Journal of Biomedical Informatics 75 (2017) 70–82

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

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

Procedure prediction from symbolic Electronic Health Records via time intervals analytics

Robert Moskovitch [a,b,c,d,e,*], Fernanda Polubriaginof [a,b,c,d], Aviram Weiss [e], Patrick Ryan [a], Nicholas Tatonetti [a,b,c,d,*]

[a] Department of Biomedical Informatics, Columbia University, NY, USA
[b] Department of Systems Biology, Columbia University, NY, USA
[c] Department of Medicine, Columbia University, NY, USA
[d] Observational Health Data Sciences and Informations (OHDSI), NY, USA
[e] Department of Software and Information Systems Engineering, Ben Gurion Univeristy, Beer Sheva, Israel

ARTICLE INFO

Article history:
Received 15 May 2016
Revised 19 June 2017
Accepted 25 July 2017
Available online 17 August 2017

Keywords:
Temporal data mining
Time intervals mining
Prediction
Electronic Health Records

ABSTRACT

Prediction of medical events, such as clinical procedures, is essential for preventing disease, understanding disease mechanism, and increasing patient quality of care. Although longitudinal clinical data from Electronic Health Records provides opportunities to develop predictive models, the use of these data faces significant challenges. Primarily, while the data are longitudinal and represent thousands of conceptual events having duration, they are also sparse, complicating the application of traditional analysis approaches. Furthermore, the framework presented here takes advantage of the events duration and gaps. International standards for electronic healthcare data represent data elements, such as procedures, conditions, and drug exposures, using *eras*, or time intervals. Such eras contain both an event and a duration and enable the application of time intervals mining – a relatively new subfield of data mining. In this study, we present Maitreya, a framework for time intervals analytics in longitudinal clinical data. Maitreya discovers frequent time intervals related patterns (TIRPs), which we use as prognostic markers for modelling clinical events. We introduce three novel TIRP metrics that are normalized versions of the horizontal-support, that represents the number of TIRP instances per patient. We evaluate Maitreya on 28 frequent and clinically important procedures, using the three novel TIRP representation metrics in comparison to no temporal representation and previous TIRPs metrics. We also evaluate the epsilon value that makes Allen's relations more flexible with several settings of 30, 60, 90 and 180 days in comparison to the default zero. For twenty-two of these procedures, the use of temporal patterns as predictors was superior to non-temporal features, and the use of the vertically normalized horizontal support metric to represent TIRPs as features was most effective. The use of the epsilon value with thirty days was slightly better than the zero.

1. Introduction

Predictive outcome modelling by employing the analysis of the data along time is one of the most important research fields in biomedical informatics [22]. Developing clinically predictive methods can improve the quality of patient treatment and act to prevent expected negative outcomes, such as in clinical procedures or diagnoses [31]. The increasing availability of Electronic Health Records (EHR) provides researchers with a unique opportunity to work closely with patient history to predict outcomes. Traditionally, such studies have focused on a specific outcome and require domain experts to define relevant variables phenotypes) that may be predictive [6,55,17,47,31,22]. Here, we propose a general approach using simple temporal abstraction of EHR data and temporal pattern recognition to automatically build predictive models for a range of clinical outcomes from symbolic data, which refers to events in the data that have a symbol or a concept, as it is more commonly referred to in the biomedical domain. The conceptual data that are used in this study are conditions, which can be a disease or a clinical state, procedures and drug exposures. These are represented by symbolic time intervals, that are defined by their start-time, end-time and symbol, or a concept id. The concept id

* Corresponding authors at: Department of Software and Information Systems Engineering, Ben Gurion Univeristy, Beer Sheva, Israel (R. Moskovitch). Department of Biomedical Informatics, Columbia University, NY, USA (N. Tatonetti).
E-mail addresses: robert.moskovitch@columbia.edu (R. Moskovitch), nick.tatonetti@columbia.edu (N. Tatonetti).

https://doi.org/10.1016/j.jbi.2017.07.018
1532-0464/

can be for example the condition or procedure id in a given vocabulary, such as ICD-9.

EHR data are recorded only when patients enter the healthcare system, providing a sparse and biased view of a patient's clinical history [19]. In addition, the data appear in several forms, from numerical values occurring at a particular time (*e.g.* lab tests) to events that may span many days, months, or years (*e.g.* conditions, and drug exposures). These challenges complicate the use of existing temporal modelling strategies ([70,22,42,11,2,45,56]). Although the study of temporal data mining has advanced in the last decade (Section 2.3), relatively little work has been dedicated to the unique challenges of the biomedical domain [22].

In recent years, the Observational Medical Outcomes Partnership OMOP) project [38] introduced a standard database scheme for pharmacovigilance research, in which conditions, drug exposures, and procedures are represented by duration events (time intervals) called 'eras'. In fact, the data we used in this study, in order to perform our experiments was in that format. In this paper, we introduce Maitreya, a framework for symbolic time intervals analysis of medical data. In our approach, we represent the patient's health records as a series of symbolic time intervals. Each symbolic time interval consists of a symbol (i.e. a clinical concept), a start time, and an end-time. We use an observation period of one year to predict the occurrence of a clinical procedure, and mask the data for month immediately preceding the event (Fig. 1).

We then use the KarmaLego algorithm [27,26] for the discovery of frequent Time Intervals Related Patterns (TIRPs). KarmaLego is a fast algorithm for TIRPs discovery from symbolic time intervals data. In a recent paper by Moskovitch and Shahar [28] a framework for the classification of multivariate temporal data was described, in which the temporal variables are abstracted using Equal Width Discretization and Symbolic Aggregation Approximation SAX) [23] and then TIRPs are discovered by KarmaLego. In that framework, unlike in this study, KarmaLego was applied to each class separately and TIRP features were unified for the classification purpose. In this study, the Maitreya framework learns a model that consists only on a single class the cohort of the patients having the outcome) and the model is learnt from the TIRPs that were discovered from this class. Then a controls class of patients is used only for the evaluation, but not for the predictive model, as in [28]. Due to the large number (i.e., hundreds) of symbols in the datasets in this study, we used an advanced version of KarmaLego that has a more efficient index structure for the pairs of symbolic time intervals (2-sized TIRPs). Each patient's clinical history is then represented by a set of TIRPs as features. Eventually, we use random forest classification to model the occurrence of a procedure as a function of patient history.

Thus, in this paper we describe a methodology for the prediction of outcomes from symbolic (i.e., conditions, procedures, drug exposures) within the Maitreya framework, which consists only on the TIRPs that are discovered from the cohort data. In comparison to the KarmaLegoS framework that was described in [28], here are the main novel contributions. In Maitreya the model consists on TIRPs discovered only from the cohort class, and not from the controls class, since it is prediction rather than classification and the controls are used only for evaluation using frequent TIRPs from the controls class would potentially improve the results and give a more optimistic estimation). Three novel metrics for TIRPs representations are introduced, and a comparison to the non use of temporal order of the data as a baseline which was not performed in [28]. We also compare the use of two temporal relations sets, the Allen's [3] seven relations and our three more general relations, and an epsilon value that enables to relax the temporal relations crispiness. TIRPs with several metrics for prediction to a baseline model that does not consider the temporal order of the data in its analysis. Finally, in order to evaluate the use of Maitreya for outcomes prediction in EHR based on symbolic time intervals data, we consider 28 of the most frequent clinical procedures at our institution (Columbia University Medical Centre) electronic health records database.

2. Related work

Multivariate temporal data exist across many scientific disciplines [46,62,61]. In many domains, such as Electronic Health Records, they are sparse, heterogeneous in both type and frequency), and have variable duration [19]. The use of Symbolic Time Intervals representation, the development of fast and efficient Time Intervals Related Patterns mining algorithms, and their use for classification and prediction in multivariate time series [14] to analyse these data are covered in this section.

2.1. Temporal data analytics in biomedical data

The development of predictive modelling in clinical medicine through data mining is an important and developing field [6,22,53]. The common and widespread use of EHR among many medical care providers creates accessible large amount of medical data and enables the application of machine learning models, in order to expose new medical knowledge regarding the prediction of medical conditions, use of resources and the effectiveness of medical care [22,31]. However, while recently increasing number of papers report methods that exploit the time dimension, such [12] that mined temporal dependencies in clinical groupings, or provide a methodology to define and represent temporal workflows [13]. Sala et al. [40] describe a framework for the discovery of evolution rules in time) in clinical data, and Yang et al. [54] who used the time dimension explicitly for the analysis in the important problem of Adverse Drug Reaction [7]. However, there are still many key papers in this field that have excluded temporal data from their analyses [31,47]. This is a significant omission, as time is a key aspect of biological processes and clinical decisions are based on complex combinations of different events which span across different time intervals. Consequentially, EHRs are inherently longitudinal, and temporal analysis is an important challenge being addressed in the biomedical [22] and data mining literature [5,39,20].

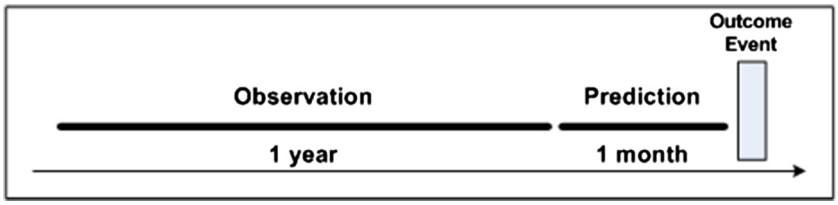

Fig. 1. Prediction of an outcome event based on an observation period. The prediction period is given between the outcome events and observation time and no data from that period is used.

The main challenge in EHR data analysis is its sparsity and irregular sampling nature, since data commonly recorded only when patients enter the healthcare system, providing a sparse and biased view of a patient's clinical history [16,18]. In addition, the data appear in several forms, from numerical values occurring at a particular time (e.g. lab tests) to events that may span many days, months, or years (e.g. conditions, and drug exposures). These challenges complicate the use of existing temporal modelling strategies [57,22,42,11,2,45,56,32]. Although the study of temporal data mining has advanced in the last decades, relatively little work has been dedicated to the unique challenges of the biomedical domain [22]. In recent years, the Observational Medical Outcomes Partnership (OMOP) project introduced a standard database scheme for pharmacovigilance research, in which conditions, drug exposures, and procedures are represented by concept events that have duration (symbolic time intervals) called 'eras'. In the biomedical domain, thinking in time durations is intrinsic, for example, a period of time a patient is prescribed for some drug, or treatment, a period of time a condition is relevant, and a procedure, which is the part of the reason in this paper we focus the methodology on the use of symbolic time intervals [26].

