

# Analyzing Patient Trajectories with Artificial Intelligence

**Review Article****Author(s):**

Allam, Ahmed; Feuerriegel, Stefan; Rebhan, Michael; Krauthammer, Michael

**Publication date:**

2021-12

**Permanent link:**

<https://doi.org/10.3929/ethz-b-000525110>

**Rights / license:**

[Creative Commons Attribution 4.0 International](#)

**Originally published in:**

Journal of Medical Internet Research 23(12), <https://doi.org/10.2196/29812>

**Funding acknowledgement:**

186932 - Data-driven health management (SNF)

Viewpoint

# Analyzing Patient Trajectories With Artificial Intelligence

Ahmed Allam<sup>1,2</sup>, PhD; Stefan Feuerriegel<sup>3,4,5</sup>, PhD; Michael Rebhan<sup>3</sup>, PhD; Michael Krauthammer<sup>1,2,6</sup>, MD, PhD

<sup>1</sup>Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland

<sup>2</sup>Biomedical Informatics, University Hospital of Zurich, Zurich, Switzerland

<sup>3</sup>Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland

<sup>4</sup>ETH Artificial Intelligence Center, ETH Zurich, Zurich, Switzerland

<sup>5</sup>Ludwig Maximilian University of Munich, Munich, Germany

<sup>6</sup>Yale Center for Medical Informatics, Yale University School of Medicine, New Haven, CT, United States

**Corresponding Author:**

Stefan Feuerriegel, PhD

Ludwig Maximilian University of Munich

Geschwister-Scholl-Platz 1

Munich, 80539

Germany

Phone: 49 8921806790

Email: [feuerriegel@lmu.de](mailto:feuerriegel@lmu.de)

## Abstract

In digital medicine, patient data typically record health events over time (eg, through electronic health records, wearables, or other sensing technologies) and thus form unique patient trajectories. Patient trajectories are highly predictive of the future course of diseases and therefore facilitate effective care. However, digital medicine often uses only limited patient data, consisting of health events from only a single or small number of time points while ignoring additional information encoded in patient trajectories. To analyze such rich longitudinal data, new artificial intelligence (AI) solutions are needed. In this paper, we provide an overview of the recent efforts to develop trajectory-aware AI solutions and provide suggestions for future directions. Specifically, we examine the implications for developing disease models from patient trajectories along the typical workflow in AI: problem definition, data processing, modeling, evaluation, and interpretation. We conclude with a discussion of how such AI solutions will allow the field to build robust models for personalized risk scoring, subtyping, and disease pathway discovery.

(*J Med Internet Res* 2021;23(12):e29812) doi: [10.2196/29812](https://doi.org/10.2196/29812)

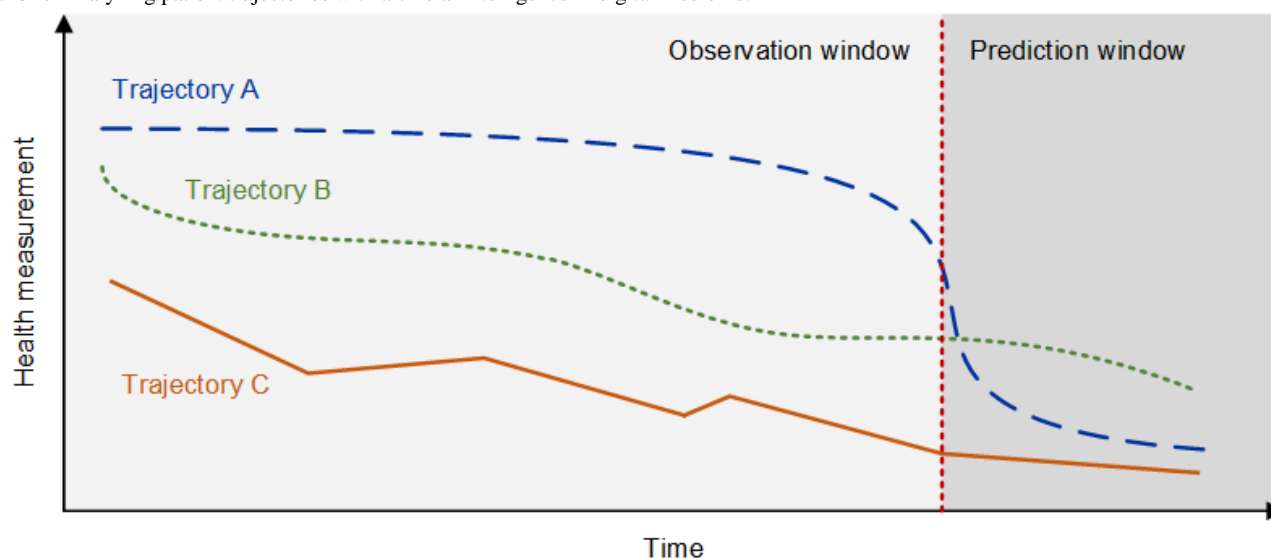
**KEYWORDS**

patient trajectories; longitudinal data; digital medicine; artificial intelligence; machine learning

## Introduction

Digital medicine facilitates broad access to large volumes of patient data, typically through recordings of health events over time. For example, electronic health records store the history of a patient's diagnoses, medications, laboratory values, and treatment plans [1-3]. Wearables collect granular sensor measurements of various neurophysiological body functions over time [4-6]. Intensive care units (ICUs) monitor disease progression via continuous physiological measurements (eg, electrocardiograms) [7-10]. As a result, patient data in digital medicine are regularly of longitudinal form (ie, consisting of health events from multiple time points) and thus form *patient trajectories*.

Analyzing patient trajectories provides opportunities for more effective care in digital medicine [2,7,11]. Patient trajectories encode rich information on the history of health states that are also predictive of the future course of a disease (eg, individualized differences in disease progression or responsiveness to medications) [9,10,12]. As such, it is possible to construct patient trajectories that capture the entire disease course and characterize the many possible disease progression patterns, such as recurrent, stable, or rapidly deteriorating disease states (Figure 1). Hence, modeling the patient trajectories allows one to build robust models of diseases that capture disease dynamics seen in patient trajectories. Here, we replace disease models with data from only a single or a small number of time points by disease models that account for the longitudinal nature of patient trajectories, thus offering vast potential for digital medicine.

**Figure 1.** Analyzing patient trajectories with artificial intelligence in digital medicine.

Several studies have previously introduced artificial intelligence (AI) in medicine for practitioners [13,14]. Some studies review potential medical applications that could benefit from AI [15,16], whereas others review specific methods (eg, deep learning [17,18]) or specific data types (eg, electronic health records [18] and medical images [19]). Some studies suggest reporting guidelines [20,21] or discuss the integration of AI into medical practice [22]. Our research contributes to the literature by discussing AI-powered digital medicine based on patient trajectories.

Patient trajectories refer to time-resolved representations of patient health events across multiple time points (eg, hospitalization or treatment events, sensor measurements from wearables, and physiological measurements from ICUs). AI-powered analysis of patient trajectories allows for an assessment of the heterogeneity of patient disease courses. In Figure 1, trajectory A predicts a sudden but sharp decline in a health state, whereas trajectories B and C depict 2 types of progressively declining disease states. Analyzing trajectory-like representations of patient data can generate new insights for better care in digital medicine.

To unlock the value of patient trajectories for digital medicine, there is a need for new AI solutions that can deal with time-resolved sequential data consisting of multiple health events. Although many models from the area of AI have become standard in digital medicine (eg, deep learning [18]), a naïve application of such models might not be effective when modeling the longitudinal nature of patient trajectories. Instead, this requires customized approaches. For example, in a study by Alaa et al [23], a direct application of deep learning has been found to be outperformed in terms of both predictive accuracy and interpretability when one instead uses a carefully engineered sequential model (ie, referred to as Hawkes process) that treats the time between medical events as informative for the course of the disease. On the basis of this background, we discuss challenges and solutions for AI that are unique to analyzing patient trajectories. Specifically, we examine the implications for developing disease models from the patient trajectories along

the typical workflow in AI: (1) problem definition, (2) data processing, (3) modeling, (4) evaluation, and (5) interpretation, as detailed in the following section.

## Applying AI to Patient Trajectories

### Problem Definition

Applying AI to patient trajectories is relevant for different objectives in digital medicine (Table 1). One objective is *risk scoring* [3,23,24], where patient trajectories are leveraged to predict patient outcomes. Here, the rationale is that the predictions based on health measurements from patient trajectories with multiple time points have greater predictive power than the predictions from those with a single or a few time points. For instance, risk scoring in ICUs becomes more accurate when traditional scores (eg, Acute Physiology and Chronic Health Evaluation II and Simplified Acute Physiology Score) are replaced with AI-based predictions incorporating data from patient trajectories [9,10,25]. Similarly, for cardiovascular diseases, existing risk scores (eg, the Framingham risk score that predicts the 10-year risk of developing coronary heart disease) become more accurate when replaced by AI solutions that work with longitudinal patient data [12]. These examples show that the underlying patient trajectory provides rich, granular insights into the disease dynamics that can then be captured by AI solutions for trajectory analysis. Therefore, additional patient information, such as past medications, comorbidities, or other risk factors, can be considered. For instance, in the context of cardiovascular diseases, it might be informative for risk scoring to analyze the patient's past journey, which comprises whether the patients have been prescribed nicotine replacements and *when* (eg, only recently or several years ago). Different patient outcomes including mortality, hospital readmission, hospital length of stay, disease onset, disease severity, or adverse drug reactions can be of interest in risk scoring. The risk score can then inform treatment planning (or in general, assess the patients' needs). In addition, AI solutions can further generate insights for defining (early) disease states.

