Journal of Biomedical Informatics 134 (2022) 104169

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

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

Original Research

Visualization of frequent temporal patterns in single or two populations

Guy Shitrit[a,*], Noam Tractinsky[a], Robert Moskovitch[a,b]

[a] Software and Information Systems Engineering, Ben-Gurion University of the Negev, Beer Sheva, Israel
[b] Population Health and Science, Ichan Medical School at Mount Sinai, NYC, USA

ARTICLE INFO

Keywords:
Temporal pattern visualization
Temporal pattern discovery
Temporal knowledge discovery
Time intervals mining
Classification

ABSTRACT

Temporal knowledge discovery in clinical problems, is crucial to investigate problems in the data science era. Meaningful progress has been made computationally in the discovery of frequent temporal patterns, which may store potentially meaningful knowledge. However, for temporal knowledge discovery and acquisition, effective visualization is essential and still stores much room for contributions. While visualization of frequent temporal patterns was relatively under researched, it stores meaningful opportunities in facilitating usable ways to assist domain experts, or researchers, in exploring and acquiring temporal knowledge. In this paper, a novel approach for the visualization of an enumeration tree of frequent temporal patterns is introduced for, whether mined from a single population, or for the comparison of patterns that were discovered in two separate populations. While this approach is relevant to any sequence-based patterns, we demonstrate its use on the most complex scenario of time intervals related patterns (TIRPs). The interface enables users to browse an enumeration tree of frequent patterns, or search for specific patterns, as well as discover the most discriminating TIRPs among two populations. For that a novel visualization of the temporal patterns is introduced using a bubble chart, in which each bubble represents a temporal pattern, and the chart axes represent the various metrics of the patterns, such as their frequency, reoccurrence, and more, which provides a fast overview of the patterns as a whole, as well as access specific ones. We present a comprehensive and rigorous user study on two real-life datasets, demonstrating the usability advantages of the novel approaches.

1. Introduction

Temporal knowledge discovery stores significant potential impact in medicine, especially now that more longitudinal electronic health records' data becomes more accessible [35]. Specifically the investigation of longitudinal clinical data is attracting focus in recent years in biomedical informatics. Frequent temporal patterns discovery is a commonly used method that stores great opportunities to discover and acquire knowledge. The use of temporal abstraction and Time Intervals Related Patterns (TIRPs) in the analysis of longitudinal data Moskovitch and Shahar [31], Moskovitch and Shahar [32], and specifically in Electronic Health Records, was demonstrated successfully in several tasks, such as classification and outcomes prediction, which even outperformed deep learning based methods, such as Long Short Term Memory (LSTM) or Convolutional Neural Network (CNN) [56,17,13,37,38].

However, with opportunities, also come challenges. While meaningful advances were made in the computational aspects of the discovery of frequent temporal patterns, one of its main challenges remains the large number of discovered patterns. Some studies suggested relevant interestingness measures to favor patterns, to the best of our knowledge, no study has yet focused on the visualization of the exploration of the tree of patterns (the result of any patterns mining method, whether not temporal (i.e., association rules mining) or temporal (i.e., sequential mining or time intervals mining) patterns. In order to acquire meaningful temporal knowledge, effective visualization is required, which is the focus of this paper and a topic relatively that was not much studied in the research community.

Visualizing frequent temporal patterns is a challenging task. First, the commonly large amount of discovered patterns makes it difficult to explore and choose those that are of a greatest interest and meaning. Note that this work focuses on the visualization of a collection of discovered temporal patterns that are the output of a mining process, rather than visualization of temporal data, and the purpose is not to provide analytics of the raw data, but rather the input data are the discovered TIRPs and their properties.

[*] Corresponding author.
E-mail addresses: shitritg@post.bgu.ac.il (G. Shitrit), robertmo@bgu.ac.il (R. Moskovitch).

https://doi.org/10.1016/j.jbi.2022.104169
Received 30 June 2021; Received in revised form 11 August 2022; Accepted 13 August 2022
Available online 28 August 2022

Thus, the input data are the discovered TIRPs, which can not be analyzed using traditional approaches for analytics. Additionally, while the visualization of temporal patterns can clearly help in knowledge acquisition and acquiring insights from the displayed patterns, current systems [42,16,54,46] do not enable to users to compare patterns discovered in different populations.