Recent studies reported in the biomedical informatics and data mining literature propose several approaches in mining EHR data to cope with the data challenges. The advancements in making EHR data available has enabled the development of new methods that incorporate temporal data analysis for mainly predictive modelling, a well known need [16,22,1,37,44]. It was shown that through modelling the lab-test values irregularity parameters, it is possible to find patterns for values forecasting and prediction [35,18,37]. Important efforts were in analysing longitudinal data from EHR and visualize it for physicians to understand better patients' history [58,59]. Recent studies had shown the advantage in using longitudinal lab tests data using kalman filter over not using longitudinal data for risk analysis [35]. Simple HbA1c testing sequence patterns were learned to examine following guidelines [36] and these challenges are further investigated with the increase in longitudinal data availability. However, in this paper the focus is on prediction based on symbolic data, including conditions, procedures and drug exposures, rather than the use of numerical values series of lab tests.

2.2. Clinical procedures and their importance

Clinical interventions, such as an outpatient clinic visit, emergency department visit or hospitalization are usually driven by one or more medical complaints of the patient. The medical team has to understand the complaint, gather additional information in order to diagnose the patient's condition (or conditions), and provide therapeutic care. Diagnostic techniques involve asking questions, physical examination and different procedures (such as imaging tests and lab tests).

Therapy also involves a variety of procedures – starting with patient education, through different types of physical therapy, drug administration to surgical and non-surgical procedures. The different purpose of procedures – diagnostic and therapeutic – are at the heart of medical care, and the ability to predict their need is highly important in several perspectives. In a resource-limited environment, in which procedures are often expensive, the ability to pre-plan the use of procedure facilities (such as operation rooms, CT and MRI suites) is essential for efficient operation of a medical organization. Additionally, procedures are often good indicators of the patient's health status. For example, mechanical ventilation is usually a procedure performed in Intensive Care Units for patients whose respiratory, cardiac or neurologic systems fail, and the beginning of Insulin injection is a marker of imbalanced diabetes mellitus. Thus, being able to predict the need of a procedure for a patient in the future based on their history has the potential to intervene and hopefully prevent the need.

2.3. Symbolic time intervals data in EHR

Representing biomedical events data using symbolic time intervals is intuitive and natural because of intrinsic nature of the temporal components in biomedical data and specifically in EHR data that often have duration, such as drug exposures, conditions, and procedures [34,5]. Symbolic time intervals data exist as raw data, as in our study when using conditions, procedures or drugs codes that have duration. Alternatively, symbolic time intervals can be extracted from time point series, such as lab-tests, in a process referred to as Temporal Abstraction [42,23,28,29], but it is not in the scope of this study. Moreover, in the OHDSI project [19] the definition of the Observational Medical Outcomes Partnership (OMOP) database scheme that was designed for pharma purposes analysis, there is an explicit emphasis of the nature of these components (i.e., conditions, procedures and drug exposures) that are called 'era's. An era has a start time, end time and a concept id, which is in fact symbolic time interval that as we defined more formally later consists on a symbol, start time and end time.

When the data are already represented by multivariate symbolic time intervals, as in our study, typically an abstraction is performed on symbolic time intervals having the same symbol. In this case, events that occur within a predefined gap will be concatenated into a single time interval. In this paper, we followed the standard aggregation process as defined in OMOP [38] to abstract EHR data. Additionally, we used the SNOMED-CT hierarchy to use more general concepts achieving higher generalizability and enabling the discovery of more frequent temporal patterns.

2.4. Symbolic time intervals mining

Allen's temporal relations [3] are used often to represent the relations between pairs of time intervals in symbolic time intervals mining algorithms. There are seven relations: *before, meet, overlap, contain, starts, finishes, and equals* (Fig. 2). However, using more general temporal relations may be more effective for prediction tasks [28,29], as will be experimented in this study. In this paper, we shall refer to a conjunction of temporal relations between pairs of time intervals as a Time Intervals Related Pattern (TIRP), as defined later in the Methods section (Definition 5). Over the past fifteen years, several algorithms were proposed for time intervals mining improving their efficiency and resulted runtime [33,24,27,8,9,21,48–51]. Among these, KarmaLego [27] uses an efficient data structure and exploits the transitivity property of the temporal relations to optimize the candidate generation process. KarmaLego discovers the complete set of TIRPs, since it discovers their complete set of the TIRPs' instances [28]. The ability to discover the complete set of TIRPs is important since it allows to represent TIRPs as classification features with other representations besides Boolean, that was used in earlier studies [34,5] allowing us to introduce novel TIRP metrics, or representations, in this paper.

2.5. Classification and prediction with time intervals related patterns

The use of frequent patterns for classification and prediction of data and specifically multivariate temporal data is increasingly reported in the past decade [10]. It is based on the idea that specific patterns, contain richer predictive information than the independent features. This is especially intuitive with temporal patterns, which represent the order of events in time as predictors [52]. Thus, once TIRPs are discovered from longitudinal data, they can be used as predictors for regression and classification [34,25,32].

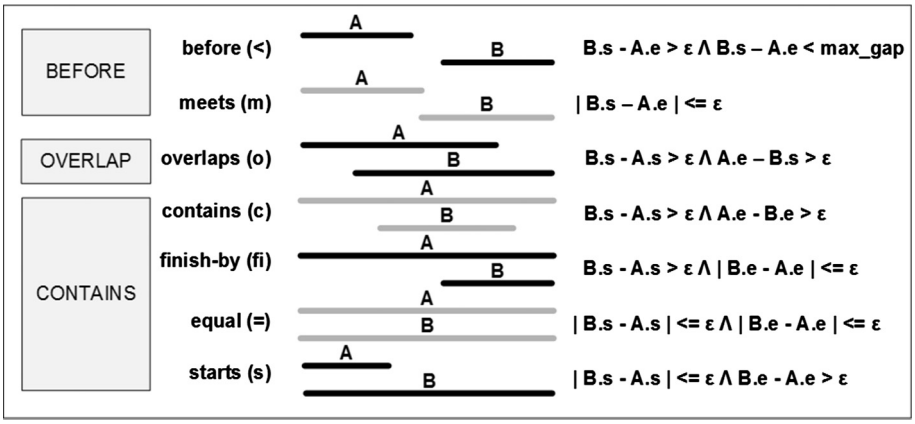

Fig. 2. A flexible extension of Allen's temporal relations, using the same epsilon for all of the relations. On the left three abstract temporal relations are presented, which are disjunctions of the seven relations.

These TIRPs are a more expressive representation of multivariate longitudinal data when compared to atemporal features that ignore the order of the events. Several groups simultaneously proposed using TIRPs as features for classifying multivariate time series ([34,60,4,5]). Interestingly, all of the studies that were published in the data mining literature reported the use of temporal abstraction and TIRPs for classification were applied to problems in the biomedical domain [28,29]. Several studies had shown the advantages of using TIRPs over atemporal representation in classifying multivariate temporal data [34,4,30,26] and other decreased the number of discovered patterns while retaining accuracy ([25,43]).

3. Methods

In this section we present a framework of definitions that is based on [57,27]. However, we add more definitions related to the novel TIRPs representation metrics for classification that are presented in this paper for the first time, such as the normalized versions of the Horizontal Support, which is a metric that represents a TIRP by the number of its instances found in a single patient's records, as will be defined soon.

3.1. Definitions

A symbolic time interval, I = ⟨s, e, sym⟩, is an ordered pair of time points, start-time (s) and end-time (e), and a symbol (sym) that represents one of the domain's temporal concepts, which in our study can be clinical procedures, conditions or drug exposures. Along the paper we will use these annotations to refer to a symbolic time interval start time – *I.s*, end time *I.e*, and symbol *I.sym*. In order to describe the temporal relations among a pair of symbolic time intervals we use Allen's [3] temporal relations. Allen defined seven temporal relations shown in Fig. 2 with thirteen descriptions, including the seven that are shown in Fig. 2 on the left, and their inverse (for example, the inverse of *A before B* is *B after A*, or of *A starts B* is *B started-by A*, while *equal* is the same for both). However, the use of all thirteen relations is not required if the symbolic time intervals data is ordered in a lexicographical order, as we defined and used here. Since we are interested in the discovery of frequent TIRPs we defined a more flexible version of Allen's temporal relations by introducing an epsilon value, as defined in Definition 1. Thus, before mining the database for frequent TIRPs, in order to allow using only Allen's seven relations, without their inverse, the symbolic time intervals are sorted in a lexicographical order, based on their *start-time*, *end-time* (if their start-times are the same), or symbol (alphabetically if their start-times and end-times are the same).

In addition to the flexible version of Allen's temporal relations, we use a maximal gap constraint in the temporal relation 'before' which limits the pairs of symbolic time intervals that will be considered in the mining process. While the max_gap is motivated mainly by computational constraints it has also clinical meaning, since in some time duration the relations among two events in time are potentially meaningless. Thus, the value of max_gap is determined both by the limitations of the memory and also based on the maximal time duration relevant from a clinical point of view. In addition to Allen's seven relations, we defined a set of more general temporal relations shown on the left in Fig. 2. These are disjunctions of Allen's relations, thus, the relation BEFORE is a disjunction of Allen's *before* or *meet*, since '*meet*' is in fact a private case of '*before*', in which the gap between the intervals is zero. Then, OVERLAP is the same as the original overlap, and CONTAIN is the disjunction of all the "containing" temporal relations.

Definition 1. To define a flexible framework of Allen's temporal relations, two relations are defined on time-stamped (point-based) data, given an epsilon value.

Given two time-points t^1 and t^2: $t^1 =^{\varepsilon} t^2$ iff $|t^2 - t^1| <=^{\varepsilon}$ and $t^1 <^{\varepsilon} t^2$ iff $t^2 - t^1 > \varepsilon$, based on the two relations $=^{\varepsilon}$ and $<^{\varepsilon}$ and the epsilon value, a flexible version of Allen's seven temporal relations is defined, as shown in Fig. 2. The use of the epsilon value is defined for each of the seven relations in order to maintain *exclusive* definitions of the temporal relations and in that way each pair of symbolic time intervals in the data can be described only by a single temporal relation. Note, the purpose of the epsilon is to make the temporal relations definitions more flexible and it is mainly effective with relations that have the same end points. For example, *meet*, in which the end-time of the first time interval (A) and the start-time of the second time-interval are identical. Similarly, *starts* or *finish-by*, in which one end point is identical, or *equal*, in which both end points have the same time. Additionally, note that Allen's seven temporal relations (1983) do not include the epsilon extension, and thus it is by default having epsilon set to zero.