**Table 1.** Overview of different objectives in artificial intelligence–based trajectory analysis.

Objective	Description	Examples	Selected references
Risk scoring	The objective is to estimate the likelihood of future health outcomes (eg, mortality, readmission, and adverse drug reactions)	<ul style="list-style-type: none"> <li>Predict the 10-year risk of developing coronary heart disease for patients as in the Framingham risk score</li> <li>Predict the need for an intensive care unit in an emergency ward through measurements from wearables</li> </ul>	[3,17,23,26-32]
Subtyping	The objective is to cluster the patient cohort into different disease dynamics (ie, subtyping) while accounting for the longitudinal form of patient trajectories	<ul style="list-style-type: none"> <li>Cluster disease progressions into “recurrent course” and “progressive decline”</li> </ul>	[26]
Pathway discovery	The objective is to detect clinically meaningful subpatterns in patient trajectories	<ul style="list-style-type: none"> <li>Identify frequent patterns in patient trajectories that are indicative of disease onset</li> </ul>	[1,33,34]

Here, we see several paths for risk scoring based on patient trajectories. First, to ensure high accuracy, AI-based risk scores must be tailored to each patient outcome while considering the desired forecast horizon and the patient cohort. So far, there are several studies that showcase the successful use of AI-based risk scores [35,36]; however, there is a need to develop other risk scores, especially for the settings in which AI-based risk scores are scarce or not yet available (eg, predicting the transition from prediabetes to diabetes or predicting specific adverse reactions to medication). Second, the AI-based risk scores will need to be integrated more extensively in clinical practice. Third, the risk scores should be extensively combined with approaches for explainability or interpretability, which allow the derivation of clinically relevant insights from patient trajectory data (eg, which information in a patient trajectory is a risk factor). Finally, if one includes data on treatments in the risk scoring model, one may infer the expected individualized treatment effect and eventually guide the treatment selection [37-42]. Here, we see further potential to transition from a purely predictive approach (ie, what is the expected risk level) to a prescriptive approach (ie, what treatment do we expect to reach a desired patient outcome). However, many AI solutions for estimating individualized treatment effects from patient trajectories have recently emerged [37-42] without being tailored to specific disease settings and patient cohorts. Therefore, further research at the interface to digital medicine should be a priority that will eventually yield effective and robust implementations for clinical practice.

Another objective of AI in digital medicine is *subtyping*, where AI can understand the heterogeneity observed in patient trajectories and identify the corresponding digital markers. A simple approach is to cluster the different patient trajectories (ie, subtyping) to match patients with similar disease dynamics, clinical pathways, or care patterns [26]. As a practical benefit, subtyping can support clinical tasks related to cohort building and, for instance, can provide patient stratification (eg, to define a cluster of patient trajectories that serves as an inclusion criterion for a clinical trial). However, subtyping requires a suitable notion of patient similarity, which can be challenging to define mathematically because of the longitudinal form of patient trajectories. Thus, it is crucial not only to cluster risk factors at baseline but also to find mathematical approaches that account for the temporal nature of the trajectories (ie, time-series

clustering). This allows clinical practice to identify subgroups or subtypes based on the underlying disease dynamics (eg, to distinguish subgroups with a recurrent course vs a progressive decline). We expect an added value from comparing different subtyping approaches in terms of their relative strengths (eg, generated insights) for future research. On this basis, digital medicine could develop a principled procedure for defining *patient trajectory similarity* in the context of subtyping.

A related objective is *pathway discovery*, where patterns in patient trajectories should be detected [1,33]. For instance, 1 application analyzes time series with laboratory measurements from patients with hepatitis B and C to discover frequent patterns indicative of liver damage [34]. This application can help to understand the underlying course of diseases and identify short- and long-term patterns (ie, motifs) in patient trajectories.

Depending on the objective and explicit assumptions, implications arise regarding the AI workflow and thus highlight the importance of selecting an appropriate modeling strategy. Additional details are provided in the following sections.

## Data Processing

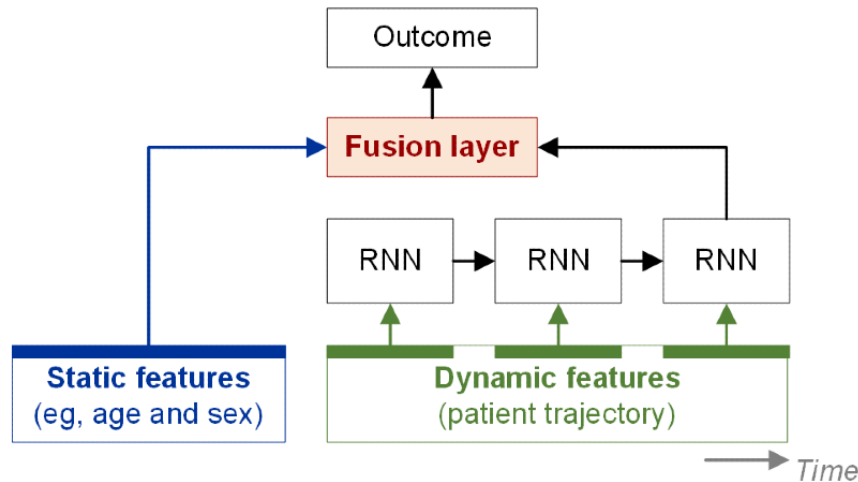
A fundamental question is concerned with data collection as this defines how time is encoded in the data. The example of nicotine replacement suggests that we should know whether such a medical event was recent or several years ago. This illustrates where we are now: when we have longitudinal data, it is often a matter of whether a medical event happened *recently* or *a while back*. Thus, the underlying time (or the underlying time lag) must be carefully considered to capture the longitudinal dimension of patient trajectories correctly. This is currently a challenge when considering the survey designs (eg, for risk scores). For example, one survey may ask a patient whether an event occurred last month or earlier, whereas another survey may ask to consider events that occurred within the last 12 months or earlier. As such, the meaning of *recent* may be inconsistent across survey designs. As a way forward, it will be necessary to develop a more consistent understanding, ideally involving data collection that considers precise time stamps (eg, by leveraging electronic health records).

The AI-based trajectory analysis often combines data from patient trajectories and baseline variables that describe risk factors at the patient level (eg, sociodemographic, genomic data,

or multimodal data). To combine sequential and static baseline data, tailored AI solutions will need to be developed [27]. Figure 2 shows an example of this approach. The literature shows larger variability regarding the modeling approaches; hence, further evaluations are needed to inform an effective approach.

Figure 2 shows an AI-based trajectory analysis in which a fusion layer combines the static (eg, age or sex) and dynamic features (eg, health measurements over time). The dynamic features have a longitudinal form and are processed by a sequential model (here, a recurrent neural network [RNN]).

**Figure 2.** Example of artificial intelligence–based trajectory analysis. RNN: recurrent neural network.



Data from patient trajectories are often complex and high-dimensional, which presents difficulties to humans and AI solutions for making accurate inferences. As a remedy, one can apply approaches that map patient trajectories onto a lower-dimensional representation that is eventually more meaningful. A simple analogy from medical practice is when one simplifies the age in years of a patient into a binary yes or no flag indicating whether a patient is above a critical age threshold. Here, we see a particular value for representation learning (eg, embeddings [31]) tailored to the unique characteristics of longitudinal health time series from clinical practice. Such AI solutions must be effective in dealing with the high-dimensional nature of medical data (clinical, genetic, social data, etc), avoid overfitting, and overcome the *curse of dimensionality* in the analysis. Mathematically, the idea of learning lower-dimensional representations is linked to manifold learning, for which embeddings are a special case [43].

Another challenge that further limits the use of AI-based trajectory analysis in medicine is data access and sharing. The current system is characterized by having data locked in *silos* where each hospital or health care institution limits access to their data and requires *lots of bureaucratic* work from researchers before allowing them to study and analyze the data. However, many initiatives are circumventing this status quo, such as the Observational Health Data Sciences and Informatics program, an international network of researchers aiming to provide reusable, collaborative, and reliable open-source solutions for large-scale health analytics [44]. Notably, there is interest in compiling extensive observational studies combining the electronic health record data from diverse health care organizations using standards to (1) design meaningful randomized controlled trials, (2) test clinical hypotheses on observational data, and (3) gain a better understanding of population characteristics, facilitated through framework efforts such as the Observational Health Data Sciences and Informatics program [44].

Furthermore, in the last couple of years, there has been an increasing number of studies focused on federated learning [45-47] that allows for AI algorithms to operate on decentralized data sets/systems in a privacy-preserving manner. In federated learning, the underlying algorithms (eg, FedAvg [48] and FedProx [49]) use the data stored in silos at different local sites for iteratively training a *global/central* model from a set of *local* models trained separately at each local site to perform prediction and classification tasks [48,50]. At the interface to health care, more research is being conducted that uses federated learning [45,51] for tasks such as clinical note phenotyping (ie, clinical natural language processing [52]) or predicting patient mortality in ICUs [53,54]. Here, we particularly point to the recent attempts to develop such approaches for patient trajectories. Examples include clustering patients through community-based federated machine learning for in-hospital mortality and length of stay prediction [55] or privacy-preserving patient similarity learning [56]. Federated learning may be further supported by secure hardware implementations, often with little computational overhead (eg, referred to as trusted execution environments [47]).

