unsupervised representation learning for time series with temporal neighborhood coding. *arXiv preprint arXiv:2106.00750* (2021).

[234] Khanh-Tung Tran, Truong Son Hy, Lili Jiang, and Xuan-Son Vu. 2024. MGLEP: multimodal graph learning for modeling emerging pandemics with big data. *Scientific Reports* 14, 1 (2024), 16377.

[235] Tim K Tsang, Peng Wu, Yun Lin, Eric HY Lau, Gabriel M Leung, and Benjamin J Cowling. 2020. Effect of changing case definitions for COVID-19 on the epidemic curve and transmission parameters in mainland China: a modelling study. *The Lancet Public Health* 5, 5 (2020), e289–e296.

[236] Moritz Vandenhirtz, Samuel Ruipérez-Campillo, Simon Böhi, Sonia Laguna, Irene Cannistraci, Andrea Agostini, Ece Ozkan, Thomas M. Sutter, and Julia E. Vogt. 2026. Foundation Model for Cardiac Time Series via Masked Latent Attention. *arXiv preprint arXiv:2603.26475* (2026).

[237] Khuong Vo, Emad Kasaeyan Naeini, Amir Naderi, Daniel Jilani, Amir M Rahmani, Nikil Dutt, and Hung Cao. 2021. P2E-WGAN: ECG waveform synthesis from PPG with conditional wasserstein generative adversarial networks. In *Proceedings of the 36th Annual ACM Symposium on Applied Computing*. 1030–1036.

[238] Patrick Wagner, Nils Strodthoff, Ralf-Dieter Bousseljot, Dieter Kreiseler, Fatima I Lunze, Wojciech Samek, and Tobias Schaeffter. 2020. PTB-XL, a large publicly available electrocardiography dataset. *Scientific data* 7, 1 (2020), 154.

[239] Guancheng Wan, Zewen Liu, Xiaojun Shan, Max SY Lau, B. Aditya Prakash, and Wei Jin. 2025. EARTH: Epidemiology-Aware Neural ODE with Continuous Disease Transmission Graph. In *Forty-second International Conference on Machine Learning*.

[240] Zhongwei Wan, Che Liu, Xin Wang, Chaofan Tao, Hui Shen, Jing Xiong, Rossella Arcucci, Huaxiu Yao, and Mi Zhang. 2025. MEIT: Multimodal electrocardiogram instruction tuning on large language models for report generation. In *Findings of the association for computational linguistics: ACL 2025*. 14510–14527.

[241] Dongdong Wang, Shunpu Zhang, and Liqiang Wang. 2021. Deep epidemiological modeling by black-box knowledge distillation: an accurate deep learning model for COVID-19. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 15424–15430.

[242] Fuying Wang, Jiacheng Xu, and Lequan Yu. 2025. From Token to Rhythm: A Multi-Scale Approach for ECG-Language Pretraining. In *International Conference on Machine Learning*. PMLR, 65059–65074.

[243] Jing Wang, Xing Niu, Juyong Kim, Jie Shen, Tong Zhang, and Jeremy C Weiss. 2025. MIMIC-IV-Ext-22MCTS: A 22 Million-Event Temporal Clinical Time-Series Dataset for Risk Prediction. *Research Square* (2025), rs–3.

[244] Lijing Wang, Aniruddha Adiga, Jiangzhuo Chen, Adam Sadilek, Srinivasan Venkatramanan, and Madhav Marathe. 2022. Causalgnn: Causal-based graph neural networks for spatio-temporal epidemic forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 36. 12191–12199.

[245] Lijing Wang, Jiangzhuo Chen, and Madhav Marathe. 2019. DEFSI: Deep learning based epidemic forecasting with synthetic information. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 9607–9612.

[246] Nana Wang, Suli Wang, Gen Li, Pengfei Ren, and Hao Su. 2026. New Synthetic Goldmine: Hand Joint Angle-Driven EMG Data Generation Framework for Micro-Gesture Recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 40. 17787–17795.