A TIRP is defined by the sequence of the symbols of the symbolic time intervals, and the temporal relations among the symbolic time intervals, for which we use Allen's temporal relations. However, in order to define a TIRP in a non ambiguous way we have to define the conjunction of all the temporal relations among each pair of symbolic time intervals, as defined in Definition 2. This definition was made first by Hoppner [15] and later was used in the following methods [33,34,57,27]. Thus, Definition 2 defines a

TIRP by a sequence of k symbolic time intervals I and the set of the conjunction of the $(k^2 - k)/2$ temporal relations. In this study we used two sets of temporal relations: Allen's seven temporal relations, and our three more general temporal relations and compared their performance. The three more general temporal relations, while used with the conjunction of temporal relations as defined in Definition 2 are less limiting and the variance of the supporting instances is larger, which results with more frequent TIRPs.

Definition 2. A non-ambiguous lexicographic *Time Intervals Related Pattern* (*TIRP*) P is defined as $P = \{I, R\}$, where $I = \{I^1, I^2, \ldots, I^k\}$ is a set of k symbolic time intervals and

$$R = \bigcap_{i=1}^{k-1} \bigcap_{j=i+1}^{k} r(I^i, I^j) = \{r_{1,2}(I^1, I^2), \ldots, r_{1,k}(I^1, I^k), \ldots, r_{k-1,k}(I^{k-1}, I^k)\}$$

defines all the conjunction of all the temporal relations among each of the $(k^2 - k)/2$ pairs of symbolic time intervals in I. Fig. 3 presents a typical TIRP, represented as a half-matrix of temporal relations. We assume such a representation throughout the description of the KarmaLego algorithm.

Definition 3. Given a database of |E| distinct entities (e.g., different patients), the *vertical support* of a TIRP P is denoted by the cardinality of the set E^P of distinct entities for which P holds, divided by the total number of entities (e.g., patients) |E|: ver_sup(P) = $|E^P|/|E|$. The vertical support is actually what is commonly used as support in association rules, item-set and sequential mining. A TIRP having above minimal vertical support threshold is referred to as frequent. We distinguish this support, and call it vertical support, since we will define later a horizontal support measure that will represent the number of a TIRP instances per an entity.

3.2. The KarmaLego and SingleKarmaLego algorithms

The KarmaLego algorithm [27] enables to perform fast time intervals mining for the discovery of TIRPs through exploiting the transitivity of temporal relations that enables an efficient candidate generation mechanism. KarmaLego consists of two main steps: Karma, in which the entire set of entities' symbolic time intervals data are scanned and indexed. Through that all the symbols are counted, and each pair of symbolic time intervals and the temporal relation (Fig. 2) among them are indexed, as a tuple, in an index that contains all the frequent 2-sized TIRPs (k = 2), based on the first symbol, the second symbol and the temporal relation among them. Later the Lego algorithm extends recursively the 2-sized TIRPs that were found frequent (above the minimal vertical support) continuously.

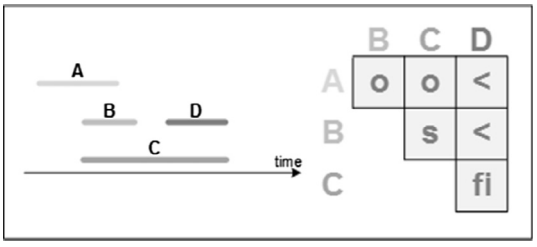

Fig. 3. An example of a Time-Interval Related Pattern (TIRP), containing a sequence of four symbolic time intervals (A, B, C, D ordered according to the lexicographical order of the start time, end time (in case the start times are the same), and symbol (in case the start times and end times are the same). The conjunction of all of the pair-wise temporal relations are shown in the half matrix on the right. For example, at the top row there are the relations of A with the following time intervals. A overlaps (o) B, A overlaps (o) C and A before (<) D, cet.

The output of the KarmaLego algorithm is an enumeration tree of all the frequent TIRPs. In this study we used an enhanced version of the KarmaLego algorithm that was designed to handle large number (i.e. thousands, such as the number of concept types typically in EHR) of symbol types, called KarmaLego^D that was introduced and rigorously evaluated [26]. KarmaLego^D indexes the 2-sized TIRPs in the Karma step using hash-tables, rather than a multi dimensional array like in [27–29]. The first KarmaLego implementation which is supposed to be faster providing a retrieval of O(1) to any two sized TIRP entry is very memory consuming, and with large number (i.e., thousands) of symbols it becomes a problem. Thus, in KarmaLego^D [26] the two sized TIRPs access was implemented by arrays and hash-tables and a hybrid version and based on an evaluation with datasets having large number of symbols it was found that the runtime is only slightly slower and even most of the time the same and the memory consumption is ideal. Detecting a set of given TIRPs can be done efficiently by SingleKarmaLego [28], compared to a naïve TIRPs detection algorithm due to the first phase of indexing of the pairs of symbolic time intervals. Inspired by KarmaLego, SingleKarmaLego indexes all the pairs of an entity's symbolic time intervals into an efficient data structure, which enables later a fast and efficient TIRPs detection.

3.3. TIRPs representation metrics

KarmaLego discovers all the complete set of instances of a TIRP, given a patient, unlike with other methods. Thus, KarmaLego discovers all the instances of a TIRP in a single patient records, and doesn't stop in the first instance as other methods do, which in a first glance seems sufficient to discovery TIRPs. However, rethinking about it is crucial in order to discover also the complete list of TIRPs, since instances that can be extended may not be the first, but rather a later instance, which may have may change the vertical support of the TIRP as explained in details in [28].

Discovering all the instances of each TIRP for each patient enables richer representations for TIRPs as classification features, such as the *Horizontal Support* (Definition 4) and *Mean Duration* (Definition 5), that were used successfully already in other clinical domains [28,29]). The horizontal support metric describes the number of instances of a TIRP that were detected for a single patient. In fact, it is analogous to the Term Frequency metric of a word in text categorization, which measures the number of times a word appears in a document [41]. The Mean Duration metric represents the average of the TIRP's instances' time length that were detected at a single patient records.

In this paper we introduce three *novel* TIRP representation metrics: two are normalized versions of the Horizontal Support, and third is an extension of the Horizontal Support which is inspired by the TFIDF metric from text categorization, which was shown to work well also in other domains, rather than text [61]. The new metrics are defined in Definitions 6-8 accordingly.

Definition 4. The *horizontal support* (HS) of a TIRP P for an entity e_i (e.g., a single patient's record), hor_sup(P, e_i) is the number of instances of the TIRP P found in e_i. For example, the number of times a particular temporal pattern P was found in a particular patient's record.

Definition 5. The *mean duration* (MND) of the n supporting instances of a given k-sized TIRP P within an entity e (e.g., within a single patient's record; note that, per Definition 3, an entity may have several supporting instances of a TIRP) is defined here. The Mean Duration is the sum of the n supporting instances durations divided by n. The instances duration is the subtraction of $I_e^{i,j}$

latest symbolic time interval j, in the k symbolic time intervals, having the latest end time (e) by $I_s^{i,1}$ the first symbolic time interval start time (s).

$$MeanDuration(P,e) = \frac{\sum_{i=1}^{n}\left(Max_{j=1}^{k} I_e^{i,j} - I_s^{i,1}\right)}{n}$$

We will use the Horizontal Support and Mean Duration metrics, and their extensions that will be defined below, for the TIRP representation as features. The horizontal support and the mean duration metrics for TIRP representation were introduced originally and applied successfully in [28]. The vertical support and horizontal support metrics are also a good way to describe the data in a hierarchical clustering way. The vertical support represents the size of the population sharing the same behaviour along time, while the horizontal support may group the supporting population according to the number of instances detected in their records. However, the horizontal support metric can misrepresent the patient when for example the patient's data has varying durations, or different data intensity. In this paper we introduce three new metrics that are normalized or extension of the horizontal support.

The first two defined in Definitions 5 and 6 are normalized versions of the HS. The first is normalized horizontally based on the patients' TIRP having the highest horizontal support, which may fit well for datasets, in which the amount of data, or the observation period, may vary. The second is normalized vertically based on the specific TIRP's maximal horizontal support value across all the patients. The classification matrix of values, in which often the rows are the patient's features' values and the columns are the features' values, orients the vertical or horizontal normalization. The third novel metric is inspired by the TFIDF from the textual domain, which was effective also in other domains, as was mentioned earlier. Here the HS is analogous to the TF and the IDF to what we call Inverse Entity Frequency. In our case the entity is a patient (which in the textual domain is a document).

Definition 6. The *horizontally normalized horizontal support* (HNHS) of a TIRP P_i for an entity e_j is the *hor_sup(P_i, e_j)* divided by the patient e_j's maximal horizontal support value across all its TIRPs.

Definition 7. The *vertically normalized horizontal support* (VNHS) of a TIRP P_i for an entity e_j is the *hor_sup(P_i, e_j)* divided by the maximal horizontal support value of the TIRP P_i across all the patients.

Definition 8. The *horizontal support inverse entity frequency* (HSIEF) of a TIRP P for an entity e_i is the result of multiplying the horizontal support of the pattern P at entity e_i (Definition 4) by the inverse entity frequency of the TIRP P, which is actually dividing by the vertical support of the TIRP P (Definition 3). As mentioned earlier, the HSIEF is analogous to the TFIDF metric from textual domain, in which the frequency of a term in the document is considered, as well as its frequency within the entire library.

$$\begin{aligned} &Horizontal\ Ssupport\ Inverse\ Entity\ Frequency = HS(P,e_i)\cdot IEF(P) \\ &= HS(P,e_i)\cdot\frac{1}{VS(P)} \end{aligned}$$

3.4. Maitreya - prediction of outcome events

In this paper, we introduce the Maitreya framework for the prediction of outcomes using time intervals mining. Fig. 3 illustrates the framework workflow, including its evaluation setup. In this setup we wanted to have a model that consists only on the Cohort data, while defining a controls group that is used only for evaluation and not the creation of the model. Additionally, since in TIRPs mining the discovered TIRPs may vary according to the mined folds, we wanted to have an additional layer of cross validation in addition to the classification cross validation. Our methodology consists of discovering TIRPs only in the set of patients having the outcome. Thus, the TIRPs are discovered and the prediction model is induced based only on the outcome cohort patients' frequent TIRPs. A control set of patients that do not have the outcome events in their records is created for the evaluation, as we describe in details in the section about the datasets that were used for the evaluation.