Moreover, there is more research on the privacy-preserving aspect of the technology, such as the differential privacy framework [57] (ie, applied to the model parameters), homomorphic encryption [58], and data anonymization techniques offering a *defensive level of privacy* as required by the General Data Protection Regulation and the Health Insurance Portability and Accountability Act [59]. Although these provide valuable tools for developers, we foresee more research that tailors them to the context of patient trajectories (eg, by offering sequential models for longitudinal data). More importantly, the availability of software packages [60,61] that allow both simulation of federated learning scenarios and their deployment in real clinical settings will accelerate the adoption of federated or distributed learning approaches and open a wide array of research exploration and experimentation possibilities. This



will eventually yield longitudinal trajectory analyses that span patient journeys across multiple hospitals or health care institutions.

## Modeling

Fundamentally, analyzing patient trajectories requires AI solutions that can effectively handle sequential data structures that can vary in length (ie, from a few time points to multiple seconds, minutes, days, months, and year time windows). Hence, AI-based trajectory analysis must carefully adapt to the sequential structures by choosing appropriate modeling approaches.

In *risk scoring*, predictions from patient trajectories are often based on neural networks that are tailored to sequential data structures. These include tailored RNNs [3,17,27-29] owing to their strength in modeling long-term dependencies. One example of RNNs is the long short-term memory networks that iteratively process a time series with physiological measurements and aim to learn a lower-dimensional representation of the complete time series, regardless of its length, in their internal code layer. We can then use this lower-dimensional representation to predict patient outcomes from the patient's trajectory. Gated recurrent units proceed analogously but have a more parsimonious structure that, in medical applications, may help in obtaining robust predictions (eg, with a lower risk of overfitting for small-sized data sets) [8,27,62]. Recently, digital medicine has also processed the patient trajectories through transformer networks [30]. Transformer networks involve attention layers that learn to weigh different parts in a patient trajectory differently while optimizing for the outcome prediction and may therefore outperform other RNNs. In our view, a particular benefit of transformer networks is that the developers from digital medicine can train them in using semisupervised learning. One can use a set of patient trajectories *without* observing the patient outcomes to learn an abstract representation (ie, via unsupervised pretraining). Subsequently, one can customize the transformer network to predict a specific patient outcome (ie, via supervised fine-tuning). Semisupervised learning often facilitates more efficient learning when the number of available observations with patient outcomes is comparatively low.

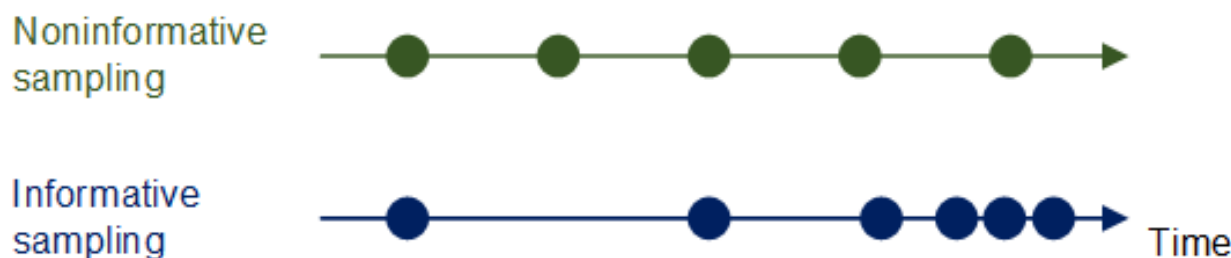
In risk scoring, other common prediction approaches are probabilistic models (eg, Markov models, point processes, and Gaussian processes) [23,32]. Here, we see several benefits for patient trajectory analyses in clinical settings. Probabilistic models often have a more parsimonious structure than the out-of-the-box neural networks, which facilitates efficient learning and reduces the risk of overfitting in data-scarce settings. In addition, such a parsimonious structure can facilitate interpretation by clinical practitioners. Probabilistic models can be naturally extended by latent structures (eg, hidden Markov models [63-67]), where latent states capture different trajectory

phases and further improve interpretability. For instance, in the context of alcoholism treatment, patient trajectories have been modeled to undergo phases of *abstinence*, *moderate drinking*, and *heavy drinking*, each of which is captured by a separate latent state. In our view, such a latent structure represents a natural way to describe the different patterns in patient trajectories (eg, acute vs stable phases) and, more importantly, relates model characteristics to established clinical terminology. In the future, we expect hybrid models that combine the benefits of probabilistic modeling and neural learning (eg, deep Markov models [25]). The former benefits from theory-informed, interpretable structures that account for different disease states, whereas the latter are particularly effective for long-term dependencies.

Across these risk models, it is further essential to consider the timing of the health measurement. Trajectories may consist of health recordings in equally spaced time intervals (eg, uniformly sampled every minute in ICUs or yearly intervals in patient registries). Often, they contain irregular time intervals, reflecting nonuniform and patient-specific interactions with the health care system. As a result, the sampling might be informative of the disease state (Figure 3). For instance, health professionals record more health measurements during deterioration in the patient's health state. Therefore, AI solutions can use shorter time intervals to predict future decline in the trajectory. If the sampling is informative, we encourage researchers to develop models that consider the time intervals between the health measurements. For instance, 1 class of such models is point processes (eg, Hawkes processes). Here, a shorter time interval between health measurements makes further health measurements more likely and influences the expected risk score [23].

Figure 3 shows individual health recordings (eg, medical events, hospital visits, and laboratory data) in the form of dots. *Top*: an example patient trajectory where all medical events are equally spaced and thus there is a uniform time interval between the events. Here, the timing of the events is not informative of the current disease state. *Bottom*: an example patient trajectory that indicates a gradual decline in the disease state. Owing to this, additional health recordings are collected with higher frequency, which are thus informative about the disease state.

For objectives beyond risk scoring, we need other modeling approaches. When using risk scoring for *prescriptive purposes* (eg, treatment planning or dose finding), we encourage broader adoption of modeling strategies designed for decision making (eg, causal machine learning [37,68,69], Markov decision processes [70,71], dynamic treatment regimens [72,73], and policy learning [74]). There is a growing traction to extend many of these modeling strategies to handle longitudinal data, where health practitioners make treatment decisions over time.

**Figure 3.** Difference between noninformative and informative sampling.

For *subtyping*, one typically draws upon time-series clustering [26]. For this objective, the definition of a similarity metric is key [75,76]. One option is to set explicit rules for establishing patient trajectory similarities, for example, according to a specific condition such as heart disease [77] or a combination of clinical or phenotypic features [78]. Domain experts may be familiar with such approaches, but their definition may not perfectly integrate with AI algorithms and may thus require customization. Hence, an alternative is to view patient trajectory analysis from a methodological, *data-driven* angle. For example, when modeling the underlying temporal dynamics for performing data-driven clustering of longitudinal data, we can loosely group the approaches into (1) model-free approaches and (2) model-based approaches. In the model-free approaches, a similarity metric on sequential data is defined and then serves as input to conventional clustering algorithms. On the other hand, in the model-based approaches [79], a representation of the patient trajectory is learned, and the model parameters are then used for clustering (eg, mixture hidden Markov models). For future research, we find model-based approaches intriguing, as they no longer focus on raw observations but cluster the underlying disease dynamics (and, as such, can account for different latent disease states).

For *pathway discovery*, several descriptive approaches have emerged that allow for the discovery of subpatterns (ie, motifs) and common patient trajectories [76]. For example, Beck et al [33] analyzed disease pathways leading to septicemia in 110,000 patients. The study revealed prototypical pathways starting from 3 initial states (alcohol abuse, diabetes, and anemia) and established the trajectory-specific probability of sepsis mortality. This and similar studies reveal great potential to further our understanding of disease etiology and the possible means of changing disease trajectories [80,81]. Similarly, Oh et al [82] constructed patient trajectories consisting of specific health events (hyperlipidemia, hypertension, and impaired fasting glucose) and evaluated the probabilities of such trajectories (and their permutation) in increasing or decreasing the log odds of developing type 2 diabetes mellitus. Zhang and Padman [83] identified the most probable clinical trajectories from patients with chronic kidney disease by first grouping the patients and then fitting a first-order hidden Markov model to infer the most probable clinical pathways given sequences of multiple laboratory test observations and other patient characteristics. Huang et al [84] proposed a probabilistic model (based on latent Dirichlet allocation) to identify clinical pathway patterns from the event logs for patients with unstable angina and cancer. Dabek and Caban [85] offered another perspective by analyzing trajectories using automata (ie, deterministic and

nondeterministic finite state automata) and using a grammar induction algorithm to identify common trajectories in neurology. Further approaches for pathway discovery are based on association rule mining [86] and functional dependencies mining [87].

Related to these objectives are models that adopt a structural lens to examine the causal mechanisms [88]. This would allow not only to understand *how* health measurements change over time but also *why*. The underlying AI algorithms are still under active research (eg, causal structure learning and neural causal discovery [89]). Here, it will be a rewarding direction for the future to develop more approaches that are tailored to longitudinal data.