[247] Xiaoda Wang, Ching Chang, Defu Cao, Kaiqiao Han, Fang Sun, Yue Huang, Minxiao Wang, Chang Xu, Xiao Luo, Runze Yan, et al. 2026. Position: Beyond Prediction: Toward Verifiable Physiological Waveform Reasoning with Foundation Models and Agentic LLMs. In *The Forty-Third International Conference on Machine Learning*.

[248] Xiaoda Wang, Kaiqiao Han, Yuhao Xu, Xiao Luo, Yizhou Sun, Wei Wang, and Carl Yang. 2026. SE-Diff: Simulator and Experience Enhanced Diffusion Model for Comprehensive ECG Generation. In *The Fourteenth International Conference on Learning Representations*.

[249] Xiaochen Wang, Junyu Luo, Jiaqi Wang, Ziyi Yin, Suhan Cui, Yuan Zhong, Yaqing Wang, and Fenglong Ma. 2023. Hierarchical pretraining on multimodal electronic health records. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. 2839–2852.

[250] Xiaoda Wang, Minxiao Wang, Kaiqiao Han, Defu Cao, Ching Chang, Yidan Shi, Runze Yan, Xiao Luo, Yan Liu, Xiao Hu, et al. 2026. PG-LRF: Physiology-Guided Latent Rectified Flow for Electro-Hemodynamic PPG-to-ECG Generation. *arXiv preprint arXiv:2605.12541* (2026).

[251] Yue Wang, Xu Cao, Yaojun Hu, Haochao Ying, Hongxia Xu, Ruijia Wu, James Matthew Rehg, Jimeng Sun, Jian Wu, and Jintai Chen. 2024. AnyECG: foundational models for multitask cardiac analysis in real-world settings. *arXiv preprint arXiv:2411.17711* (2024).

[252] Yihe Wang, Nan Huang, Taida Li, Yujun Yan, and Xiang Zhang. 2024. Medformer: A multi-granularity patching transformer for medical time-series classification. *Advances in Neural Information Processing Systems* 37 (2024), 36314–36341.

[253] Yuanlong Wang, Changchang Yin, and Ping Zhang. 2024. Multimodal risk prediction with physiological signals, medical images and clinical notes. *Heliyon* 10, 5 (2024).

[254] Yuqing Wang, Yun Zhao, Rachael Callcut, and Linda Petzold. 2022. Integrating physiological time series and clinical notes with transformer for early prediction of sepsis. *arXiv preprint arXiv:2203.14469* (2022).

[255] Zifeng Wang and Jimeng Sun. 2022. Survtrace: Transformers for survival analysis with competing events. In *Proceedings of the 13th ACM international conference on bioinformatics, computational biology and health informatics*. 1–9.

[256] Mingyang Wei, Dehai Min, Zewen Liu, Yuzhang Xie, Guanchen Wu, Carl Yang, Max SY Lau, Qi He, Lu Cheng, and Wei Jin. 2026. EpiQAL: Benchmarking Large Language Models in Epidemiological Question Answering for Enhanced Alignment and Reasoning. *arXiv preprint arXiv:2601.03471* (2026).

[257] Riling Wei, Xiaogang Xu, Yue Li, Yiyi Zhang, Jun Wang, and Hanjie Chen. 2024. PulseID: Multi-scale photoplethysmographic identification using a deep convolutional neural network. *Biomedical Signal Processing and Control* 88 (2024), 105609.

[258] Yishu Wei, Yu Deng, Cong Sun, Mingquan Lin, Hongmei Jiang, and Yifan Peng. 2024. Deep learning with noisy labels in medical prediction problems: a scoping review. *Journal of the American Medical Informatics Association* 31, 7 (2024), 1596–1607.