Since in TIRPs mining in each fold a different set of frequent TIRPs can be discovered, and to avoid overfitting, we designed a rigorous evaluation strategy, in which the cohort and control datasets are divided into three folds, as shown in Fig. 4. In each iteration TIRPs are discovered from a third of the cohort (one fold), and then SingleKarmaLego is used to detect the TIRPs that were discovered in the other two thirds (two folds) of the cohort and of the controls respectively. Once the TIRPs' instances were detected using SingleKarmaLego, a matrix of TIRPs-features is created. Note, the TIRPs are discovered from a distinct cohort population than the patients used for the classification evaluation. We used ten-fold cross validation for the classification evaluation.

After we apply SingleKarmaLego to the cohort and control, in order to detect the TIRPs (that were discovered by KarmaLego in the first fold) instances in the patient data, a matrix is constructed. The rows of this matrix are the patients and the columns are the

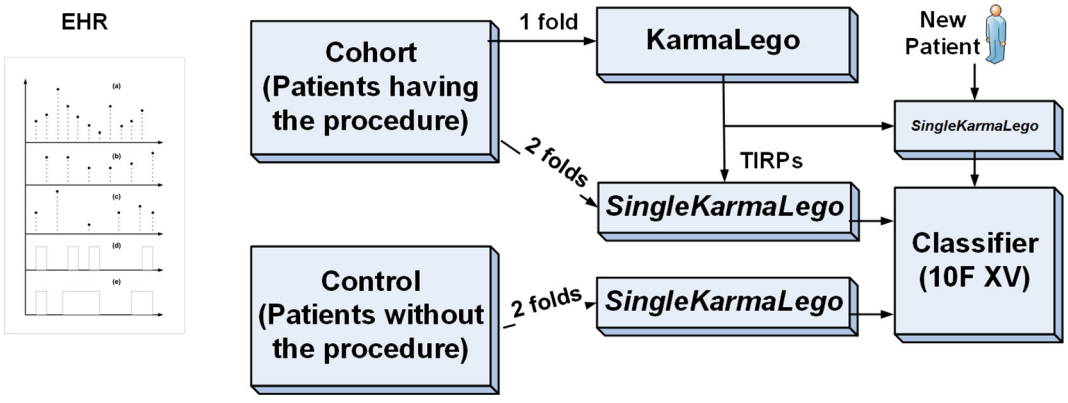

Fig. 4. The workflow of the Maitreya framework for outcomes prediction. A cohort of patient having the outcome procedure is created. TIRPs for the model are discovered by KarmaLego only from the Cohort training folds. A control group patients that do not have the outcome procedure are used for evaluation only (not for TIRPs discovery). Using SingleKarmaLego the TIRPs that were discovered for the model are detected in the testing folds of both the Cohort and the Controls, which is given to a classifier that learns a model and evaluates it using 10 folds cross validation classification is performed.

TIRP features, containing the values of the various metrics. Thus, each matrix entry contains a value, which in default is Boolean, meaning whether the given TIRP instances were detected for the specific patient or not. Additionally, we also implemented the Horizontal Support (Definition 4) and Mean Duration (Definition 5) representations, as well as the three novel metrics that are the normalized versions of the Horizontal Support: *Horizontally Normalized Horizontal Support* (Definition 6) and *Vertically Normalized Horizontal Support* (Definition 7), as well as the *Horizontal Support Inverse Entity Frequency* (Definition 8).

4. Evaluation

4.1. Datasets

We extracted clinical data from the New York Presbyterian Hospital - Columbia University Medical Center (NYP-CUMC) clinical data warehouse. In total, the NYP-CUMC EHR contains medical record data for approximately 4.5 million patients going back to 1989, containing approximately 30 million diagnosis billing codes, 20 million prescription orders, 9 million procedures, and 500 million laboratory results. We used only coded data for this analysis, including drug exposures, conditions billing codes), and procedures. As part of the migration of the data into the Observational Medical Outcomes Partnership [38] these concepts were mapped to RxNorm, SNOMED-CT, and ICD-9-Procedure, respectively, to conform to the Common Data Model. Thus, Medical concepts were transformed into symbolic time intervals (called "eras" in OMOP) and were further concatenated according to Definition 9. According to the OMOP standards, as we define officially in Definition 9 each two time intervals ("eras") having the same concept id (symbol) that follow each other, having a time gap period that is less than a given time period threshold of time are concatenated into a single time interval having the same symbol that starts at the first time interval start-time and ends at the end-time of the second time interval.

Definition 9. Abstraction function: Given two symbolic time intervals having the same symbol I_i and I_j, and I_i is before I_j according to the lexicographical order, and $I_{js} - I_{ie} < max\_time$ holds, it will be abstracted into a new single time interval having the start time of I_i and the end time of I_j. In our case *max_time* was set to 30 days, according to the OMOP standard.

$$\forall I^i,\ I^j \in TIS \wedge (i<j) \wedge (I^i_{sym} = I^j_{sym}) \wedge (I^j_s - I^i_e < \max_{time}\text{time}) : TIS \leftarrow TIS - I^i \wedge TIS \leftarrow TIS - I^j \wedge newI \cdot s = I^i_S \wedge newI \cdot e = I^j_e \wedge TIS \leftarrow TIS + newI$$

However, since we used patients that have at least two out of three event types (for example, conditions and drug exposures) we had in fact more than hundred thousands patients. Then a cohort of patients for each predicted procedure was created. The cohorts were created by selecting the patients that have at least one procedure in their records. We went back in time for the prediction period of two months, and then went back two years of observation time. The patient's data along the observation time were used for the cohort. Each patient that had events spread along more than a year within the two years was added to the cohort. For each procedure that had more than five hundred patients, we created a control group. The controls included patients that were randomly selected among the rest of the patients in the database who have not the said procedure in their records, but have another procedure in their records.

Once a procedure was detected, we prepared its data in the same way as with the cohort given the prediction and observation time. Thus, as illustrated in Fig. 1, we went backwards in time the prediction-time period before the procedure, and then used the data that was found in the observation-time period prior to the beginning of the prediction-time period. Thus, the data that was used for the prediction was taken from the observation-time that ends at the beginning of the prediction period, prior to the outcome event. Finally, from the most common procedures available in our dataset, we excluded procedure codes that were uninformative such as "Other interview and evaluation". We also excluded codes that describe small routine procedures for specimen collection such as "Routine venipuncture or finger/heel/ear stick for collection of specimen(s)". Since our method uses the laboratory test results to identify temporal patterns, such procedure code would not add information to the described method. We performed our analysis based on a list of 28 procedures codes that represented meaningful clinical conditions. The resulted datasets, that are listed in Table 1, contained hundreds of distinct symbols, which were up to two thousand in specific datasets, and in average the number of symbolic time intervals per patient were around seventy. Table 1 presents the entire list of the procedures datasets, including the cohort and controls that were used for the prediction experiments. Each row presents the procedure name, the size of the dataset, and the percentages of the cohort and controls of races, and their diagnoses (based on the codes in the database).

4.2. Research questions and experimental plan

For the evaluation of the prediction of outcome events on the 28 procedures that were described in Table 1, we defined several research questions. We wanted to answer the following questions: (1) whether using TIRPs as features is better than symbols (without considering their temporal order); (2) whether the TIRPs representation will be better than the default Boolean that were used in previous studies. We wanted to clarify these questions for each procedure. However, before that we wanted to answer two questions regarding the optimal settings of the Maitreya framework for the task of procedures prediction, namely (3) the set of temporal relations (Allen's seven relations, or our three more general relations) and (4) the size of the epsilon (which enables to make Allen's temporal relations more flexible, as defined in Definition 1 and Fig. 2).

To answer these questions, we applied Maitreya on the entire list of twenty-eight procedures using two years of observation time and two months of prediction time (Fig. 1) with *three* versus *seven* temporal relations, and we set epsilon to 0, 30, 60, 90 and 180 days. We used 20% minimal vertical support for all the procedures (datasets) and maximal gap (Fig. 2) of 730 days. All the experiments were performed with three folds of mining cross validation, as was explained in Section 3.4 and ten folds of classification cross validation using weka 3.6.11. Thus, each result that is reported is based on thirty experimental runs.

5. Results

5.1. Temporal relations and epsilon values

We ran Maitreya for each procedure using the three abstract and Allen's seven temporal relations, with the following epsilon values: 0, 30, 60, 90 and 180 days. In Fig. 5 we present the mean results for each epsilon value, in comparison to the use of symbols (with no temporal representation). We calculated the means based on the top twelve procedures (shown in Table 2). We present here the results when using the Boolean, HS and MND metrics for the TIRPs, and the mean of the maximal differences between the TIRPs versus the symbols (Sym) based prediction. Fig. 5a shows the mean results of using TIRPs with binary (whether a TIRP instance was detected, or not) representation.

Table 1
The outcome clinical procedures datasets.