### Evaluation

Evaluations through randomized controlled trials are needed to confirm the effectiveness of AI-based analysis of patient trajectories in clinical practice. Recently, there have been such trials for traditional AI solutions that rely on snapshot data from a single or few time points [90]. However, similar trials for patient trajectory analysis are rare [90]. We expect significant value in conducting such trials and foresee challenges owing to the unique characteristics of patient trajectories. Foremost, evaluations through rigorous randomized controlled trials are a prerequisite to building trust in clinical practice, thereby expediting further AI-based trajectory analysis. However, evaluating such an AI solution might be a multiyear undertaking depending on the time window of the patient's trajectory. Thus, it is also essential to recognize that the evaluations are likely to involve a 2-step procedure. In the first step, trajectory data are collected to train the AI solution. In the second step, the previously trained AI solution is then deployed to analyze how the AI solution generalizes to new patient trajectories. When conducting such trials, it is crucial to acknowledge that the patient trajectories capture data from multiple time points and might thus be subject to an inherent domain shift (ie, where data distributions change over time, that is, over the patient journey) [91]. Such domain shifts might affect the performance of AI solutions [92], especially when the patient trajectories span a long time horizon. For example, AI-based predictions have been recently applied to patients with COVID-19 to compare the predicted health trajectory with the observed trajectory in a prospective study, finding that the performance of some risk scores decreased over time [93]. One reason was because of temporal domain shifts over time [94] as medical professionals learned about the emerging infectious disease and adapted their clinical routines over time, thus yielding different and, in particular, better outcomes than in the data used for training.

Furthermore, there is a need to ensure reproducibility of the AI-based analyses. Here, we consider 3 priorities. First, there is a need to develop a framework. For instance, the existing AI frameworks (such as scikit learn in Python) are designed for modeling static data. Conversely, more effort is necessary to define standardized building blocks for longitudinal data to effectively model the patient trajectories. Moreover, such frameworks should also involve tools for automation so that the disease models can be trained in a semiautomated manner (subsumed under the term *automated machine learning* [AutoML]). Although AutoML has become widespread for static data sets, only a few libraries are designed for time-series AutoML [95]. Here, we see enormous potential for future research at the interface to digital medicine. On the basis of our own experience, we expect such frameworks to play a critical role in achieving scalable and increased adoption of AI-based patient trajectory analysis in clinical settings. Second, future research should carefully assess and, if needed, revise the best practice guidelines for AI in medicine [96], so that they consider the longitudinal form of patient trajectories. For example, the reporting guidelines should involve information on whether the period between health measurements is informative (however, such information is not part of existing reporting guidelines for static patient data). Other examples could be the choice of the model (eg, whether latent dynamics were considered and why or which alternative architectures of recurrent neural networks were tested and eventually discarded), how the timing of the health care event was collected (eg, whether a time stamp was retrieved from a medical health record or whether this was a survey question referring to the *last 12 months* and the resulting uncertainty about the correct time), or the rationale behind how patient similarity was measured in subtyping. Finally, more data sets with patient trajectories should be made publicly available for benchmarking. Although data access is a common issue for AI research in medicine in general, the challenges are exacerbated in the context of patient journeys, where it is common to merge the health measurements from different sources (eg, from other health registries). So far, only a few longitudinal data sets are public (as compared with static data sets with patient data) [76]. Notable exceptions are large initiatives, such as the Healthcare Cost and Utilization Project, that offer longitudinal data sets for evaluating trajectory-based AI approaches [97].

### Interpretations

To generate insights for clinical practice, it is often necessary that AI solutions overcome their black-box nature [98]. Here, we see enormous potential for new AI solutions that adhere to the needs of clinical practice with the objective of knowledge discovery.

One approach is explainability, which aims to understand how a model arrives at a particular outcome [10,99]. However, AI explainability is typically developed in the context of static data and therefore, meaningful time-varying patterns from the course of a disease might not be revealed. For instance, SHAP values [99] inform which health measurements are used by the AI model and what values indicate risk. However, SHAP values

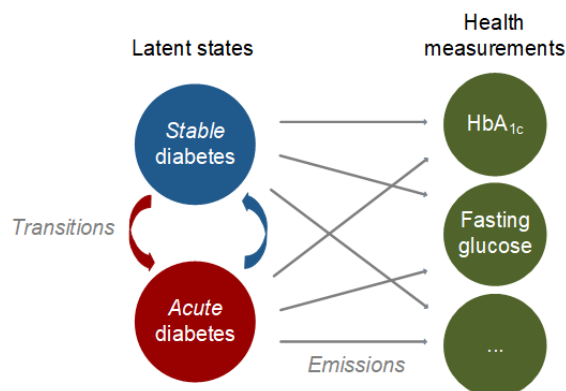
cannot directly interpret the *dynamics* in health measurements (eg, whether there is an increase or decrease or large variability in health measurements, which would be needed to characterize changes in disease states over time). As a road map for research in digital medicine, we require techniques that interpret the temporal dynamics of disease progression, thereby being closely aligned with the demands of clinical practice. Here, digital medicine might find inspiration in other disciplines, for instance, financial technical analysis [100], where a systematic set of short-term movements of stock prices is used for interpretation. Similar patterns in patient trajectories could be inferred by researchers via short-term patterns (ie, trajectory markers) that characterize a disease course or a combination of short- and long-term motifs that help identify distinct disease states.

Another approach to generating insights is via interpretability, which builds upon inferences where the underlying logic is transparent. Interpretability often requires tailored modeling approaches. For static data, this is usually achieved through (penalized) linear regression or decision trees, whereas interpretability for longitudinal data is typically achieved through parsimonious models. Different strategies exist to obtain parsimonious models. For neural networks, there are techniques that tweak a neural network to provide a sparse one with similar performance (eg, enforcing feature sparsity through architecture design, modifying the objective function and the weight updating scheme [101], post hoc via reservoir computing or pruning, or a priori via cognitive networks [102]). Alternatively, one can draw upon structural formalizations (eg, dynamic fuzzy cognitive maps [78] to simulate patient trajectories) and probabilistic models (eg, Markov models, hidden Markov models, and Hawkes processes [23,25,63-67]).

Out of these modeling approaches, we expect hidden Markov models to be beneficial for interpretability, especially for risk scoring. The reason being hidden Markov models use latent variables to capture different disease phases in patient trajectories. These latent variables often have clinically relevant meanings and can thus be mapped onto existing clinical terminology (Figure 4). For instance, in diabetes mellitus, the latent states are characterized as *acute* and *stable* disease states [64] and thus are of clinical meaning. In addition, recent evidence suggests that interpretable models might also improve prediction performance [23]; however, more effort is needed to explore this further. In the future, we expect to see other modeling approaches that combine the strengths of hidden Markov models (for interpretability) with neural learning (for representation learning and capturing long-term dependencies in patient trajectories).

The health measurements are observable and thus obtained via standard data collection practices. In contrast, the latent states cannot be observed directly and, instead, are recovered from the health measurements. The latent states then describe different disease states in a patient trajectory (eg, *acute* vs *stable* disease states). During estimation, the latent states and health measurements are mathematically linked via components for both transition and emission probabilities.



**Figure 4.** Example of a hidden Markov model. HbA<sub>1c</sub>: hemoglobin A<sub>1c</sub>.

## Implications for Digital Medicine

Analyzing patient trajectories using AI has multiple benefits. Dissecting the variability of disease pathways allows research to better understand both the disease etiology and disease course, facilitating a more extensive personalization of care. For instance, it enables the identification of short-term patterns predictive of future health states, which are then used during risk scoring. Similarly, patient trajectories capture the responsiveness of patients to treatments, and by leveraging this information in patient trajectories, AI solutions can guide treatment planning. In summary, AI-based trajectory analysis promises to strengthen the existing computational approaches to prevent, detect, diagnose, and treat diseases.

To address these challenges, we see particular importance in community building and in the development of computational patient trajectory tools that lower the barrier of entry.

Community building will help to set a clear agenda and define an established terminology, bridging both practice and research in digital medicine. Here, we point to several valuable directions: (1) further effort will be needed to extend traditional clinical terminology (eg, cohort building and patient similarity) to AI-based trajectory analysis, thereby facilitating communication between the experts in AI and medicine; (2) it is essential to build communities by connecting different actors from regulation, law, data science, and medicine, as this will eventually be a prerequisite for deploying AI solutions in medical practice; and (3) such communities may promote data exchange, thus allowing for more extensive benchmarking of AI solutions. Similarly, we also suggest hosting leaderboard competitions as conducted in other fields (eg, the SemEval benchmark competitions in natural language processing). Leaderboard competitions will eventually help to identify robust AI solutions and thus to condense best practices during modeling.

## Acknowledgments

SF received funding from the Swiss National Science Foundation (grant 186932). AA and MK were funded by grant 201184 from the Swiss National Science Foundation. The opinions expressed in this manuscript are those of the authors' and do not necessarily reflect those of the Novartis Institutes for Biomedical Research.

## Conflicts of Interest

MR declares employment with the Novartis Institutes for Biomedical Research, Switzerland.