Thus, the motivations behind the development of the introduced interface are the following: exploration of an enumeration TIRPs tree to enable access to TIRPs of interest, based on various TIRPs' metrics, such as the vertical support, horizontal support, or mean duration. This paper introduces KLW, a novel visual interface that enables exploration of an enumeration tree of frequent patterns through navigation, search and retrieval, and visualization of the pattern's related metrics. Specifically, the interface was designed and implemented for frequent TIRPs, of which Sequential Patterns are a private case, and these principles can be used easily for Sequential Patterns visualization as well. In addition to visualization and exploration of the discovered tree of frequent TIRPs from a single population, in this paper, the option to visualize and explore TIRPs that were discovered separately in two populations is investigated for the first time, as far as we know. Thus, based on two populations (that can be two classes of a classification problem) from which two trees of frequent TIRPs were discovered, the novel visualization enables to access easily the TIRPs that are most discriminatively efficiently through an advanced informative bubble chart based visualization.

To evaluate the usability and usefulness of the KLW interface, we conducted a comprehensive user study in which the participants had to perform tasks using the new novel interfaces. We evaluated their performance according to their answers' correctness (i.e., accuracy), task execution time duration, and level of ease of task execution (a subjective measure).

To summarize, the contributions of the paper are the following:

- Navigation and exploration of a frequent patterns enumeration tree through event types (symbols) or various pattern's metrics
- Visualization of a TIRP, including its supporting instances, and entities' demographic distributions and metrics
- Novel bubble chart based presentation of the frequent patterns in a bubble chart that presents several metrics using the axes and bubble' properties
- Novel presentation of patterns in two populations and their discriminative metrics
- Novel presentation of the discriminative patterns on a bubble chart illustrating their frequency in the populations, and more metrics using a bubble chart

The rest of the paper is organized accordingly. We start with the Background section covering relevant topics, propose our new visualization in the Methods section, introduce the user study in the Evaluation section, and then its findings in the Results section.

2. Background

Temporal data visualization has evolved over the past years [1], offering various visualization tools and techniques in order to present temporal data or knowledge. Different types of time-oriented data require visualization for various purposes; however, visualization of frequent temporal patterns has been significantly under-investigated.

2.1. Temporal data and temporal data mining

Temporal data is data that represents measurements or states in time, which comes from many sources ranging from manual data entry to data collected using sensors. Sampling might occur at a fixed frequency, which is often the case for automated sampling or at random periods, as occurs in manual measurements. Often, specific time-stamped data points might be missing, or their measurements might include an error. Raw data might also be represented by time intervals, such as medication-administration periods. The duration of the events can be fixed or vary.

Temporal data mining is an extension and a sub-field of data mining. Data mining is the process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems. Temporal data mining has the capability of mining activity rather than just states, and makes it possible to infer relationships of contextual and temporal proximity that may indicate a cause–effect association [52]. This sub-field offers an understanding of scientific phenomena and may create more productive and accurate classification models [33,34].

2.2. Temporal abstraction

A major challenge in mining multivariate temporal data is the heterogeneity of the sampling forms and appearances of the data, which can be time-point series sampled regularly or not, or events that may or not have duration. Abstracting time-point series into symbolic time intervals representation enables a uniform representation of the heterogeneous temporal variables.

State Abstraction (see Fig. 1), which we use in this study, includes mainly-two steps. First classifying each time-point value into a state, given relevant cutoffs, and then concatenating adjacent states that have the same symbol. A detailed description of the process can be found in Moskovitch and Shahar [31].

Several common discretization methods exist, such as Equal Width Discretization (EWD); Symbolic Aggregate approXimation (SAX) [27], which consists on the Gaussian distribution of the data; Temporal Discretization for Classification (TD4C), and more. TD4C is a supervised method that learns the cutoffs that result in the most different state distribution between the classes [31].

Previous studies showed the advantages of TD4C [32] in comparison to EWD and SAX, especially for the classification of time-oriented clinical data. In the evaluation of the current study, we performed a comparison of the use of both knowledge-based temporal abstraction (KB), EWD, SAX, and TD4C. After the data is transformed into a uniform representation of symbolic time intervals, which are defined in Definition 1, TIRPs can be discovered [31,32].