[259] Weining Weng, Yang Gu, Shuai Guo, Yuan Ma, Zhaohua Yang, Yuchen Liu, and Yiqiang Chen. 2025. Self-supervised learning for electroencephalogram: A systematic survey. *Comput. Surveys* 57, 12 (2025), 1–38.

[260] Michael Wornow, Rahul Thapa, Ethan Steinberg, Jason Fries, and Nigam Shah. 2023. Ehrshot: An ehr benchmark for few-shot evaluation of foundation models. *Advances in Neural Information Processing Systems* 36 (2023), 67125–67137.

[261] Yuexin Wu, Yiming Yang, Hiroshi Nishiura, and Masaya Saitoh. 2018. Deep learning for epidemiological predictions. In *The 41st international ACM SIGIR conference on research & development in information retrieval*. 1085–1088.

[262] Ting Xiang, Yanwei Jin, Zijun Liu, Lei Clifton, David A Clifton, Yiming Zhang, Quan Zhang, Nan Ji, and Yuanting Zhang. 2025. Dynamic beat-to-beat measurements of blood pressure using multimodal physiological signals and a hybrid CNN-LSTM model. *IEEE journal of biomedical and health informatics* 29, 8 (2025).

[263] Xiao Xiang, David Restrepo, Hyewon Jeong, Yugang Jia, and Leo Anthony Celi. 2026. Learning Representations from Incomplete EHR Data with Dual-Masked Autoencoding. *arXiv preprint arXiv:2602.15159* (2026).

[264] Feng Xie, Han Yuan, Yilin Ning, Marcus Eng Hock Ong, Mengling Feng, Wynne Hsu, Bibhas Chakraborty, and Nan Liu. 2022. Deep learning for temporal data representation in electronic health records: A systematic review of challenges and methodologies. *Journal of biomedical informatics* 126 (2022), 103980.

[265] Maxwell A Xu, Jaya Narain, Gregory Darnell, Haraldur T Hallgrimsson, Hyewon Jeong, Darren Forde, Richard Andres Fineman, Karthik Jayaraman Raghuram, James Matthew Rehg, and Shirley You Ren. 2025. RelCon: Relative Contrastive Learning for a Motion Foundation Model for Wearable Data. In *The Thirteenth International Conference on Learning Representations*.

[266] Yuhao Xu, Xiaoda Wang, Yi Wu, Wei Jin, Xiao Hu, and Carl Yang. 2025. ECG-MoE: Mixture-of-Expert Electrocardiogram Foundation Model. In *NeurIPS 2025 Workshop on Learning from Time Series for Health*.

[267] Yanbo Xu, Shangqing Xu, Manav Ramprassad, Alexey Tumanov, and Chao Zhang. 2023. TransEHR: Self-supervised transformer for clinical time series data. In *Machine Learning for Health (ML4H)*. PMLR, 623–635.

[268] Chaoqi Yang, M Westover, and Jimeng Sun. 2023. BIOT: Biosignal Transformer for Cross-data Learning in the Wild. In *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Eds.), Vol. 36. Curran Associates, Inc., 78240–78260.

[269] Jenny Yang, Hagen Triendl, Andrew AS Soltan, Mangal Prakash, and David A Clifton. 2024. Addressing label noise for electronic health records: insights from computer vision for tabular data. *BMC Medical Informatics and Decision Making* 24, 1 (2024), 183.

[270] Kai Yang, Massimo Hong, Jiahuan Zhang, Yizhen Luo, Suyuan Zhao, Ou Zhang, Xiaomao Yu, Jiawen Zhou, Liuqing Yang, Ping Zhang, et al. 2025. ECG-LM: understanding electrocardiogram with a large language model. *Health Data Science* 5 (2025), 0221.

[271] Xue Yang, Xuejun Qi, and Xiaobo Zhou. 2023. Deep learning technologies for time series anomaly detection in healthcare: A review. *Ieee Access* 11 (2023), 117788–117799.