Procedure name	Dataset size	Female%	White%	Other%	Asian%	Black%	Congestive heart failure%	Diabetes%	High cholesterol%
Electrocardiogram	1046	0.46	0.401	38.4	1.3	18.9	29.8	30.2	36.5
		0.581	0.414	42	9	14.5	21.2	28.1	35.1
Platelets transfusion	1116	0.503	0.577	25.8	0.8	15.2	26.8	26.3	25.2
		0.566	0.41	43.1	1.4	13.7	18.9	25	34.4
Steroid injection	1322	0.546	0.44	41.6	0.6	12.7	22.9	24.5	30.4
		0.586	0.438	41.6	0.6	13.1	19.6	26.4	35.2
Urinary tract ultrasound	1570	0.505	0.448	35.5	0.7	18.3	30.7	34.1	37.4
		0.579	0.424	42.9	1.7	12.1	17.8	24.5	32.8
Insulin injection	1738	0.571	0.338	52.3	1	11.9	34.1	56.5	50.2
		0.578	0.425	42	1.1	13.3	17.3	21	31.6
Occupational therapy	1174	0.565	0.504	37.9	0.3	10.7	27.2	33.7	41.9
		0.565	0.456	41.9	1	11	19.9	24.3	31.3
Anticoagulant injection	2360	0.576	0.3	57.2	0.5	11.1	26.3	33.1	39.4
		0.58	0.433	0.422	1	12.7	17.1	24	31.1
Chemotherapeutic administration	1142	0.56	0.588	0.253	1.4	14.1	16.4	22.2	25.2
		0.57	0.399	0.453	1.2	12.9	19.7	27.6	33.9
Endoscopy of small intestine	1596	0.561	0.469	0.362	1	14.9	27.9	31.8	34
		0.58	0.416	0.427	0.7	14	14.2	22	30.5
Invasive mechanical ventilation for >=96 h	1522	0.49	0.546	0.24	1.3	18.6	31.5	28.7	30.8
		0.584	0.374	0.483	1.7	11.5	18	25.3	31.2
Invasive mechanical ventilation for <96 h	1718	0.525	0.484	0.316	1.1	17.9	30.2	29.2	32.9
		0.61	0.398	0.471	1.5	10.8	18.9	27.7	35.1
Thoracic CT	2478	0.535	0.443	0.381	1.5	15.1	25.7	28.4	33
		0.587	0.402	0.433	1.1	13.9	15.8	25.4	33
Abdomen and retroperitoneum ultrasound	2358	0.565	0.418	0.416	1.1	14.6	24	29.2	32.4
		0.569	0.441	0.428	0.7	11.2	17.3	23.7	32.4
Peripheral vascular ultrasound	3168	0.558	0.436	0.393	1.1	15.1	30.3	32.3	35.6
		0.559	0.419	0.421	0.9	14.1	15.8	22.8	31.1
Spinal tap	1178	0.561	0.456	0.356	0.6	17.1	22.7	28.5	30.7
		0.578	0.41	0.441	2	11.7	18.1	23.2	32.7
Non-invasive mechanical ventilation	1356	0.523	0.436	0.415	1	12.9	33	31.8	36.1
		0.572	0.414	0.435	0.4	13.7	19.4	27.1	33.3
Cataract extraction surgery with intraocular lens insertion	1032	0.608	0.352	0.565	0.9	07.1	34.4	45.7	59.4
		0.575	0.405	0.428	1.1	14.1	19.3	24.8	32.7
Physical therapy	2200	0.57	0.478	0.364	0.005	14.3	29	35	4
		0.569	0.414	0.446	0.011	11.7	16	23.6	30.4
Abdominal CT	3038	0.594	0.438	0.398	0.007	14.9	23.9	28.3	34.6
		0.57	0.454	0.403	0.011	12.7	18.2	23.1	30.4
Electrographic monitoring	2536	0.549	0.421	0.395	0.004	17.5	39.8	43.2	46.1
		0.583	0.402	0.451	0.013	12.3	14.6	21.9	31.5
Large intestine endoscopic biopsy	1642	0.584	0.401	0.459	0.006	12.7	23.6	32	39
		0.56	0.406	0.427	0.006	14.7	18.7	25	32.7
Endotracheal tube	2654	0.513	0.527	0.271	0.012	17.7	31.1	30.6	32.3
		0.574	0.407	0.452	0.008	12.1	19	28.4	36.8
Brain and brain stem MRI	2108	0.595	0.468	0.352	0.01	0.159	24.3	31.3	35.4
		0.573	0.409	0.436	0.009	0.138	17.5	24.9	32.8
Head CT	3464	0.575	0.446	0.386	0.008	0.149	29.1	34	39.3
		0.577	0.433	0.419	0.01	0.127	15.1	21.1	29.6
Colonoscopy	2096	0.548	0.34	0.537	0.009	0.107	27.6	35.4	45.8
		0.575	0.433	0.41	0.006	0.137	19.5	24.2	29.5
Packed cells transfusion	3842	0.537	0.505	0.312	0.01	0.166	27.2	25.7	3
		0.603	0.388	0.477	0.007	0.121	17.4	26.4	35.1
Coronary arteriography	1936	0.403	0.479	0.365	0.011	0.135	32.1	30.7	38.7
		0.6	0.378	0.467	0.007	0.139	13.7	24.3	30.1
Cardiac ultrasound	3898	0.544	0.424	0.405	0.009	0.154	32.4	32.8	39
		0.616	0.419	0.445	0.006	0.121	10.9	21.9	28.7
Total		0.587	0.435	0.414	0.011	0.119	9.9	15.5	23.1

These results show that while the performance did not change much for the varying epsilon values, using epsilon of thirty days, especially with Allen's original seven temporal relations. Similarly, Fig. 5b shows the results of using horizontal support to represent the TIRPs, representing how many instances of a TIRP were detected at the patient'. Here, epsilon of 30 and 60 days was slightly better than the use of zero, but not significantly, as can be seen in Figs. 5 using the confidence interval bars. However, above epsilon = 60 the mean performance decreased. Overall the use of three and seven temporal relations preformed similarly. Fig. 5c shows the results when using the mean duration to represent the TIRPs and here the use of seven relations performed overall slightly better than three relations. Above epsilon = 60 the performance with three temporal relations dropped quite meaningfully. Fig. 5d shows the mean results of the maximal differences among the best performing TIRP representation and the symbols.

Since the results in Fig. 5d are the mean of the best performing TIRP metrics for each we couldn't calculate their confidence intervals. Overall the use of three temporal relations is better, especially with the maximal difference (Fig. 5d), and epsilon with thirty days performs slightly better than the use of zero. Having an epsilon with sixty days performed similar to zero, while larger epsilon

Table 2
The prediction results of the twenty-eight procedures, including the baseline Symbols, the TIRPs with the six representation metrics of the Bianry (BIN), the Horizontal Support (HS) and its extensions the Horizontally Normalized Horizontal Support (HNHS), the Vertically Normalized Horizontal Support (VNHS) and Horizontal Support Inverse Entity Frequency (HSIEF), as well as the Mean Duration (MND), and the subtracted results of the TIRPs metrics from the Symbols baseline.