Definition 1. *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 symbolic concepts, or their abstraction, which in our study can be a state of a lab result (i.e., high blood pressure), abstraction of vital signs, or exposure to a medication.*

2.3. TIRPs mining and TIRPs' metrics' definitions

In the past two decades, meaningful advancements were made in the problem of TIRPs mining from symbolic time intervals (STIs) data, in which the temporal relations among the STIs are represented using entire or some subset of Allen's temporal relationsAllen [2]. Allen defined seven mutually exclusive temporal relations including: *before, meets, overlaps, finish-by, contains, starts,* and *equal*, as shown in Fig. 3. Several algorithms were proposed for time intervals mining, improving their efficiency and runtime [55], Hoppner [19,22], Mörchen and Ultsch [30], Moskovitch and Shahar, [32]. Among these, ARMADA, by Winarko and Roddick [53], is a projection-based efficient time-intervals mining algorithm, based on an efficient sequential patterns mining algorithm. Patel et al., [40] introduced IEMiner, which extends the patterns directly. The KarmaLego method [33], which we use in this study, uses an efficient data structure and exploits the transitivity property of the temporal relations for an efficient candidate generation process that filters candidates that cannot exist in reality. The analysis of its complexity is described in [31].

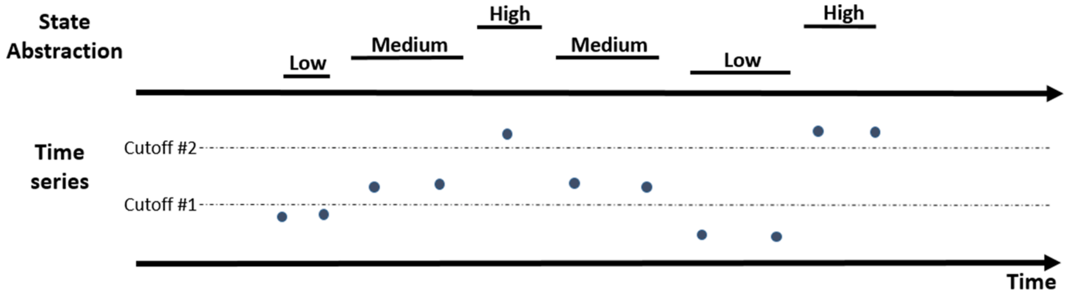


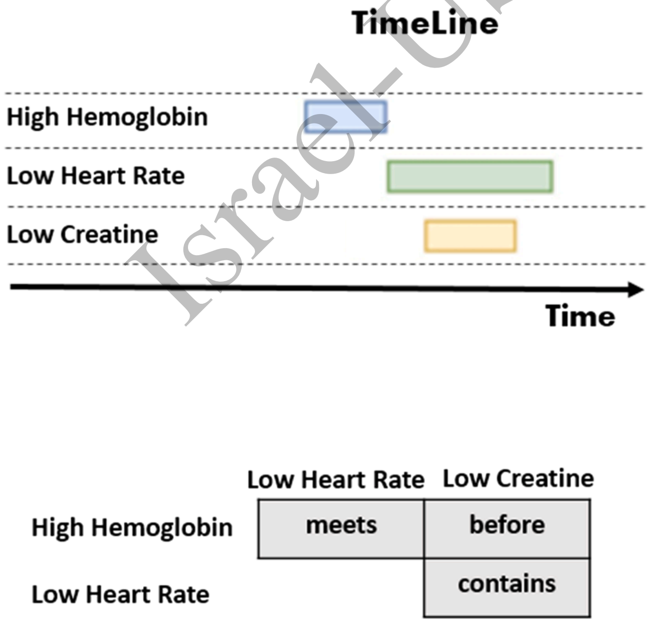
Fig. 1. A series of raw time-series (at the bottom) is abstracted into an interval-based state abstraction that has three discrete values: Low, Medium, and High (at the top).

Unlike previous methods, KarmaLego discovers the complete set of frequent TIRPs since it discovers their entire set of horizontally supporting instances [31,32]. This enables a complete and comprehensive TIRP representation as features for classification, in addition to Binary (i.e., whether the TIRP exists at least once in the records), that was used in earlier studies [5,40]. Recently Harel and Moskovitch [18] published TIRPClo which discovers closed TIRPs, and is complete as well.

Definition 2. *A non-ambiguous Time Interval-Related Pattern (TIRP) P is defined as $P = \{I,R\}$, where $I = \{I^1, I^2,.., I^k\}$ is a set of k symbolic time intervals ordered lexicographically (ordered by their start-time, end-time, or symbol) and $R = \bigwedge_{i=1}^{k} \bigwedge_{j=i+1}^{k} r(I^i, I^j) = \{r_{1,2}(I^1,I^2),.., r_{1,k}(I^1,I^k), \ldots, r_{k-1,k}(I^{k-1},I^k)\}$ which defines all the temporal relations among each of the $(k^2\text{-}k)/2$ pairs of symbolic time intervals in I.*

Fig. 2 presents an example of a 3-sized TIRP, having three lexicographically ordered symbolic time intervals displayed graphically along the timeline on the top. The half-matrix on the bottom represents the conjunction of the temporal relations among the symbolic time intervals that enables a non-ambiguous representation.