## References

- Jensen AB, Moseley PL, Oprea TI, Ellesøe SG, Eriksson R, Schmock H, et al. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. *Nat Commun* 2014 Jun 24;5(1):4022 [FREE Full text] [doi: [10.1038/ncomms5022](https://doi.org/10.1038/ncomms5022)] [Medline: [24959948](https://pubmed.ncbi.nlm.nih.gov/24959948/)]
- Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med* 2018 May 8;1(1):18 [FREE Full text] [doi: [10.1038/s41746-018-0029-1](https://doi.org/10.1038/s41746-018-0029-1)] [Medline: [31304302](https://pubmed.ncbi.nlm.nih.gov/31304302/)]
- Xiao C, Choi E, Sun J. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. *J Am Med Inform Assoc* 2018 Oct 01;25(10):1419-1428 [FREE Full text] [doi: [10.1093/jamia/ocy068](https://doi.org/10.1093/jamia/ocy068)] [Medline: [29893864](https://pubmed.ncbi.nlm.nih.gov/29893864/)]
- Woldaregay AZ, Årsand E, Botsis T, Albers D, Mamykina L, Hartvigsen G. Data-driven blood glucose pattern classification and anomalies detection: machine-learning applications in type 1 diabetes. *J Med Internet Res* 2019 May 01;21(5):e11030 [FREE Full text] [doi: [10.2196/11030](https://doi.org/10.2196/11030)] [Medline: [31042157](https://pubmed.ncbi.nlm.nih.gov/31042157/)]
- Noah B, Keller MS, Mosadeghi S, Stein L, Johl S, Delshad S, et al. Impact of remote patient monitoring on clinical outcomes: an updated meta-analysis of randomized controlled trials. *NPJ Digit Med* 2018 Jan 15;1(1):20172. [doi: [10.1038/s41746-017-0002-4](https://doi.org/10.1038/s41746-017-0002-4)]

6. Porumb M, Stranges S, Pescapè A, Pecchia L. Precision medicine and artificial intelligence: a pilot study on deep learning for hypoglycemic events detection based on ECG. *Sci Rep* 2020 Jan 13;10(1):170 [FREE Full text] [doi: [10.1038/s41598-019-56927-5](https://doi.org/10.1038/s41598-019-56927-5)] [Medline: [31932608](https://pubmed.ncbi.nlm.nih.gov/31932608/)]
7. Contreras I, Vehi J. Artificial intelligence for diabetes management and decision support: literature review. *J Med Internet Res* 2018 May 30;20(5):e10775 [FREE Full text] [doi: [10.2196/10775](https://doi.org/10.2196/10775)] [Medline: [29848472](https://pubmed.ncbi.nlm.nih.gov/29848472/)]
8. Che Z, Purushotham S, Cho K, Sontag D, Liu Y. Recurrent neural networks for multivariate time series with missing values. *Sci Rep* 2018 Apr 17;8(1):6085 [FREE Full text] [doi: [10.1038/s41598-018-24271-9](https://doi.org/10.1038/s41598-018-24271-9)] [Medline: [29666385](https://pubmed.ncbi.nlm.nih.gov/29666385/)]
9. Purushotham S, Meng C, Che Z, Liu Y. Benchmarking deep learning models on large healthcare datasets. *J Biomed Inform* 2018 Jul;83:112-134 [FREE Full text] [doi: [10.1016/j.jbi.2018.04.007](https://doi.org/10.1016/j.jbi.2018.04.007)] [Medline: [29879470](https://pubmed.ncbi.nlm.nih.gov/29879470/)]
10. Thorsen-Meyer H, Nielsen AB, Nielsen AP, Kaas-Hansen BS, Toft P, Schierbeck J, et al. Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records. *Lancet Digit Health* 2020 Apr;2(4):179-191. [doi: [10.1016/s2589-7500\(20\)30018-2](https://doi.org/10.1016/s2589-7500(20)30018-2)]
11. Harutyunyan H, Khachatrian H, Kale DC, Ver Steeg G, Galstyan A. Multitask learning and benchmarking with clinical time series data. *Sci Data* 2019 Jun 17;6(1):96 [FREE Full text] [doi: [10.1038/s41597-019-0103-9](https://doi.org/10.1038/s41597-019-0103-9)] [Medline: [31209213](https://pubmed.ncbi.nlm.nih.gov/31209213/)]
12. Zhao J, Feng Q, Wu P, Lupu RA, Wilke RA, Wells QS, et al. Learning from longitudinal data in electronic health record and genetic data to improve cardiovascular event prediction. *Sci Rep* 2019 Jan 24;9(1):717 [FREE Full text] [doi: [10.1038/s41598-018-36745-x](https://doi.org/10.1038/s41598-018-36745-x)] [Medline: [30679510](https://pubmed.ncbi.nlm.nih.gov/30679510/)]
13. Meskó B, Görög M. A short guide for medical professionals in the era of artificial intelligence. *NPJ Digit Med* 2020 Sep 24;3(1):126 [FREE Full text] [doi: [10.1038/s41746-020-00333-z](https://doi.org/10.1038/s41746-020-00333-z)] [Medline: [34556791](https://pubmed.ncbi.nlm.nih.gov/34556791/)]
14. Pencina MJ, Goldstein BA, D'Agostino RB. Prediction models — development, evaluation, and clinical application. *N Engl J Med* 2020 Apr 23;382(17):1583-1586. [doi: [10.1056/nejmp2000589](https://doi.org/10.1056/nejmp2000589)]
15. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019 Jan 7;25(1):44-56. [doi: [10.1038/s41591-018-0300-7](https://doi.org/10.1038/s41591-018-0300-7)] [Medline: [30617339](https://pubmed.ncbi.nlm.nih.gov/30617339/)]
16. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med* 2019 Apr 04;380(14):1347-1358. [doi: [10.1056/nejmra1814259](https://doi.org/10.1056/nejmra1814259)]
17. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. *Nat Med* 2019 Jan 7;25(1):24-29. [doi: [10.1038/s41591-018-0316-z](https://doi.org/10.1038/s41591-018-0316-z)] [Medline: [30617335](https://pubmed.ncbi.nlm.nih.gov/30617335/)]
18. Tomašev N, Harris N, Baur S, Mottram A, Glorot X, Rae JW, et al. Use of deep learning to develop continuous-risk models for adverse event prediction from electronic health records. *Nat Protoc* 2021 Jun 05;16(6):2765-2787. [doi: [10.1038/s41596-021-00513-5](https://doi.org/10.1038/s41596-021-00513-5)] [Medline: [33953393](https://pubmed.ncbi.nlm.nih.gov/33953393/)]
19. Esteva A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, et al. Deep learning-enabled medical computer vision. *NPJ Digit Med* 2021 Jan 08;4(1):5 [FREE Full text] [doi: [10.1038/s41746-020-00376-2](https://doi.org/10.1038/s41746-020-00376-2)] [Medline: [33420381](https://pubmed.ncbi.nlm.nih.gov/33420381/)]
20. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK, SPIRIT-AICONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. *Br Med J* 2020 Sep 09;370:m3164 [FREE Full text] [doi: [10.1136/bmj.m3164](https://doi.org/10.1136/bmj.m3164)] [Medline: [32909959](https://pubmed.ncbi.nlm.nih.gov/32909959/)]
21. Rivera SC, Liu X, Chan A, Denniston AK, Calvert MJ, SPIRIT-AICONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. *Br Med J* 2020 Sep 09;370:m3210 [FREE Full text] [doi: [10.1136/bmj.m3210](https://doi.org/10.1136/bmj.m3210)] [Medline: [32907797](https://pubmed.ncbi.nlm.nih.gov/32907797/)]
22. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med* 2019 Jan 7;25(1):30-36 [FREE Full text] [doi: [10.1038/s41591-018-0307-0](https://doi.org/10.1038/s41591-018-0307-0)] [Medline: [30617336](https://pubmed.ncbi.nlm.nih.gov/30617336/)]
23. Alaa A, Hu S, Schaar M. Learning from clinical judgments: Semi-Markov-modulated marked Hawkes processes for risk prognosis. In: Proceedings of the 34th International Conference on Machine Learning. 2017 Presented at: 34th International Conference on Machine Learning; August 6-11, 2017; Sydney, Australia p. 60-69 URL: <https://proceedings.mlr.press/v70/alaal7a.html>
24. Ghassemi M, Naumann T, Schulam P, Beam A, Chen I, Ranganath R. A review of challenges and opportunities in machine learning for health. *AMIA Jt Summits Transl Sci Proc* 2020;2020:191-200 [FREE Full text] [Medline: [32477638](https://pubmed.ncbi.nlm.nih.gov/32477638/)]
25. Oezuyurt Y, Kraus M, Hatt T, Feuerriegel S. AttDMM: an attentive deep Markov model for risk scoring in intensive care units. In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021 Presented at: KDD '21: The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining; August 14 - 18, 2021; Virtual Event Singapore p. 3452-3462. [doi: [10.1145/3447548.3467143](https://doi.org/10.1145/3447548.3467143)]
26. Schulam P, Wigley F, Saria S. Clustering longitudinal clinical marker trajectories from electronic health data: applications to phenotyping and endotype discovery. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015 Presented at: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence; January 25 - 30, 2015; Austin Texas p. 2956-2964. [doi: [10.5555/2888116.2888128](https://doi.org/10.5555/2888116.2888128)]
27. Choi E, Bahadori MT, Schuetz A, Stewart WF, Sun J. Doctor AI: predicting clinical events via recurrent neural networks. *JMLR Workshop Conf Proc* 2016 Aug;56:301-318 [FREE Full text] [Medline: [28286600](https://pubmed.ncbi.nlm.nih.gov/28286600/)]
28. Lipton Z, Kale D, Elkan C, Wetzell R. Learning to diagnose with LSTM recurrent neural networks. In: Proceedings of the International Conference on Learning Representations (ICLR). 2015 Presented at: International Conference on Learning