Procedure name	Symbols (mean −/+ 95%ci)	TIRPs (mean −/+ 95% ci)						Subtracted from the Symbols					
		BIN	HS	HNHS	VNHS	HSIEF	MND	BIN	HS	HNHS	VNHS	HSIEF	MND
Endotracheal tube	0.51 −/+ 0.07	0.66 −/+ 0.05	0.63 −/+ 0.06	0.58 −/+ 0.08	0.63 −/+ 0.06	0.67 −/+ 0.08	0.66 −/+ 0.07	15.58	12.25	7.66	12.6	**16.39**	15.2
Occupational therapy	0.53 −/+ 0.08	0.61 −/+ 0.08	0.65 −/+ 0.07	0.61 −/+ 0.07	0.65 −/+ 0.07	0.67 −/+ 0.07	0.64 −/+ 0.07	7.74	12.08	7.74	12.39	**14.09**	10.59
Thoracic CT	0.60 −/+ 0.06	0.61 −/+ 0.07	0.68 −/+ 0.07	0.62 −/+ 0.06	0.65 −/+ 0.08	0.71 −/+ 0.07	0.59 −/+ 0.07	1.20	7.95	2.12	5.66	**10.93**	−0.81
Large intestine endoscopic biopsy	0.46 −/+ 0.08	0.45 −/+ 0.07	0.44 −/+ 0.07	0.41 −/+ 0.09	0.46 −/+ 0.08	0.50 −/+ 0.08	0.56 −/+ 0.09	−0.69	−1.66	−4.58	0.41	4.02	**9.86**
Brain and brain stem MRI	0.61 −/+ 0.08	0.64 −/+ 0.07	0.67 −/+ 0.08	0.63 −/+ 0.08	0.66 −/+ 0.06	0.64 −/+ 0.06	0.71 −/+ 0.07	2.18	5.31	1.73	4.58	2.15	**9.44**
Platelet transfusion	0.63 −/+ 0.07	0.71 −/+ 0.07	0.68 −/+ 0.07	0.70 −/+ 0.06	0.68 −/+ 0.08	0.71 −/+ 0.08	0.66 −/+ 0.08	**8.64**	5.64	7.64	5.43	8	3.10
Cardiac ultrasound	0.57 −/+ 0.07	0.57 −/+ 0.09	0.57 −/+ 0.09	0.57 −/+ 0.07	0.60 −/+ 0.09	0.57 −/+ 0.08	0.66 −/+ 0.07	−0.13	−0.06	−0.45	3.02	−0.31	**8.43**
Anticoagulant injection	0.65 −/+ 0.06	0.72 −/+ 0.06	0.69 −/+ 0.06	0.72 −/+ 0.07	0.73 −/+ 0.07	0.74 −/+ 0.06	0.74 −/+ 0.06	6.97	4.06	6.35	7.83	**8.39**	8.33
Steroid injection	0.65 −/+ 0.05	0.67 −/+ 0.07	0.72 −/+ 0.06	0.68 −/+ 0.06	0.73 −/+ 0.07	0.71 −/+ 0.07	0.67 −/+ 0.07	2.58	7.08	3.25	**8.18**	6.14	2.79
Electrographic monitoring	0.64 −/+ 0.07	0.68 −/+ 0.08	0.72 −/+ 0.07	0.70 −/+ 0.07	0.71 −/+ 0.07	0.66 −/+ 0.08	0.69 −/+ 0.08	4.09	**8.02**	6.63	7.11	2.04	5.76
VAD insertion	0.65 −/+ 0.08	0.7 −/+ 0.06	0.69 −/+ 0.07	0.64 −/+ 0.07	0.71 −/+ 0.06	0.73 −/+ 0.06	0.66 −/+ 0.06	4.33	3.85	−1.06	6.12	**7.54**	0.29
Endoscopy of small intestine	0.59 −/+ 0.06	0.63 −/+ 0.06	0.63 −/+ 0.07	0.56 −/+ 0.08	0.66 −/+ 0.08	0.58 −/+ 0.07	0.66 −/+ 0.07	3.81	3.71	−3.43	6.11	−1.63	**6.87**
Chemotherapy administration	0.59 −/+ 0.09	0.62 −/+ 0.09	0.65 −/+ 0.08	0.61 −/+ 0.09	0.65 −/+ 0.09	0.65 −/+ 0.09	0.58 −/+ 0.10	3.75	6.20	1.94	**6.57**	6.29	−0.55
Invasive mechanical ventilation for >=96 h	0.67 −/+ 0.06	0.73 −/+ 0.06	0.68 −/+ 0.06	0.67 −/+ 0.07	0.69 −/+ 0.05	0.71 −/+ 0.06	0.73 −/+ 0.06	**6.06**	1.20	0.35	1.45	3.45	5.52
Peripheral Vascular Ultrasound	0.60 −/+ 0.07	0.64 −/+ 0.08	0.63 −/+ 0.07	0.59 −/+ 0.08	0.64 −/+ 0.07	0.64 −/+ 0.08	0.64 −/+ 0.07	3.27	2.91	−1.43	3.20	3.75	**3.81**
Abdomen and Retroperitoneum Ultrasound	0.63 −/+ 0.05	0.55 −/+ 0.06	0.66 −/+ 0.06	0.58 −/+ 0.06	0.65 −/+ 0.07	0.65 −/+ 0.06	0.66 −/+ 0.07	−7.83	**3.31**	−4.27	2.43	2.29	3.14
Head CT	0.61 −/+ 0.07	0.62 −/+ 0.09	0.61 −/+ 0.09	0.58 −/+ 0.09	0.65 −/+ 0.08	0.59 −/+ 0.09	0.63 −/+ 0.09	0.76	−0.17	−3.22	**3.29**	−1.8	1.42
Invasive mechanical ventilation for <96 h	0.68 −/+ 0.06	0.64 −/+ 0.06	0.69 −/+ 0.06	0.69 −/+ 0.05	0.70 −/+ 0.06	0.65 −/+ 0.06	0.71 −/+ 0.06	−3.49	0.83	0.83	1.83	−3.41	**2.91**
Abdominal CT	0.66 −/+ 0.06	0.64 −/+ 0.07	0.65 −/+ 0.07	0.61 −/+ 0.08	0.65 −/+ 0.08	0.69 −/+ 0.07	0.66 −/+ 0.06	−1.93	−0.75	−4.72	−1.2	**2.60**	−0.02
Non-invasive mechanical ventilation	0.68 −/+ 0.06	0.67 −/+ 0.06	0.66 −/+ 0.07	0.67 −/+ 0.07	0.68 −/+ 0.07	0.68 −/+ 0.07	0.70 −/+ 0.07	−0.58	−1.81	−0.93	0.58	0.14	**2.39**
Cataract extraction surgery with intraocular lens insertion	0.73 −/+ 0.05	0.71 −/+ 0.07	0.71 −/+ 0.07	0.70 −/+ 0.08	0.73 −/+ 0.08	0.71 −/+ 0.07	0.74 −/+ 0.07	−1.75	−2.26	−3	−0.27	−1.34	**0.74**
Colonoscopy	0.61 −/+ 0.08	0.56 −/+ 0.08	0.56 −/+ 0.09	0.61 −/+ 0.08	0.55 −/+ 0.09	0.55 −/+ 0.10	0.61 −/+ 0.08	−4.76	−5.23	**0.277**	−6.11	−6.43	−0.37
Coronary arteriography	0.77 −/+ 0.05	0.76 −/+ 0.06	0.74 −/+ 0.07	0.75 −/+ 0.07	0.76 −/+ 0.07	0.74 −/+ 0.07	0.74 −/+ 0.06	−1.02	−3.37	−2.41	−0.68	−2.89	−2.95
Urinary Tract Ultrasound	0.77 −/+ 0.05	0.72 −/+ 0.06	0.76 −/+ 0.06	0.73 −/+ 0.06	0.75 −/+ 0.06	0.74 −/+ 0.06	0.72 −/+ 0.07	−4.83	−0.97	−3.83	−1.91	−2.74	−4.49
Electrocardiogram	0.67 −/+ 0.06	0.63 −/+ 0.07	0.64 −/+ 0.07	0.65 −/+ 0.07	0.63 −/+ 0.07	0.63 −/+ 0.07	0.65 −/+ 0.07	−3.33	−2.79	−1.75	−3.62	−3.35	−1.58
Insulin injection	0.81 −/+ 0.06	0.72 −/+ 0.07	0.73 −/+ 0.06	0.77 −/+ 0.06	0.70 −/+ 0.07	0.70 −/+ 0.07	0.74 −/+ 0.07	−8.95	−8.5	−4.2	−10.9	−11.1	−7.08
Spinal tap	0.63 −/+ 0.05	0.58 −/+ 0.07	0.55 −/+ 0.07	0.57 −/+ 0.07	0.53 −/+ 0.07	0.52 −/+ 0.07	0.55 −/+ 0.08	−4.37	−7.5	−5.97	−9.61	−10.9	−7.91
Packed cell transfusion	0.67 −/+ 0.07	0.60 −/+ 0.08	0.58 −/+ 0.07	0.62 −/+ 0.08	0.58 −/+ 0.07	0.58 −/+ 0.07	0.57 −/+ 0.09	−7.32	−9.47	−5.31	−9.47	−9.58	−10
Mean	0.64 −/+ 0.06	0.64 −/+ 0.07	0.65 −/+ 0.07	0.64 −/+ 0.07	0.66 −/+ 0.07	0.65 −/+ 0.07	0.66 −/+ 0.07	0.516	1.334	−0.316	1.837	1.387	1.841

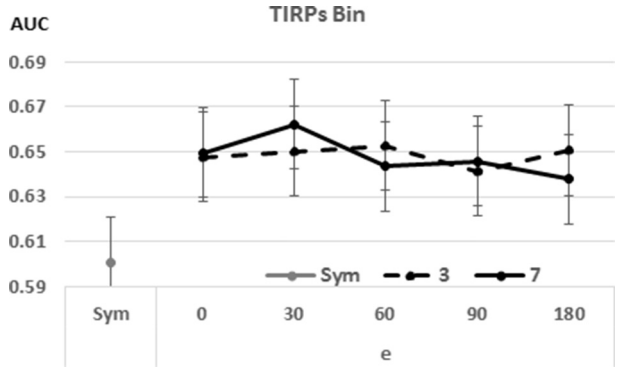

Fig. 5a. The mean results with the Bin representation, in which using epsilon zero worked better.

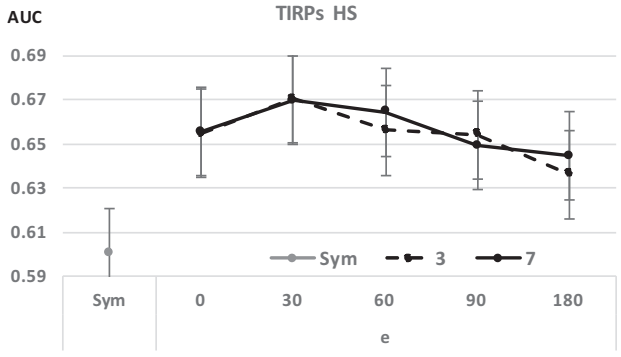

Fig. 5b. The mean results with the horizontal support representation, in which the epsilon and the number of relations seem to perform the same.

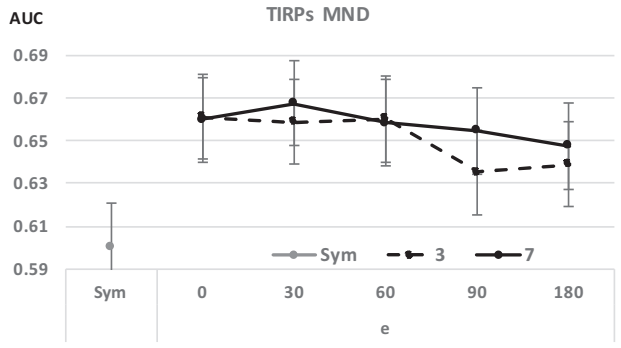

Fig. 5c. The mean results of the Mean Duration representation, in which for the 7 relations the use of larger epsilon decreased the performance.

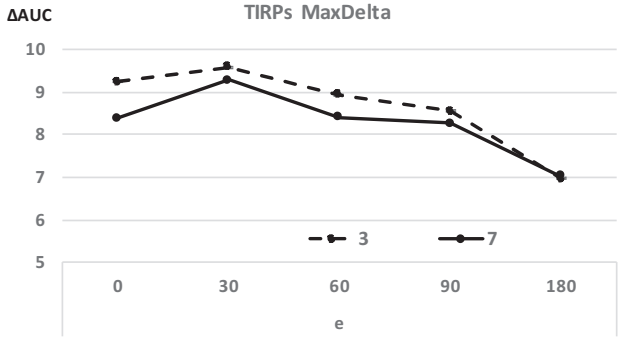

Fig. 5d. The mean results of the maximal difference, in which the use of 7 relations performed better than the 3, especially with epsilon zero.

was less effective and performed worse than using zero. For the TIRPs metrics mean results shown in Figs. 5a–5c there were no meaningful differences when using epsilon larger than zero (which was the reason we used zero at the next experiment for the entire set of procedures), but with the best performing TIRP metric, which we defined as maximal difference the use of 30 days epsilon mean behaviour seems quite meaningful.

5.2. Procedures prediction

After evaluating the varying settings of the Maitreya framework we set the temporal relations to the three broader relations and the epsilon to zero. Although the results of the epsilon set to 30 days were better (mainly for the metrics in which there was the best performance as shown in Fig. 5d), since it was not meaningfully better, we present the results of the epsilon with zero. In Table 2 we report the results of our experiments for each of the procedures presented in Table 1. The table reports the results for the various representations, whether without temporal or with temporal representation. Thus, starting with the use of symbols (no temporal representation, in which the predictors were the concepts) and then the use of TIRPs with the six metrics, including the three novel metrics that are extensions of the horizontal support: *binary* (BIN), which represent whether the a TIRP was found for a patient or not, *horizontal support* (HS), which is the number of occurrences a TIRP instance was detected for a patient, *horizontally normalized horizontal support* (HNHS), which is aa normalized version of the HS based on the TIRP of a patient having the highest HS value, *vertically normalized horizontal support* (VNHS), which is similar but normalized based on the patients values of a TIRP having the highest value, *horizontal support inverse entity frequency* (HSIEF), which is metrics inspired by the TFIDF, and *mean duration* (MND), which is the average of the TIRP instances length in time. The next part (on the right) in the table is the subtraction of each TIRP metric performance from the symbols performance, in which the best result is in bold.

The results report our prediction experiments of the twenty-eight procedures that are presented in Table 2. The results report the baseline of using only Symbols (without any temporal analysis), and the use of TIRPs with several representations.