Since a temporal pattern can occur and be discovered multiple times, which appear as multiple instances within a single entity (an entity can be the entire records of a patient, or a specific time window of 6 h in the Intensive Care Unit of a specific patient for example), we distinguish between two types of support: vertical support (Definition 3), which refers to the percentage of the entities in the database having a pattern, and the horizontal support (Definition 4), which represents the number of patterns discovered within the symbolic time intervals of a specific entity.



	Low Heart Rate	Low Creatine
High Hemoglobin	meets	before
Low Heart Rate		contains

Fig. 2. A 3-sized TIRP presented graphically on a timeline (at the top) and the temporal relations as a half-matrix (at the bottom).

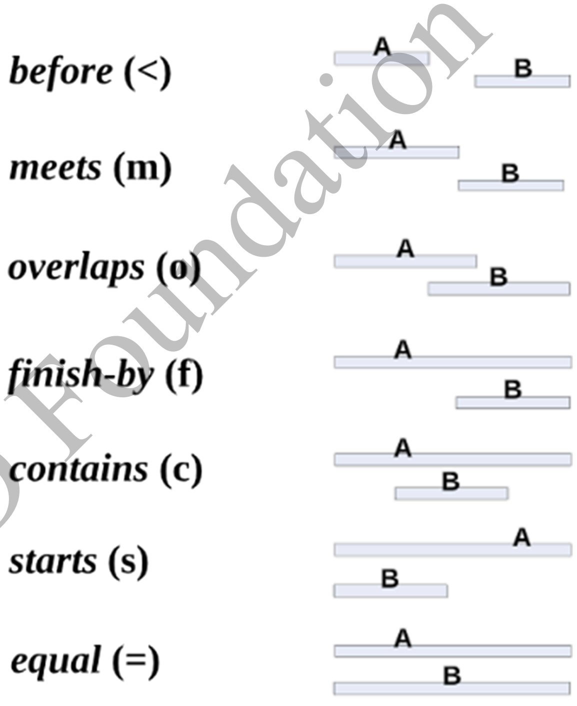


Fig. 3. Allen's temporal relations among two symbolic time intervals.

Definition 3. *Given a database of $|E|$ distinct entities, the vertical support (VS) of a TIRP P is denoted by the cardinality of the set EP of distinct entities within which P holds at least once, divided by the total number of entities $|E|$: $Ver\_sup(P) = \frac{|E^P|}{|E|}$.*

In the context of patterns mining, the term support is usually used and refers to what we call vertical support (VS). When a TIRP has vS above a minimal predefined threshold, it is referred to as frequent.

The task of mining time-intervals is to find all the TIRPs whose vS is above a predefined minimal vertical support threshold. However, in order to discover the complete set of TIRPs, the entire set of horizontally supporting (Definition 4) instances also have to be discovered [32].

The horizontal support (Defintion 4) and Mean Duration (Definition 6) metrics are discovered as part of the TIRPs discovery process. The following definitions describe those TIRP's metrics.

Definition 4. *The horizontal support (HS) of a TIRP P for an entity e_i, is the number of instances of the TIRP P found in e_i. However, to represent the average HS in a group of entities, for example all the patients that have the TIRP P (its VS), we use the Mean Horizontal Support (MHS).*

Definition 5. *The Mean Horizontal Support (MHS) of a TIRP P describes* *the average HS values of all its $|E^p|$ supporting entities (each entity I may have a different number of hs_i instances). Thus, given a TIRP P and a set of $|E^p|$ supporting entities, having a hs_i >=1 $\{hs_1,\ hs_2\cdots..hs_{|E^p|}\}$, $MHS(P) = \frac{\sum_{i=1}^{e}(hs_i)}{|E^p|}$.*

Since the duration of the STIs in a TIRP are not part of its definition, the duration of an STI may vary in each of its supporting instances. For that, the mean duration metric describes the average duration of the entire set of a TIRP's instances accordingly.

Definition 6. *The mean duration (MD) of the n horizontally supporting instances of the same k-sized pattern P within an entity e is defined by the average of the durations of all the instances, where each instance duration is defined from its earliest time point till its last time point. We here define the MD for TIRP:MeanDuration$(P, e) = \frac{\sum_{i=1}^{n}(max_{j=1}^{k} I_{e}^{i,j} - I_{s}^{i,1})}{n}$, where the Max operator selects the symbolic time interval having the latest end-time (e) among the k symbolic time intervals of an instance i (among n instances), and $I_{s}^{i,1}$ is the start-time (s) of the first symbolic time interval in the i^{th} instance. Their subtraction is the time duration of the instance.*

While MD is a metric that is defined per entity, in order to represent the TIRP, similar to the MHS, here is the definition of the Mean MD, which is at the TIRP level.