[272] Zhichao Yang, Avijit Mitra, Weisong Liu, Dan Berlowitz, and Hong Yu. 2023. TransformEHR: transformer-based encoder-decoder generative model to enhance prediction of disease outcomes using electronic health records. *Nature communications* 14, 1 (2023), 7857.

[273] Hugo Yèche, Gideon Dresdner, Francesco Locatello, Matthias Hüser, and Gunnar Rätsch. 2021. Neighborhood contrastive learning applied to online patient monitoring. In *International conference on machine learning*. PMLR, 11964–11974.

[274] Hugo Yèche, Rita Kuznetsova, Marc Zimmermann, Matthias Hüser, Xinrui Lyu, Martin Faltys, and Gunnar Rätsch. 2021. HiRID-ICU-Benchmark–A comprehensive machine learning benchmark on high-resolution ICU data. *arXiv preprint arXiv:2111.08536* (2021).

[275] Yakir Yehuda and Kira Radinsky. 2026. ODE-Constrained Generative Modeling of Cardiac Dynamics for 12-Lead ECG Synthesis. *Transactions on Machine Learning Research* (2026).

[276] Yakir Yehuda and Kira Radinsky. 2026. PDE-Driven Spatiotemporal Generative Modeling for Multilead ECG Synthesis. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 40. 1453–1461.

[277] Jinsung Yoon, Michel Mizrahi, Nahid Farhady Ghalaty, Thomas Jarvinen, Ashwin S Ravi, Peter Brune, Fanyu Kong, Dave Anderson, George Lee, Arie Meir, et al. 2023. EHR-Safe: generating high-fidelity and privacy-preserving synthetic electronic health records. *NPJ digital medicine* 6, 1 (2023), 141.

[278] Jinsung Yoon, William R Zame, and Mihaela Van Der Schaar. 2018. Estimating missing data in temporal data streams using multi-directional recurrent neural networks. *IEEE Transactions on Biomedical Engineering* 66, 5 (2018), 1477–1490.

[279] Mina Youssef and Caterina Scoglio. 2011. An individual-based approach to SIR epidemics in contact networks. *Journal of theoretical biology* 283, 1 (2011), 136–144.

[280] Han Yu, Peikun Guo, and Akane Sano. 2023. Zero-shot ECG diagnosis with large language models and retrieval-augmented generation. In *Machine learning for health (ML4H)*. PMLR, 650–663.

[281] Han Yu, Peikun Guo, and Akane Sano. 2024. ECG Semantic Integrator (ESI): A Foundation ECG Model Pretrained with LLM-Enhanced Cardiological Text. *Transactions on Machine Learning Research* (2024).

[282] Daoze Zhang, Zhizhang Yuan, Yang Yang, Junru Chen, Jingjing Wang, and Yafeng Li. 2023. Brant: Foundation model for intracranial neural signal. *Advances in Neural Information Processing Systems* 36 (2023), 26304–26321.

[283] Jiawen Zhang, Shun Zheng, Wei Cao, Jiang Bian, and Jia Li. 2023. 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*. 3273–3285.

[284] Weijia Zhang, Chenlong Yin, Hao Liu, Xiaofang Zhou, and Hui Xiong. 2024. Irregular Multivariate Time Series Forecasting: A Transformable Patching Graph Neural Networks Approach. In *Forty-first International Conference on Machine Learning*.

[285] Xiyuan Zhang, Boran Han, Haoyang Fang, Abdul Fatir Ansari, Shuai Zhang, Danielle C Maddix, Cuixiong Hu, Andrew Gordon Wilson, Michael W Mahoney, Hao Wang, et al. 2025. When Does Multimodality Lead to Better Time Series Forecasting? *arXiv preprint arXiv:2506.21611* (2025).

[286] Xi Zhang, Yuan Pu, Yuki Kawamura, Andrew Loza, Yoshua Bengio, Dennis L Shung, and Alexander Tong. 2024. Trajectory flow matching with applications to clinical time series modelling. In *Advances in Neural Information Processing Systems*, Vol. 37. 107198–107224.