The first is the baseline of Boolean that was used by [34,5], and the rest are more expressive and refer to the instances of the TIRPs for each patient. Overall, for 22 of 28 procedures the use of the TIRPs compared to the atemporal features was more effective. The use of the richer TIRP metrics was better than the use of the Boolean representation for TIRPs that was used in past studies [34,5]. The procedures in Table 2 are ordered based on the maximal difference of the TIRPs use compared to the Symbols baseline. As can be seen in the bottom of the table the vertically normalized HS and the HSIEF performed better than the HS and similar to the mean duration, which was the best with the VNHS.

We grouped the procedures into two types of therapeutic and diagnostic and several purposes, such as *imaging,* drug infusion or transfusion, ventilation procedures, endoscopy, in order to compare the advantages of using the TIRPs compared to the baseline within these groupings, but in our analysis we couldn't find any meaningful differences among these groupings.

6. Discussion

The increasing availability and accumulation of longitudinal data in electronic health records provides many opportunities in biomedical research, especially in prediction of outcomes. However, the use of these data comes with many challenges, including sparseness and variability of temporal variables and their representation along time whether as time point values, or as time inter-

vals events having duration. In this paper, we described Maitreya, a framework for the prediction of clinical procedures. Maitreya uses an enhanced version of KarmaLego that enables its use in datasets with many symbols [26], a fast TIRPs discovery algorithm, to discover TIRPs from a cohort of patients with a given outcome, such as a procedure. After a set if frequent TIRPs are discovered, they are detected using SingleKarmaLego, which constructs a matrix that describes the TIRPs' instances appearances in each patient for the classification evaluation. The TIRPs are represented using five types of TIRP representations, in addition to the previously used default Boolean representation [5]. The five metrics for representation of the TIRP instances in a patient records includes two novel normalized versions of the horizontal support, that represents the number of TIRPs instances per patient.

Maitreya allows for several sets of temporal relations to define the temporal relations among the symbolic time intervals in a TIRP. The first option is Allen's seven temporal relations, alternatively we proposed here a set of three general temporal relations that are disjunctions of the seven relations, and an epsilon value is defined to make the temporal relations more flexible. In this paper we performed experiments both with the original seven and the broader three relations, as well as with setting the epsilon for 30, 60, 90 and 180. We wanted to know which set of temporal relations will be more effective and whether setting the epsilon to more than zero will influence the results. We expected that the epsilon value will make the temporal relations framework more flexible, which is expected to be more meaningful in the seven relations and less meaningful with the three temporal relations that are already broad in their definitions. The introduction of the epsilon value intends to make the relations more flexible. This is since often in real data due to noise A meets B may be A before B, but in fact B is after A in a single time unit. Thus, the difference is meaningless and we would like to consider it still as meet and not as before. For that the epsilon value can be defined according to the specific dataset and required granularity. The distribution of the instances supporting the relations, for example in the corresponding meet and before example, it may change according to the dataset and the epsilon. In a specific dataset with epsilon = 1 it may not change at all (compared to the default Allen's relations, in which epsilon is actually zero). The same may happen with a month or more, depending on the data distribution. Thus, the choice of the epsilon can be chosen according to domain knowledge or through exploration of the data. In our dataset which contains hospitalizations, considering epsilons of dozens of days, as we report for a month, two or three, had no meaningful effect on the prediction accuracy. It might be that with data that is more intense the effect may be more meaningful, which we would like to explore in the future. Additionally, Maitreya allows for several metrics for the representations of the TIRPs in the classification task. In this paper specifically we present three new TIRPs metrics and their use in single patient representation. Two new normalized versions of the Horizontal Support metric that counts the number of TIRP instances in the observation period, horizontally and vertically, and a Horizontal Support that is multiplied by the inverse entity (i.e., patient) frequency.

We first evaluated the settings of the Maitreya framework and based on the mean results of the procedures we saw that the use of TIRPs was significantly better than the use of the Symbols (in which the temporal relations and order among different concepts in the data, such as procedures, conditions and drug exposures, are ignored), based on the analysis of the twelve procedures in which there was the highest difference.

When we increased the value of the epsilon above zero, using 30, 60, 90 and 180 days, we saw an increase only in the performance for epsilon of thirty days, but for a larger epsilon the results were worse. Even with thirty days they were not significantly better, but just a little better. We tried to figure out the reason the results for 30 days was still slightly better, and we think it may be since 30 days is a common time period for inspection, following a hospitalization or procedure and may be reflected in the data. However, since it is not meaningful it is hard to investigate and determine. The epsilon value is not part of Allen's original framework [3], probably since it was defined for natural language processing and for temporal reasoning. However, the purpose was to define a more flexible framework of temporal relations, but introducing the epsilon value that is expected to be mostly effective at the equal end points, while being exclusive. By equal end points we refer to the temporal relations, in which the end points have the same time. For example, for the temporal relation *meet* the first time interval end time and the second time interval start time are identical, and the same happens with *starts*, in which both time intervals' start times are the same. For the other end points, which are not identical the epsilon effect is less meaningful (for example, the relation overlap has no end points that are the same, and as such the epsilon is less important). However, the epsilon was introduced in order to make the relations having the same end point more flexible to enable to discover more frequent TIRPs. Nevertheless, in this study the use of larger epsilon was slightly better with 30 days for the mean results, while for specific procedure it may have been also more effective with other epsilon values, but not with other values. We think that it happens since it is inpatient data and the time gaps between hospitalizations may be larger. Additionally, the slightly higher performance with 30 days may be since re-hospitalization may occur in such period of time and it is a meaningful time period for post hospitalization inspections.

The results for each procedure were reported in Table 2. Our results show that for most of the procedures the use of TIRPs performed better than the use of no temporal representation that we called symbols, and for some significantly better. Using the mean duration in average had the best performance and also the difference in the performance compared to the symbols baseline was mainly achieved by the MND. The new normalized versions of the horizontal support propose normalization vertically (the values of the same TIRP divided by the highest) and the horizontal, in which the values are divided by the TIRPs highest value for each patient, which was less effective in these experiments than the vertically normalized. The HSIEF metric performed in average better than the horizontal support metric and its normalized versions, and not as good as the mean duration, but the procedures that had the highest maximal differences compared to the symbols was with the HSIEF metric. Nevertheless, looking at the mean results the mean duration seems to perform best.

7. Conclusions

Based on the current use of TIRPs that were discovered from the conditions, procedures and drug exopsurs, we couldn't see any meaningful differences, but we expect to see such in our future work when we will add lab tests as part of the predictive data, in addition to the symbolic data that were used in this study (conditions, procedures and drug exposures). With the lab tests values that will have to go through temporal abstraction in order to be represented as symbolic time intervals data, we expect them to be more correlative and thus predictive of the outcomes.

We conclude that the use of TIRPs is more effective than the use Symbols for procedures prediction in EHR, even with only symbolic data. The use of epsilon as a value that is larger than the default zero seems not that meaningful, and this may be different with outpatient data. In this study we had inpatient data that has hospitalization data that may have big time gaps among them, which may be the result for the epsilon being not that meaningful here.

The use of more advanced metrics of the horizontal support was indeed better but not that meaningful, and the mean duration seems to be most effective yet. We plan to add lab-tests, which are time points data, and evaluate their influence on prediction performance. However, laboratory data are complex – represented by categorical, ordinal, and continuous values – and will require the use of temporal abstraction methods [42,23,28,29]. Typically, in EHR records there are many lab tests and much more than symbolic data, thus we expect also to discover more frequent TIRPs, which is expected to require more advanced TIRPs selection method that we are plan to develop.

Acknowledgements

RM and NPT are supported by funding from NIGMS R01GM107145. The authors wish to thank Yuval Shahar, George Hripcsak, Carlo Combi, Christos Faloutsos, Christian Freksa, Fabian Moerchen, Dhaval Patel and Iyad Batel for insightful discussions about temporal data analysis and prediction in EHR data and Alexandra Jacunski for editorial review.

Conflict of interests

The authors have no conflict of interests to report.