Definition 7. *The Mean Mean Duration (MMD) of a TIRP P is the average of the MD values of its $|E^p|$ supporting entities. Each entity i may have several instances of P (having different lengths of MDs), whose mean is the md_i (Definition 6), while the MMD refers to the average of the MD values of the entire set of entities. Thus, given a TIRP P and a set of $|E^p|$ supporting entities $\{md_1,\ md_2....,\ md^p_{|E}\}$, the MMD of P is: $MMD(P) = \frac{\sum_{i=1}^{|E^p|} mdi}{|E^p|}$.*

After running a TIRPs' discovery process, such as KarmaLego that we use in this study, or another method, the output is an enumeration tree of all the discovered frequent TIRPs, which is shown in Fig. 4. Each node represents a frequent TIRP that was discovered (including its supporting instances data). In the first level of the tree (below the root), there are patterns that consist on a single STI (1-sized pattern), and going deeper in the tree levels, there are expanded patterns, including additional new symbolic time intervals and the corresponding temporal relations to all the earlier pattern's symbolic time intervals. Going down through the tree, the TIRPs typically have lower, or equal, vertical support, but always above the minimal vS threshold. This discovered TIRPs enumeration tree is the input to the visual interface.

2.3.1. The discovered TIRPs tree visualization challenge

The discovered TIRPs enumeration tree and the patterns' related metrics' data is the input to the visual interface, whose exploration introduces several visualization challenges. The challenge in visualizing such a patterns tree is that, in order to enable browsing according to the various patterns' metrics and the pattern's details, which are essential for the user, and may influence her choices, is that it cannot be displayed as a tree that represents only the structure of the extended TIRPs. For that, here the interface shows each time a node in the tree – a TIRP – the user is on. The main requirements from such visualization is to enable browsing within the patterns tree by navigation, to present a pattern visually and to enable exploration of the patterns by their metrics' values.

2.4. Visualization of temporal data and knowledge

The field of visualization tools for temporal data is crucial in clinical data [15,21,23,36], in which temporal patterns are discovered in patients' records data [7,8,9,43]. The main goal of this kind of visualizations is to allow exploring and querying Electronic Health Records (EHRs) in order to support clinical decision-making and research [28].