[287] Xinmeng Zhang, Chao Yan, Yuyang Yang, Zhuohang Li, Yubo Feng, Bradley A Malin, and You Chen. 2025. Optimizing large language models for discharge prediction: best practices in leveraging electronic health record audit logs. In *AMIA Annual Symposium Proceedings*, Vol. 2024. 1323.

[288] Xiang Zhang, Marko Zeman, Theodoros Tsiligkaridis, and Marinka Zitnik. 2022. Graph-Guided Network for Irregularly Sampled Multivariate Time Series. In *International Conference on Learning Representations*.

[289] Yuanyuan Zhang, Runwei Guan, Lingxiao Li, Rui Yang, Yutao Yue, and Eng Gee Lim. 2025. radarODE: An ODE-embedded deep learning model for contactless ECG reconstruction from millimeter-wave radar. *IEEE Transactions on Mobile Computing* (2025).

[290] Yunjia Zhang, Zhixiong Hu, and Mei Li. 2026. Modeling Localized PPG for Blood Pressure Forecasting With MoE and Quantile Regression. *IEEE Signal Processing Letters* 33 (2026), 531–535.

[291] Zichi Zhang, Phi Hung Nguyen, Ngoc Phu Doan, Viet-Hung Tran, Xuan Hoang Nguyen, Hui Wang, Hans Vandierendonck, and Son Thai Mai. 2026. HierarNet: Independent Interactive Hierarchical Disease Outbreak Forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 40. 39628–39636.

[292] Zhilin Zhang, Zhouyue Pi, and Benyuan Liu. 2014. TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. *IEEE Transactions on biomedical engineering* 62, 2 (2014), 522–531.

[293] Jianwei Zheng, Jianming Zhang, Sidy Danioko, Hai Yao, Hangyuan Guo, and Cyril Rakovski. 2020. A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. *Scientific data* 7, 1 (2020), 48.

[294] Hao Zhou, Simon A Lee, Cyrus Tanade, Keum San Chun, Juhyeon Lee, Migyeong Gwak, Megha Thukral, Justin Sung, Eugene Hwang, Mehrab Bin Morshed, et al. 2026. Physiology-Aware Masked Cross-Modal Reconstruction for Biosignal Representation Learning. *arXiv preprint arXiv:2605.00973* (2026).

[295] Rushuang Zhou, Lei Lu, Zijun Liu, Ting Xiang, Zhen Liang, David A Clifton, Yining Dong, and Yuan-Ting Zhang. 2023. Semi-supervised learning for multi-label cardiovascular diseases prediction: a multi-dataset study. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46, 5 (2023), 3305–3320.

[296] Ya Zhou, Xiaolin Diao, Yanni Huo, Yang Liu, Xiaohan Fan, and Wei Zhao. 2024. Masked transformer for electrocardiogram classification. *arXiv preprint arXiv:2309.07136* (2024).

[297] Yuchen Zhou, Jiamin Wu, Zichen Ren, Zhouheng Yao, Weiheng Lu, Kunyu Peng, Qihao Zheng, Chunfeng Song, Wanli Ouyang, and Chao Gou. 2025. CSBrain: A Cross-scale Spatiotemporal Brain Foundation Model for EEG Decoding. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

[298] Mingcheng Zhu, Zhiyao Luo, Yu Liu, and Tingting Zhu. 2025. Medtpe: Compressing long ehr sequence for llm-based clinical prediction with token-pair encoding. In *Proceedings of the 11th Mining and Learning from Time Series Workshop@ KDD*, Vol. 2025.

[299] Brendan P Zietsch, Jonathan L Hansen, Narelle K Hansell, Gina M Geffen, Nicholas G Martin, and Margaret J Wright. 2007. Common and specific genetic influences on EEG power bands delta, theta, alpha, and beta. *Biological psychology* 75, 2 (2007), 154–164.