References

[1] D. Albers, N. Elhadad, E. Tabak, A. Perotte, G. Hripcsak, Dynamical phenotyping: using temporal analysis of clinically collected physiologic data to stratify populations, PLoS One 9 (6) (2014).
[2] K.P. Adlassnig, C. Combi, A.K. Das, E.T. Keravnou, G. Pozzi, Temporal representation and reasoning in medicine: research directions and challenges, Artif. Intell. Med. 38 (2006) 101–113.
[3] J.F. Allen, Maintaining knowledge about temporal intervals, Commun. ACM 26 (11) (1983) 832–843.
[4] I. Batal, H. Valizadegan, G. Cooper, M. Hauskrecht, A temporal pattern mining approach for classifying electronic health record data, ACM Trans. Intell. Syst. Technol. 4 (4) (2013).
[5] I. Batal, D. Fradkin, J. Harrison, F. Moerchen, M. Hauskrecht, Mining recent temporal patterns for event detection in multivariate time series data, in: Proceedings of Knowledge Discovery and Data Mining (KDD), Beijing, China, 2012.
[6] R. Bellazzi, B. Zupan, Predictive data mining in clinical medicine: current issues and guidelines, Int. J. Med. Inf. 77 (2) (2008) 81–97.
[7] M.R. Boland, A. Jacunski, T. Lorberbaum, J. Romano, R. Moskovitch, N.P. Tatonetti, Systems biology approaches for identifying adverse drug reactions and elucidating their underlying biological mechanisms, Wiley Interdisc. Rev.: Syst. Biol. Med. 8 (2) (2016) 104–122.
[8] Y. Chen, W. Peng, S. Lee, Mining temporal patterns in time interval-based data, IEEE Trans. Knowl. Data Eng. 27 (12) (2015) 3318–3331.
[9] Y. Chen, J.T. Weng, L. Hui, A novel algorithm for mining closed temporal patterns from interval based data, Knowl. Inf. Syst. 46 (1) (2016) 151–183.
[10] H. Cheng, X. Yan, J. Han, C. Hsu, Discriminative frequent patterns analysis for effective classification, in: International Conference on Data Engineering (ICDE), 2007.
[11] C. Combi, G. Pozzi, Temporal representation and reasoning in medicine, Artif. Intell. Med. 38 (2006) 97–100.
[12] C. Combi, M. Mantovani, A. Sabaini, P. Sala, F. Amaddeo, U. Moretti, G. Pozzi, Mining approximate temporal functional dependencies with pure temporal grouping in clinical databases, Comp. Biol. Med. 62 (2015) 306–324.
[13] C. Combi, B. Oliboni, A. Zardini, F. Zerbato, Seamless Design of Decision Intensive Care Pathways, IEEE International Conference on Health Informatics, Chicago, USA, 2016.
[14] P. Esling, C. Agon, Time series data mining, ACM Comput. Surv. 45 (1) (2012) 12.
[15] F. Höppner, Learning temporal rules from state sequences, in: Proceedings of WLTSD, 2001.
[16] G. Hripcsak, D.J. Albers, A. Perotte, Exploiting time in electronic health record correlations, J. Am. Med. Inf. Assoc. 18 (1) (2011) 109–115.
[17] G. Hripcsak, D. Albers, Next-generation phenotyping of electronic health records, J. Am. Med. Inf. Assoc. 20 (2013) 117–121.
[18] G. Hripcsak, D.J. Albers, A. Perotte, Parameterizing time in electronic health record studies, J. Am. Med. Inf. Assoc. 22 (4) (2015) 794–804.
[19] G. Hripcsak, Physics of the Medical Record: Handling Time in Health Record Studies, Artificial Intelligence in Medicine (AIME), Pavia, Italy, 2015.
[20] Huang, W. Dong, P. Bath, L. Ji, H. Duan, On mining latent treatment patterns from electronic medical records, Data Min. Knowl. Discov. 29 (2015) 914–949.
[21] Lin Hui, Yi-Cheng Chen, Julia Tzu-Ya Weng, Suh-Yin Lee, Incremental mining of temporal patterns in interval-based database, Knowl. Inf. Syst. 46 (2) (2016) 423–448.
[22] P.B. Jensen, L.J. Jensen, S. Brunak, Mining electronic health records: towards better research applications and clinical care, Nat. Rev. Genet. 13 (6) (2012) 395–405.
[23] J. Lin, E. Keogh, S. Lonardi, B. Chiu, A symbolic representation of time series with implications for streaming algorithms, in 8th ACM SIGMOD DMKD Workshop, 2003.
[24] F. Mörchen, A. Ultsch, Optimizing time series discretization for knowledge discovery, in Proceeding of SIG-KDD, 2005.
[25] F. Mörchen, D. Fradkin, Robust mining of time intervals with semi-interval partial order patterns, in: Proceedings of SIAM Data Mining, 2010.
[26] R. Moskovitch, C. Walsh, F. Wang, G. Hripsack, N. Tatonetti, Outcomes prediction via time intervals related patterns, in: IEEE International Conference on Data Mining (ICDM), Atlantic City, USA, 2015.
[27] R. Moskovitch, Y. Shahar, Fast time intervals mining using transitivity of temporal relations, Knowl. Inf. Syst. 42 (1) (2015) 21–48.
[28] R. Moskovitch, Y. Shahar, Classification of multivariate time series via temporal abstraction and time intervals mining, Knowl. Inf. Syst. 45 (1) (2015) 35–74.
[29] R. Moskovitch, Y. Shahar, Classification driven temporal discretization of multivariate time series, Data Min. Knowl. Discov. 29 (4) (2015) 871–913.
[30] R. Moskovitch, H. Choi, G. Hripsack, N. Tatonetti, Prognosis of clinical procedures with temporal patterns and one class feature selection, ACM/IEEE Trans. Comput. Biol. Bioinf. 14 (3) (2016).
[31] K. Ng, A. Ghoting, S.R. Steinhubl, W.F. Stewart, B. Malin, J. Sun, PARAMO: a PARAllel predictive MOdeling platform for healthcare analytic research using electronic health records, J. Biomed. Inf. 48 (2014) 160–170.
[32] K. Orphanou, A. Dagliati, L. Sacchi, A. Stassopoulou, E. Keravnou, R. Bellazzi, Combining Naïve Bayes classifiers with temporal association rules for coronary heart disease diagnosis, in: IEEE International Conference on Health Informatics, Chicago, US, 2016.
[33] P. Papapetrou, G. Kollios, S. Sclaroff, D. Gunopulos, Mining frequent arrangements of temporal intervals, Knowl. Inf. Syst. 21 (2009) 133.
[34] D. Patel, W. Hsu, M. Lee, Mining relationships among interval-based events for classification, in: Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data, 2008.
[35] A. Perotte, G. Hripcsak, Temporal properties of diagnoses code time series in aggregate, IEEE Trans. Biomed. Health Inf. 17 (2) (2013) 477–483.
[36] R. Pivovarov, D. Albers, J. Sepulveda, G. Hripcsak, N. Elhadad, Temporal trends of hemoglobin A1c testing, J. Am. Med. Inf. Assoc. 21 (2014) 1038–1044.
[37] R. Pivovarov, A. Perotte, E. Grave, J. Angiolillo, C. Wiggins, N. Elhadad, Learning probabilistic phenotypes from heterogeneous EHR data, J. Biomed. Inf. 58 (2015) 156–165.
[38] S.J. Reisinger, P.B. Ryan, D.J. O'Hara, G.E. Powell, J.L. Painter, E.N. Pattishall, J.A. Morris, Development and evaluation of a common data model enabling active drug safety surveillance using disparate healthcare databases, J. Am. Med. Inf. Assoc. 17 (6) (2010) 652–662.
[39] L. Sacchi, C. Larizza, C. Combi, R. Bellazzi, Data mining with temporal abstractions: learning rules from time series, Data Min. Knowl. Discov. 15 (2) (2007) 217–247.
[40] P. Sala, C. Combi, M. Cuccato, A. Galvani, A. Sabaini, A Framework for mining evolution rules and its application to the clinical domain, in: IEEE International Conference on Health Informatics, 2015, pp. 293–302.
[41] G. Salton, MJ. McGill, Introduction to Modern Retrieval, New York, McGraw-Hill Book Company, 1983.
[42] Y. Shahar, A framework for knowledge-based temporal abstraction, Artif. Intell. 90 (1–2) (1997) 79–133.
[43] A. Shknevsky, R. Moskovitch, Y. Shahar, Semantic considerations in time intervals mining, in: ACM KDD on Workshop on Connected Health at Big Data Era, NYC, USA, 2014.
[44] A. Singh, G. Nadkarni, O. Gottesman, S.B. Ellis, E.P. Bottinger, J.V. Guttag, Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration, J. Biomed. Inf. 53 (2015) 220–228.
[45] M. Stacey, C. McGregor, Temporal abstraction in intelligent clinical data analysis: a survey, Artif. Intell. Med. 39 (2007) 1–24.
[46] D. Stopel, Z. Boger, R. Moskovitch, Y. Shahar, Y. Elovici, Improving worm detection with artificial neural networks through feature selection and temporal analysis techniques, in: Proceedings Third International Conference on Neural Networks, Barcelona, 2006.
[47] J. Sun, C.D. McNaughton, P. Zhang, A. Perer, A. Gkoulalas-Divanis, J.C. Denny, J. Kirby, T. Lasko, A. Saip, B.A. Malin, Predicting changes in hypertension control using electronic health records from a chronic disease management program, J. Am. Med. Inf. Assoc. 21 (2014) 337–344.
[48] R. Villafane, K. Hua, D. Tran, B. Maulik, Knowledge discovery from time series of interval events, J. Intell. Inf. Syst. 15 (1) (2000) 71–89.
[49] E. Winarko, J. Roddick, Armada - an algorithm for discovering richer relative temporal association rules from interval-based data, Data Knowl. Eng. 63 (1) (2007) 76–90.
[50] S. Wu, Y. Chen, Mining non-ambiguous temporal patterns for interval-based events, IEEE Trans. Knowl. Data Eng. 19 (6) (2007) 1041–4347.
[51] J. Wu, J. Roy, W.F. Stewart, Prediction modelling using EHR data: challenges strategies, and a comparison of machine learning approaches, Medical Care 48 (2010) 106–113.

[52] Z. Xing, J. Pei, E. Keogh, A brief survey on sequence classification, SIGKDD Expl. 12 (2010) 1.
[53] P. Yadav, M. Steinbach, V. Kumar, G. Simon, Mining electronic health records (EHR): a survey, ACM Comput. Surv. 15 (2016).
[54] H. Yang, C.C. Yang, Using health-consumer-contributed data to detect adverse drug reactions by association mining with temporal analysis, ACM Trans. Intell. Syst. Technol. 6 (4) (2015) 55.
[55] I. Yoo, P. Alafaireet, M. Marinov, K. Pena-Hernandez, R. Gopidi, J.F. Chang, L. Hua, Data mining in healthcare and biomedicine: a survey of the literature, J. Med. Syst. 36 (4) (2012) 2431–2448.
[56] K. Orphanou, A. Stassopoulou, E. Keravnou, Temporal abstraction and temporal bayesian networks in clinical domains: a survey, Artif. Intell. Med. 60 (2014) 133–149.
[57] R. Moskovitch, Y. Shahar, Medical temporal-knowledge discovery via temporal abstraction, American Medical Informatics Association (AMIA) 2009, San Francisco, USA, 2009. 1–9.
[58] D. Reichert, D. Kaufman, B. Bloxham, H. Chase, N. Elhadad, Cognitive analysis of the summarization of longitudinal patient records, in: AMIA Annual Symposium, Washington, DC, 2010, pp. 667–671
[59] J. Hirsch, J. Tanenbaum, S.L. Gorman, C. Liu, E. Schmitz, D. Hashorva, A. Ervits, D. Vawdrey, M. Sturm, N. Elhadad, HARVEST, a longitudinal patient record summarizer, J. Am. Med. Inform. Assn. 22 (2) (2015) 263–274.
[60] R. Moskovitch, N. Peek, Y. Shahar, Classification of ICU patients via temporal bbstraction and temporal patterns mining, IDAMAP 2009, Verona, Italy, 2009.
[61] R. Moskovitch, I. Gus, S. Pluderman, D. Stopel, C. Glezer, Y. Shahar, Y. Elovici, Detection of unknown computer worms activity based on computer behavior using data mining, in: IEEE Symposium on Computational Intelligence in Security and Defense Applications, Honolulu, Hawaii, 2007.
[62] T. Shimshon, R. Moskovitch, L. Rokach, Y. Elovici, Continuous verification using keystroke dynamics. in: The International Conference on Computational Intelligence and Security (CIS 2010), Nanning, China, 2010.