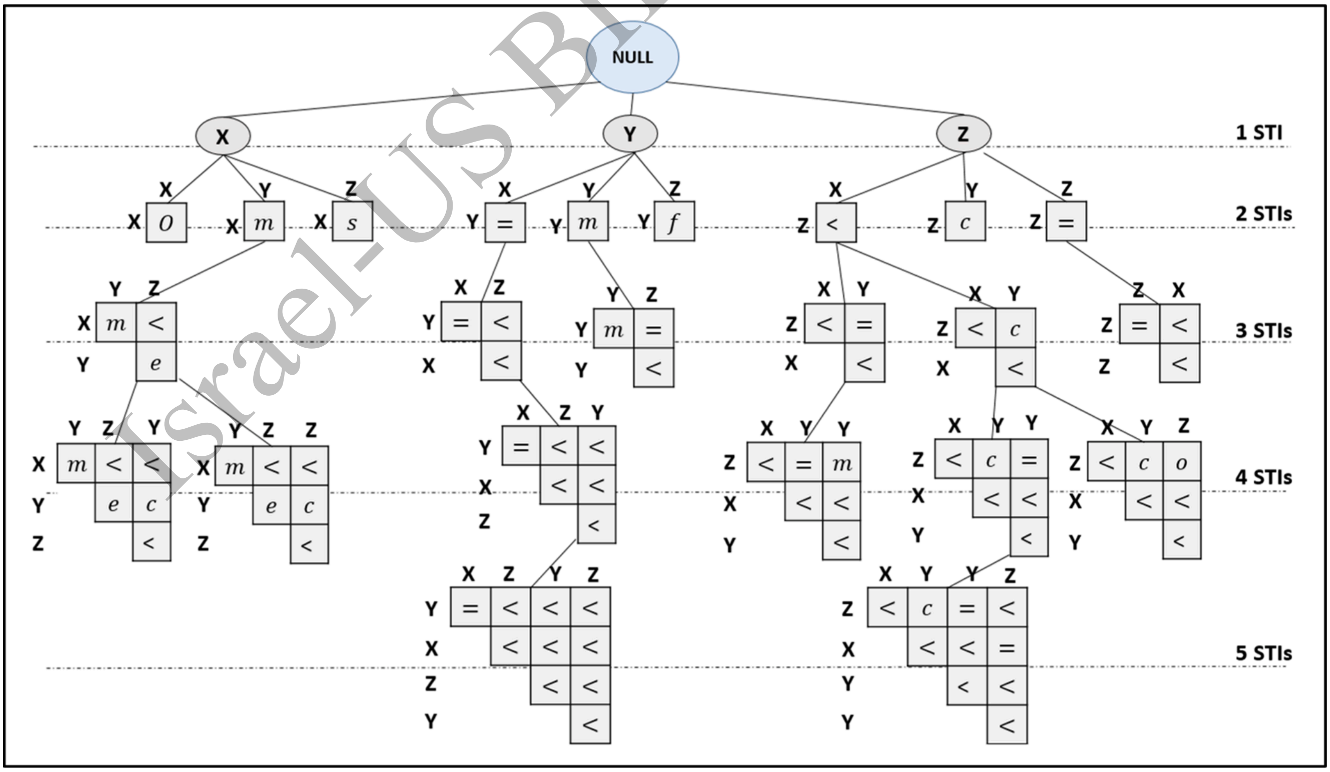


Fig. 4. A frequent TIRP enumeration tree, which is the result of a TIRPs mining process. For simplicity, it includes only three symbol types, X, Y and Z, that are frequent. Below are the frequent TIRPs, starting with the 2-sized TIRPs, and going down to larger TIRPs. Each node is a TIRP represented by its sequence of STIs and the conjunction of the temporal relations represented by the half-matrix.

Visualization of time intervals data was studied in the past in several papers for various purposes, focusing on different aspects and making meaningful contributions, in the visualization of reasoning in time intervals [10,6]. However, these studies mainly referred to the visualization of the patterns themselves, and not the visualization of an enumeration tree of a set of discovered frequent TIRPs, as we do in this paper, nor on the visualization of patterns in two populations. Wongsuphasawat and Gotz [55] introduced a visualization tool which aims at assisting in clinical studies and in diagnosis of patients. Temporal data visualization has evolved over the years [1] proposing various visualization tools and techniques in order to present temporal data [4,44,45] or knowledge, such as temporal patterns [43,15]. Moreover, as mentioned, temporal data can be transformed using temporal data abstraction into meaningful symbolic time intervals, that describe periods of time the signal is increasing or decreasing, or being in a specific state continuously (which is also what we do in this paper for purposes of temporal knowledge discovery - see section 2.2). KNAVE [29,48] and VISITORS [24], Klimov et al. [25,26,47] are visualization systems that visualize the process of temporal abstraction on raw numerical data for temporal reasoning purposes over time. VISITORS enables different types of queries that allow to retrieve the list of entities that satisfy a set of constraints and allows searching for both raw data and abstracted data among entities collection, given various associations among the temporal variables' states.

The primary purpose of visualization of temporal data or knowledge [51,52,56,55,28] is to present the data longitudinally in a way that facilitates the recognition of temporal patterns or models.

However, the presentation or visualization of multivariate heterogeneous data is highly challenging due to the variety of types of temporal data [49], such as time-points, time-intervals, temporal patterns, and more, and the longitudinal functionality that is required in order to present this type of data. Consequently, visualization tools for knowledge in the form of temporal patterns is scarce and has seldom been investigated, as we discuss later (Section 2.4.1).

2.4.1. *Visualization of temporal patterns*

Notable work in temporal patterns visualization systems was introduced by Perer and Wang [42], Gotz et al. [16],Vrotsou and Nordman [54], and Poon et al. [46]. All these systems present a visual interface that aims to display the patterns discovered by a temporal data mining algorithm and tries to allow the user to explore the pattern in an interactive manner. Most temporal patterns' visualization systems are based on the output of frequent sequential mining algorithms rather than on frequent time interval-related patterns, as we do in this paper.

There are systems that use the SPAM sequential mining algorithm [3,16] or an enhanced SPAM approach in order to make SPAM capable of detecting temporal patterns with constraints on their duration [42], while other [46,54] used the PrefixSpan [41] sequential mining algorithm.