- Representations (ICLR); May 7-9, 2015; San Diego, CA, USA URL: [https://scholar.google.fr/citations?view\\_op=view\\_citation&hl=en&user=MN9Kfg8AAAAJ&citation\\_for\\_view=MN9Kfg8AAAAJ:Tyk-4Ss8FVUC](https://scholar.google.fr/citations?view_op=view_citation&hl=en&user=MN9Kfg8AAAAJ&citation_for_view=MN9Kfg8AAAAJ:Tyk-4Ss8FVUC)
29. Zhang D, Yin C, Hunold KM, Jiang X, Caterino JM, Zhang P. An interpretable deep-learning model for early prediction of sepsis in the emergency department. *Patterns (N Y)* 2021 Feb 12;2(2):100196 [FREE Full text] [doi: [10.1016/j.patter.2020.100196](https://doi.org/10.1016/j.patter.2020.100196)] [Medline: [33659912](https://pubmed.ncbi.nlm.nih.gov/33659912/)]
  30. Li Y, Rao S, Solares JR, Hassaine A, Ramakrishnan R, Canoy D, et al. BEHRT: Transformer for electronic health records. *Sci Rep* 2020 Apr 28;10(1):7155 [FREE Full text] [doi: [10.1038/s41598-020-62922-y](https://doi.org/10.1038/s41598-020-62922-y)] [Medline: [32346050](https://pubmed.ncbi.nlm.nih.gov/32346050/)]
  31. Choi Y, Chiu CY, Sontag D. Learning low-dimensional representations of medical concepts. *AMIA Jt Summits Transl Sci Proc* 2016;2016:41-50 [FREE Full text] [Medline: [27570647](https://pubmed.ncbi.nlm.nih.gov/27570647/)]
  32. Alaa AM, Yoon J, Hu S, van der Schaar M. Personalized risk scoring for critical care prognosis using mixtures of Gaussian processes. *IEEE Trans Biomed Eng* 2018 Jan;65(1):207-218. [doi: [10.1109/tbme.2017.2698602](https://doi.org/10.1109/tbme.2017.2698602)]
  33. Beck MK, Jensen AB, Nielsen AB, Perner A, Moseley PL, Brunak S. Diagnosis trajectories of prior multi-morbidity predict sepsis mortality. *Sci Rep* 2016 Nov 04;6(1):36624 [FREE Full text] [doi: [10.1038/srep36624](https://doi.org/10.1038/srep36624)] [Medline: [27812043](https://pubmed.ncbi.nlm.nih.gov/27812043/)]
  34. Hirano S, Tsumoto S. Mining similar temporal patterns in long time-series data and its application to medicine. In: *Proceedings of the IEEE International Conference on Data Mining*. 2002 Presented at: IEEE International Conference on Data Mining; Dec. 9-12, 2002; Maebashi City, Japan. [doi: [10.1109/icdm.2002.1183906](https://doi.org/10.1109/icdm.2002.1183906)]
  35. Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 2019 Aug 31;572(7767):116-119 [FREE Full text] [doi: [10.1038/s41586-019-1390-1](https://doi.org/10.1038/s41586-019-1390-1)] [Medline: [31367026](https://pubmed.ncbi.nlm.nih.gov/31367026/)]
  36. Ravizza S, Huschto T, Adamov A, Böhm L, Büsser A, Flöther FF, et al. Predicting the early risk of chronic kidney disease in patients with diabetes using real-world data. *Nat Med* 2019 Jan 7;25(1):57-59. [doi: [10.1038/s41591-018-0239-8](https://doi.org/10.1038/s41591-018-0239-8)] [Medline: [30617317](https://pubmed.ncbi.nlm.nih.gov/30617317/)]
  37. Schulam P, Saria S. Reliable decision support using counterfactual models. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017 Presented at: 31st International Conference on Neural Information Processing Systems; December 4 - 9, 2017; Long Beach California USA p. 1696-1706. [doi: [10.5555/3294771.3294933](https://doi.org/10.5555/3294771.3294933)]
  38. Lim B, Alaa A, van der Schaar M. Forecasting treatment responses over time using recurrent marginal structural networks. In: *Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018)*. 2018 Presented at: 32nd Conference on Neural Information Processing Systems (NeurIPS 2018); 2018; Montréal, Canada URL: <https://proceedings.neurips.cc/paper/2018/file/5666a93212e4482d99c84a639d254b67-Paper.pdf>
  39. Liu R, Yin C, Zhang P. Estimating individual treatment effects with time-varying confounders. In: *Proceedings of the IEEE International Conference on Data Mining (ICDM)*. 2020 Presented at: IEEE International Conference on Data Mining (ICDM); Nov. 17-20, 2020; Sorrento, Italy. [doi: [10.1109/icdm50108.2020.00047](https://doi.org/10.1109/icdm50108.2020.00047)]
  40. Hatt T, Feuerriegel S. Sequential deconfounding for causal inference with unobserved confounders. *arXiv*. 2021. URL: <https://arxiv.org/abs/2104.09323> [accessed 2021-11-27]
  41. Bica I, Alaa A, Jordon J, van der Schaar M. Estimating counterfactual treatment outcomes over time through adversarially balanced representations. In: *Proceedings of the International Conference on Learning Representations (ICLR)*. 2019 Presented at: International Conference on Learning Representations (ICLR); May 6-9, 2019; New Orleans, LA, USA URL: [https://scholar.google.co.uk/citations?view\\_op=view\\_citation&hl=en&user=mnU3HpcAAAAJ&citation\\_for\\_view=mnU3HpcAAAAJ:9yKSN-GCB0IC](https://scholar.google.co.uk/citations?view_op=view_citation&hl=en&user=mnU3HpcAAAAJ&citation_for_view=mnU3HpcAAAAJ:9yKSN-GCB0IC)
  42. Kuzmanovic M, Hatt T, Feuerriegel S. Deconfounding Temporal Autoencoder: estimating treatment effects over time using noisy proxies. In: *Proceedings of the Machine Learning for Health (ML4H) Workshop 2021*. 2021 Presented at: Machine Learning for Health (ML4H) Workshop 2021; December 4, 2021; Online (Upcoming) URL: <https://www.research-collection.ethz.ch/handle/20.500.11850/512184>
  43. Izenman AJ. Introduction to manifold learning. *WIREs Comp Stat* 2012 Jul 16;4(5):439-446. [doi: [10.1002/wics.1222](https://doi.org/10.1002/wics.1222)]
  44. Hripcsak G, Ryan PB, Duke JD, Shah NH, Park RW, Huser V, et al. Characterizing treatment pathways at scale using the OHDSI network. *Proc Natl Acad Sci U S A* 2016 Jul 05;113(27):7329-7336 [FREE Full text] [doi: [10.1073/pnas.1510502113](https://doi.org/10.1073/pnas.1510502113)] [Medline: [27274072](https://pubmed.ncbi.nlm.nih.gov/27274072/)]
  45. Rieke N, Hancox J, Li W, Milletari F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. *NPJ Digit Med* 2020 Sep 14;3(1):119 [FREE Full text] [doi: [10.1038/s41746-020-00323-1](https://doi.org/10.1038/s41746-020-00323-1)] [Medline: [34518641](https://pubmed.ncbi.nlm.nih.gov/34518641/)]
  46. Sheller MJ, Edwards B, Reina GA, Martin J, Pati S, Kotrotsou A, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Sci Rep* 2020 Jul 28;10(1):12598 [FREE Full text] [doi: [10.1038/s41598-020-69250-1](https://doi.org/10.1038/s41598-020-69250-1)] [Medline: [32724046](https://pubmed.ncbi.nlm.nih.gov/32724046/)]
  47. Kaissis GA, Makowski MR, Rückert D, Braren RF. Secure, privacy-preserving and federated machine learning in medical imaging. *Nat Mach Intell* 2020 Jun 08;2(6):305-311. [doi: [10.1038/s42256-020-0186-1](https://doi.org/10.1038/s42256-020-0186-1)]
  48. McMahan B, Moore E, Ramage D, Hampson S, y Arcas B. Communication-efficient learning of deep networks from decentralized data. In: *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*. 2017 Presented at: 20th International Conference on Artificial Intelligence and Statistics (AISTATS); April 20-22, 2017; Fort Lauderdale, FL, USA.