Temporal patterns visualization systems also differ in terms of the visual technique in which the patterns are displayed. Some of the systems use Sankey diagram layouts [14] and alluvial diagrams such as Frequence [42]. Events in the frequent sequences are represented as nodes, and events that belong to the same frequent sequence (i.e., pattern) are connected by edges, where their frequency is presented by the thickness of the edges, while their color represents their correlation with some specified outcomes. Vrotsou and Nordman [54] proposed a visual technique to present the patterns, in which the patterns are presented as a tree, where each node represents an event, its color represents the event category, and the size of the nodes reflects the support of each pattern, decreasing as the tree expands. Another type of visual patterns presentation is 3D sunburst visualization [50], where each event (i.e., pattern) is presented as a colored block, and when a mouse hovers over an event, the visualization highlights the full path to that event (the pattern) and its events sequence, and the support values are displayed. In Gotz et al. [16], the patterns are presented in an event chart, while the distance between the events represents the mean duration between them.

Finally, we provide examples of how the main features that the systems need to support, according to Shneiderman [52], are implemented in the systems. Only one system, Frequence, allows zooming by displaying patterns that include events at different hierarchy levels by their category. This enables the user to notice more patterns by different hierarchical levels of the categories of the pattern's events. Filtering temporal patterns can be done either by temporal or non-temporal conditions. An example of filtering by non-temporal conditions can be found in Frequence, which allows filtering of the presented patterns by specific population or by some outcome. Frequence also allows filtering under temporal conditions, and Gotz et al. [16] allow filtering patterns that contain only events that occurred within a specific time range or by a specific event sequence. These researchers presented the retrieved patterns in a scatterplot, where each pattern is located according to its support in the negative and positive outcomes, and color and size represent their correlation to certain outcome measures. Another essential feature to support visualization systems is aggregation.

The surveyed pattern-based systems support relatively simple aggregations, such as the average number of events per pattern and mean duration per pattern. Gotz et al. [16] show gender and age distributions for a set of patients returned by a query module. While the visualization of temporal patterns can clearly improve knowledge and insights from the displayed patterns, the main drawback of the surveyed systems is that they do not support users in comparing patterns. For example, none of the systems support the comparison of patterns that are found in different populations. Another drawback is that the systems provide very few relevant statistical details on the displayed patterns. For example, these visualizations provide no information regarding supporting entity distributions by demographic data such as gender and age. This makes it impossible to draw conclusions from the visualized patterns, for example, whether patterns are more common in men or in women, or among young or old people. In addition, most of these systems are based on frequent sequential mining algorithms rather than on frequent time interval-related patterns, and so the duration of the events in the pattern sequence has no effect.

3. Materials and methods

In this study, we introduce KLW, a novel visual interface that enables presentation and browsing of a tree of frequent time interval-related patterns (TIRPs) that were discovered by the KarmaLego algorithm (See Section 2.3). While KLW is demonstrated on TIRPs, its principles can also be used for other types of patterns, such as sequential patterns, which are in fact a private case of TIRPs (having time intervals without duration).

The KLW interface was designed with three main purposes: 1. to conveniently and effectively navigate and explore the frequent patterns enumeration trees that were discovered by a patterns mining process, given various metrics (such as a TIRP's frequency, reoccurrence, duration, and more), as well as their distribution in a population; 2. to identify patterns that differ significantly between two populations, which we refer to as "Discriminative TIRPs"; and 3. to enable the query and retrieval of patterns of interest that were discovered in either a single population or two populations. The query results are presented in two different ways: by graphs or by tables which are part of the comparison in the evaluation phase. This section will present the KLW interface by focusing on three main visualizations. We implemented the server-side project using Python and the front-side using Angularjs and JavaScript, which was applied on an Apache server.

3.1. *KLW – visualization of TIRPs analytics*

The KLW interface for TIRPs Analytics is introduced here in details, which focuses on enabling the TIRP enumeration tree to be browsed, along with each TIRP and its related information. Fig. 5 shows an overview of the KLW home page. The menu bar enables navigation between various functions. The DataSets option opens a table that includes all available data sets and related information. The user can now find and choose the working data set.

Table 1 presents the main KLW interfaces, and their main purpose and different properties. The objective of the interfaces is to enable users to explore the discovered enumeration tree of frequent TIRPs, and to query the TIRP tree for specific TIRP criteria. As we explain later, KLW enables two visualization modes: KLW1 for TIRPs discovered from a single population, and KLW2 for TIRPs mined from two populations.

For each option, three interfaces exist: 1. Explore, which navigates and explores the TIRP tree; Query, which enables querying the TIRP tree for specific TIRPs according to a given criterion, for which there are two optional result views: 2. Table, in which the results are displayed in a tabular presentation, and 3. Graph, in which the results are shown in a two-dimensional bubble chart graph.