49. Li T, Sahu A, Zaheer M, Sanjabi M, Talwalkar A, Smith V. Federated optimization in heterogeneous networks. *MLSys*. 2020. URL: <https://arxiv.org/abs/1812.06127> [accessed 2021-11-27]
50. Konečný J, McMahan H, Ramage D, Richtárik P. Federated optimization: distributed machine learning for on-device intelligence. *arXiv*. 2016. URL: <https://arxiv.org/abs/1610.02527> [accessed 2021-11-27]
51. Scheibner J, Raisaro JL, Troncoso-Pastoriza JR, Ienca M, Fellay J, Vayena E, et al. Revolutionizing medical data sharing using advanced privacy-enhancing technologies: technical, legal, and ethical synthesis. *J Med Internet Res* 2021 Feb 25;23(2):e25120 [FREE Full text] [doi: [10.2196/25120](https://doi.org/10.2196/25120)] [Medline: [33629963](https://pubmed.ncbi.nlm.nih.gov/33629963/)]
52. Liu D, Dligach D, Miller T. Two-stage federated phenotyping and patient representation learning. In: *Proceedings of the 18th BioNLP Workshop and Shared Task*. 2019 Presented at: 18th BioNLP Workshop and Shared Task; August, 2019; Florence, Italy p. 283-291. [doi: [10.18653/v1/w19-5030](https://doi.org/10.18653/v1/w19-5030)]
53. Liu D, Miller T, Sayeed R, Mandl KD. FADL: Federated-autonomous deep learning for distributed electronic health record. In: *Proceedings of the NIPS Machine Learning for Health (ML4H) Workshop*. 2018 Presented at: NIPS Machine Learning for Health (ML4H) Workshop; December 08, 2018; Montreal, Canada URL: <https://arxiv.org/abs/1811.11400>
54. Huang L, Yin Y, Fu Z, Zhang S, Deng H, Liu D. LoAdaBoost: Loss-based AdaBoost federated machine learning with reduced computational complexity on IID and non-IID intensive care data. *PLoS One* 2020 Apr 17;15(4):e0230706 [FREE Full text] [doi: [10.1371/journal.pone.0230706](https://doi.org/10.1371/journal.pone.0230706)] [Medline: [32302316](https://pubmed.ncbi.nlm.nih.gov/32302316/)]
55. Huang L, Shea AL, Qian H, Masurkar A, Deng H, Liu D. Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *J Biomed Inform* 2019 Nov;99:103291 [FREE Full text] [doi: [10.1016/j.jbi.2019.103291](https://doi.org/10.1016/j.jbi.2019.103291)] [Medline: [31560949](https://pubmed.ncbi.nlm.nih.gov/31560949/)]
56. Lee J, Sun J, Wang F, Wang S, Jun C, Jiang X. Privacy-preserving patient similarity learning in a federated environment: development and analysis. *JMIR Med Inform* 2018 Apr 13;6(2):e20 [FREE Full text] [doi: [10.2196/medinform.7744](https://doi.org/10.2196/medinform.7744)] [Medline: [29653917](https://pubmed.ncbi.nlm.nih.gov/29653917/)]
57. Li W, Milletari F, Xu D, Rieke N, Hancox J, Zhu W, et al. Privacy-preserving federated brain tumour segmentation. In: *Proceedings of the International Workshop on Machine Learning in Medical Imaging*. 2019 Presented at: International Workshop on Machine Learning in Medical Imaging; October 13, 2019; Shenzhen, China p. 133-141. [doi: [10.1007/978-3-030-32692-0\\_16](https://doi.org/10.1007/978-3-030-32692-0_16)]
58. Gentry C. Computing arbitrary functions of encrypted data. *Commun ACM* 2010 Mar;53(3):97-105. [doi: [10.1145/1666420.1666444](https://doi.org/10.1145/1666420.1666444)]
59. Choudhury O, Gkoulalas-Divanis A, Salonidis T, Sylla I, Park Y, Hsu G, et al. Anonymizing data for privacy-preserving federated learning. In: *Proceedings of the European Conference on Artificial Intelligence (ECAI)*. 2020 Presented at: European Conference on Artificial Intelligence (ECAI); 2020; Santiago de Compostela, Spain URL: <https://arxiv.org/abs/2002.09096>
60. Kholod I, Yanaki E, Fomichev D, Shalugin E, Novikova E, Filippov E, et al. Open-source federated learning frameworks for IoT: a comparative review and analysis. *Sensors* 2020 Dec 29;21(1):167 [FREE Full text] [doi: [10.3390/s21010167](https://doi.org/10.3390/s21010167)] [Medline: [33383803](https://pubmed.ncbi.nlm.nih.gov/33383803/)]
61. Beutel D, Topal T, Mathur A, Qiu X, Parcollet T, Gusmão PD, et al. Flower: a friendly federated learning research framework. *arXiv*. 2020. URL: <https://arxiv.org/abs/2007.14390> [accessed 2021-11-27]
62. Brouwer ED, Simm J, Arany A, Moreau Y. Deep ensemble tensor factorization for longitudinal patient trajectories classification. *arXiv*. 2018. URL: <https://arxiv.org/abs/1811.10501> [accessed 2021-11-26]
63. Naumzik C, Feuerriegel S, Nielsen A. Data-driven dynamic treatment planning for chronic diseases. *Working Paper* 2022:1-30 [FREE Full text]
64. Maag B, Feuerriegel S, Kraus M, Saar-Tsechansky M, Züger T. Modeling longitudinal dynamics of comorbidities. In: *Proceedings of the Conference on Health, Inference, and Learning (CHIL)*. 2021 Presented at: Conference on Health, Inference, and Learning (CHIL); April 8 - 10, 2021; Virtual. [doi: [10.1145/3450439.3451871](https://doi.org/10.1145/3450439.3451871)]
65. Shirley KE, Small DS, Lynch KG, Maisto SA, Oslin DW. Hidden Markov models for alcoholism treatment trial data. *Ann Appl Stat* 2010 Mar 1;4(1):366-395. [doi: [10.1214/09-aoas282](https://doi.org/10.1214/09-aoas282)]
66. Scott SL, James GM, Sugar CA. Hidden Markov models for longitudinal comparisons. *J Am Stat Assoc* 2005;100(470):359-369. [doi: [10.1198/016214504000001592](https://doi.org/10.1198/016214504000001592)]
67. Bartolomeo N, Trerotoli P, Serio G. Progression of liver cirrhosis to HCC: an application of hidden Markov model. *BMC Med Res Methodol* 2011;11(1):38 [FREE Full text] [doi: [10.1186/1471-2288-11-38](https://doi.org/10.1186/1471-2288-11-38)] [Medline: [21457586](https://pubmed.ncbi.nlm.nih.gov/21457586/)]
68. Daniel R, Cousens S, De Stavola B, Kenward MG, Sterne JA. Methods for dealing with time-dependent confounding. *Stat Med* 2013;32(9):1584-1618. [doi: [10.1002/sim.5686](https://doi.org/10.1002/sim.5686)] [Medline: [23208861](https://pubmed.ncbi.nlm.nih.gov/23208861/)]
69. Hatt T, Feuerriegel S. Estimating average treatment effects via orthogonal regularization. In: *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM)*. 2021 Presented at: ACM International Conference on Information and Knowledge Management (CIKM); November 1 - 5, 2021; Virtual. [doi: [10.1145/3459637.3482339](https://doi.org/10.1145/3459637.3482339)]
70. Schaefer A, Bailey M, Shechter S, Roberts M. Modeling medical treatment using Markov decision processes. *Oper Res Health Care* 2005;70:593-612. [doi: [10.1007/1-4020-8066-2\\_23](https://doi.org/10.1007/1-4020-8066-2_23)]



71. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med* 2018 Nov 22;24(11):1716-1720. [doi: [10.1038/s41591-018-0213-5](https://doi.org/10.1038/s41591-018-0213-5)] [Medline: [30349085](https://pubmed.ncbi.nlm.nih.gov/30349085/)]
72. Murphy SA, van der Laan MJ, Robins JM, CPPRG. Marginal mean models for dynamic regimes. *J Am Stat Assoc* 2001 Dec 01;96(456):1410-1423 [FREE Full text] [doi: [10.1198/016214501753382327](https://doi.org/10.1198/016214501753382327)] [Medline: [20019887](https://pubmed.ncbi.nlm.nih.gov/20019887/)]
73. Murphy S. Optimal dynamic treatment regimes. *J Royal Stat Soc : Series B (Stat Methodol)* 2003;65(2):331-355. [doi: [10.1111/1467-9868.00389](https://doi.org/10.1111/1467-9868.00389)]
74. Liao P, Klasnja P, Murphy S. Off-policy estimation of long-term average outcomes with applications to mobile health. *J Am Stat Assoc* 2021 Oct 01;116(533):382-391. [doi: [10.1080/01621459.2020.1807993](https://doi.org/10.1080/01621459.2020.1807993)] [Medline: [33814653](https://pubmed.ncbi.nlm.nih.gov/33814653/)]
75. Brown S. Patient similarity: emerging concepts in systems and precision medicine. *Front Physiol* 2016 Nov 24;7:561 [FREE Full text] [doi: [10.3389/fphys.2016.00561](https://doi.org/10.3389/fphys.2016.00561)] [Medline: [27932992](https://pubmed.ncbi.nlm.nih.gov/27932992/)]
76. Allam A, Dittberner M, Sintsova A, Brodbeck D, Krauthammer M. Patient similarity analysis with longitudinal health data. arXiv. 2020. URL: <https://arxiv.org/abs/2005.06630> [accessed 2021-11-26]
77. Carrasco-Ribelles LA, Pardo-Mas JR, Tortajada S, Sáez C, Valdivieso B, García-Gómez JM. Predicting morbidity by local similarities in multi-scale patient trajectories. *J Biomed Inform* 2021 Aug;120:103837. [doi: [10.1016/j.jbi.2021.103837](https://doi.org/10.1016/j.jbi.2021.103837)] [Medline: [34119690](https://pubmed.ncbi.nlm.nih.gov/34119690/)]
78. Giabbanelli PJ, Torsney-Weir T, Mago VK. A fuzzy cognitive map of the psychosocial determinants of obesity. *Appl Soft Comput* 2012 Dec;12(12):3711-3724. [doi: [10.1016/j.asoc.2012.02.006](https://doi.org/10.1016/j.asoc.2012.02.006)]
79. Ghassempour S, Girosi F, Maeder A. Clustering multivariate time series using hidden Markov models. *Int J Environ Res Public Health* 2014 Mar 06;11(3):2741-2763 [FREE Full text] [doi: [10.3390/ijerph110302741](https://doi.org/10.3390/ijerph110302741)] [Medline: [24662996](https://pubmed.ncbi.nlm.nih.gov/24662996/)]
80. Westergaard D, Moseley P, Sørup FKH, Baldi P, Brunak S. Population-wide analysis of differences in disease progression patterns in men and women. *Nat Commun* 2019 Feb 08;10(1):666 [FREE Full text] [doi: [10.1038/s41467-019-08475-9](https://doi.org/10.1038/s41467-019-08475-9)] [Medline: [30737381](https://pubmed.ncbi.nlm.nih.gov/30737381/)]
81. Lademann M, Lademann M, Jensen AB, Brunak S. Incorporating symptom data in longitudinal disease trajectories for more detailed patient stratification. *Int J Med Inform* 2019 Sep;129:107-113 [FREE Full text] [doi: [10.1016/j.ijmedinf.2019.06.003](https://doi.org/10.1016/j.ijmedinf.2019.06.003)] [Medline: [31445244](https://pubmed.ncbi.nlm.nih.gov/31445244/)]
82. Oh W, Kim E, Castro MR, Caraballo PJ, Kumar V, Steinbach MS, et al. Type 2 diabetes mellitus trajectories and associated risks. *Big Data* 2016;4(1):25-30 [FREE Full text] [doi: [10.1089/big.2015.0029](https://doi.org/10.1089/big.2015.0029)] [Medline: [27158565](https://pubmed.ncbi.nlm.nih.gov/27158565/)]
83. Zhang Y, Padman R. Innovations in chronic care delivery using data-driven clinical pathways. *Am J Manag Care* 2015 Dec 01;21(12):661-668 [FREE Full text] [Medline: [26760429](https://pubmed.ncbi.nlm.nih.gov/26760429/)]
84. Huang Z, Dong W, Ji L, Gan C, Lu X, Duan H. Discovery of clinical pathway patterns from event logs using probabilistic topic models. *J Biomed Inform* 2014 Feb;47:39-57 [FREE Full text] [doi: [10.1016/j.jbi.2013.09.003](https://doi.org/10.1016/j.jbi.2013.09.003)] [Medline: [24076435](https://pubmed.ncbi.nlm.nih.gov/24076435/)]
85. Dabek F, Caban J. A grammar-based approach to model the patient's clinical trajectory after a mild traumatic brain injury. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2015 Presented at: IEEE International Conference on Bioinformatics and Biomedicine (BIBM); November 9 - 12, 2015; Washington, DC. [doi: [10.1109/bibm.2015.7359775](https://doi.org/10.1109/bibm.2015.7359775)]
86. Li X, Liu H, Mei J, Yu Y, Xie G. Mining temporal and data constraints associated with outcomes for care pathways. *Stud Health Technol Inform* 2015;216:711-715. [Medline: [26262144](https://pubmed.ncbi.nlm.nih.gov/26262144/)]
87. Combi C, Mantovani M, Sala P. Discovering quantitative temporal functional dependencies on clinical data. In: Proceedings of the IEEE International Conference on Healthcare Informatics (ICHI). 2017 Presented at: IEEE International Conference on Healthcare Informatics (ICHI); August 23 - 26, 2017; Park City, UT. [doi: [10.1109/ichi.2017.80](https://doi.org/10.1109/ichi.2017.80)]
88. Giabbanelli PJ. Analyzing the complexity of behavioural factors influencing weight in adults. In: Giabbanelli PJ, Mago V, Papageorgiou E, editors. *Advanced Data Analytics in Health*. Cham, Switzerland: Springer; 2018:163-181.
89. Scherrer N, Bilaniuk O, Annadani Y, Goyal A, Schwab P, Schölkopf B, et al. Learning neural causal models with active interventions. arXiv. 2021. URL: <https://arxiv.org/abs/2109.02429> [accessed 2021-11-26]
90. Topol EJ. Welcoming new guidelines for AI clinical research. *Nat Med* 2020 Sep 09;26(9):1318-1320. [doi: [10.1038/s41591-020-1042-x](https://doi.org/10.1038/s41591-020-1042-x)] [Medline: [32908274](https://pubmed.ncbi.nlm.nih.gov/32908274/)]
91. Vokinger KN, Feuerriegel S, Kesselheim AS. Mitigating bias in machine learning for medicine. *Commun Med* 2021;1(1):25. [doi: [10.1038/s43856-021-00028-w](https://doi.org/10.1038/s43856-021-00028-w)] [Medline: [34522916](https://pubmed.ncbi.nlm.nih.gov/34522916/)]
92. Finlayson SG, Subbaswamy A, Singh K, Bowers J, Kupke A, Zittrain J, et al. The clinician and dataset shift in artificial intelligence. *N Engl J Med* 2021 Jul 15;385(3):283-286. [doi: [10.1056/nejmc2104626](https://doi.org/10.1056/nejmc2104626)]
93. Arif SO, Shor J, Sinha R, Yoon J, Ledsam JR, Le LT, et al. A prospective evaluation of AI-augmented epidemiology to forecast COVID-19 in the USA and Japan. *NPJ Digit Med* 2021 Oct 08;4(1):146 [FREE Full text] [doi: [10.1038/s41746-021-00511-7](https://doi.org/10.1038/s41746-021-00511-7)] [Medline: [34625656](https://pubmed.ncbi.nlm.nih.gov/34625656/)]
94. Roland T, Boeck C, Tschoellitsch T, Maletzky A, Hochreiter S, Meier J, et al. Machine learning based COVID-19 diagnosis from blood tests with robustness to domain shifts. medRxiv. 2021. URL: <https://www.medrxiv.org/content/10.1101/2021.04.06.21254997v1> [accessed 2021-11-26]
95. Jarrett D, Yoon J, Bica I, Qian Z, Ercole A, van der Schaar M. Clairvoyance: a unified, end-to-end AutoML pipeline for medical time series. In: Proceedings of the International Conference on Learning Representations (ICLR). 2021 Presented

- at: International Conference on Learning Representations (ICLR); May 3 - 7, 2021; Vienna, Austria URL: <https://openreview.net/forum?id=xnC8YwKUE3k>
96. Kakarmath S, Esteva A, Arnaout R, Harvey H, Kumar S, Muse E, et al. Best practices for authors of healthcare-related artificial intelligence manuscripts. NPJ Digit Med 2020 Oct 16;3(1):134 [FREE Full text] [doi: [10.1038/s41746-020-00336-w](https://doi.org/10.1038/s41746-020-00336-w)] [Medline: [33083569](https://pubmed.ncbi.nlm.nih.gov/33083569/)]
  97. Allam A, Nagy M, Thoma G, Krauthammer M. Neural networks versus Logistic regression for 30 days all-cause readmission prediction. Sci Rep 2019 Jun 26;9(1):9277 [FREE Full text] [doi: [10.1038/s41598-019-45685-z](https://doi.org/10.1038/s41598-019-45685-z)] [Medline: [31243311](https://pubmed.ncbi.nlm.nih.gov/31243311/)]
  98. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. NPJ Digit Med 2018 Mar 14;1(1):5 [FREE Full text] [doi: [10.1038/s41746-017-0012-2](https://doi.org/10.1038/s41746-017-0012-2)] [Medline: [31304291](https://pubmed.ncbi.nlm.nih.gov/31304291/)]
  99. Lundberg SM, Nair B, Vavilala MS, Horibe M, Eisses MJ, Adams T, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. Nat Biomed Eng 2018 Oct 10;2(10):749-760 [FREE Full text] [doi: [10.1038/s41551-018-0304-0](https://doi.org/10.1038/s41551-018-0304-0)] [Medline: [31001455](https://pubmed.ncbi.nlm.nih.gov/31001455/)]
  100. Kirkpatrick C, Dahlquist J. Technical Analysis: The Complete Resource for Financial Market Technicians. Upper Saddle River, NJ: Financial Times Press; 2006.
  101. Lemhadri I, Ruan F, Abraham L, Tibshirani R. LassoNet: a neural network with feature sparsity. J Machin Learn Res 2021;22(127):1-29 [FREE Full text]
  102. Nápoles G, Jastrzębska A, Salgueiro Y. Pattern classification with evolving long-term cognitive networks. Inform Sci 2021 Feb;548:461-478. [doi: [10.1016/j.ins.2020.08.058](https://doi.org/10.1016/j.ins.2020.08.058)]

## Abbreviations

- AI:** artificial intelligence  
**AutoML:** automated machine learning  
**ICU:** intensive care unit  
**RNN:** recurrent neural network

*Edited by R Kukařka; submitted 21.04.21; peer-reviewed by Z Ren, P Giabbanelli, E Mahmoudi; comments to author 27.05.21; revised version received 26.07.21; accepted 29.10.21; published 03.12.21*

*Please cite as:*

Allam A, Feuerriegel S, Rebhan M, Krauthammer M  
Analyzing Patient Trajectories With Artificial Intelligence  
J Med Internet Res 2021;23(12):e29812  
URL: <https://www.jmir.org/2021/12/e29812>  
doi: [10.2196/29812](https://doi.org/10.2196/29812)  
PMID:

©Ahmed Allam, Stefan Feuerriegel, Michael Rebhan, Michael Krauthammer. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 03.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.