The purpose of Table 1 is to provide a high level of organization and a summary of KLW, with additional properties of each interface to make their differences clear. Thus, for example, in the Query Graph, the chart axes have different properties for KLW1 versus KLW2: in the first, these are the TIRP metrics; and in the second, the metric is the same, but for the different populations. In the next subsections, we describe each interface in detail.

3.1.1. *KLW1-Explore: Visualization of TIRPs in an enumeration tree*

The output of the TIRPs discovery process results from running any TIRPs discovery method such as KarmaLego, and is an enumeration tree of TIRPs, as shown in Fig. 4. The TIRPs are displayed in the way they were generated and discovered in the discovery (mining) process in which each tree node is a TIRP, including its supporting instances.

The order of the enumeration is determined by the symbols' alphabetical order, which is the way the TIRPs are generated in the mining process, and shown in the visualization. This tree of TIRPs serves as the input to the visualization system. The main challenge is to present it visually and allow users to navigate and explore it effectively and conveniently. To support the discovery of interesting patterns, KLW allows browsing of the tree according to several metrics, such as by the TIRP properties (symbols or temporal relations) or its metrics. Due to the various TIRP metrics, and aspects, based on which a potential user may want to browse the resulting TIRP tree, visualization based on the TIRPs' tree structure is not sufficient to enable the TIRPs to be explored based on various metrics such as the TIRP's vertical or mean horizontal support, or mean duration. There are several requirements from this KLW1-Explore component. It must: 1. enable browsing of the TIRPs within the patterns tree by navigation according to the TIRP's properties or metrics; 2. visually display the patterns along with their metrics to facilitate their exploration; and 3. demonstrate demographic data distribution of the entities that had each TIRP, according to various classifications in order to gain more knowledge about the discovered patterns. An example of this is to provide the answer to the question: How are the supporting entities of the pattern divided by gender? A detailed explanation will now be given of how all of these requirements in the system were met.

In order to enable the navigation in the TIRPs' enumeration tree based on the various metrics, in each step, the interface displays the selected TIRP (node in the tree) using a dynamic table (see Fig. 6a) that shows the following optional extended TIRP choices. Each row in the dynamic table represents an optional extended pattern based on its additional symbol, and the temporal relation between the last STI of the current TIRP and the new symbol that creates the extended TIRP.

Note, when adding the new last symbol, according to the TIRP representation (Definition 2), there might be several temporal relations between the new STI and the current TIRP's STIs, depending on the size of the TIRP (if the extended TIRP's size is 2, then it is only one; if it is 3, then it is 2, and so on). However, in the table, only the temporal relation between the current TIRP's last STI and the new STI is displayed.

Initially, the table shows the root level of the tree, listing all the frequent patterns that include all the frequent 1-sized TIRPs, which are actually the frequent symbols (Subsection 2.3). Selecting a row in the table in Fig. 6a means selecting the next symbol related to the last STI in the current TIRP. The extended TIRP information is loaded to the page with all its corresponding data.

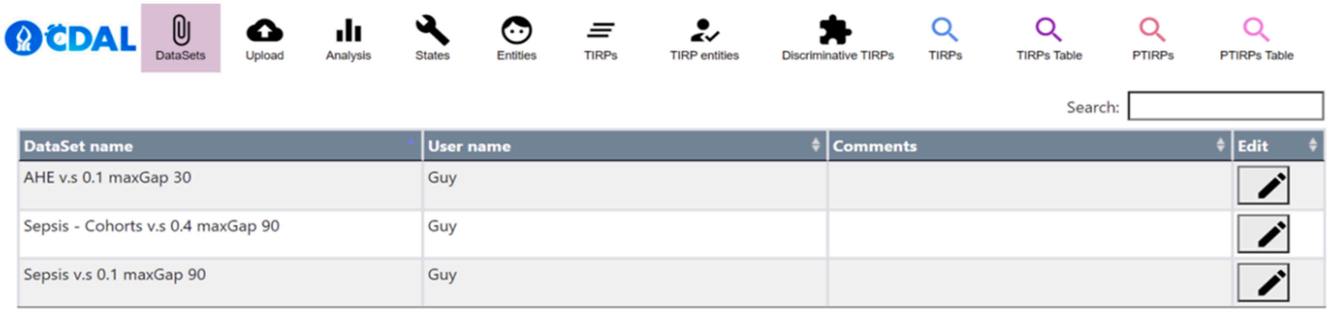
The color of each row in the table can either be white, meaning it is a leaf and that this TIRP has no further extensions; or dark, meaning the TIRP has further extensions. Thus, the user can navigate along the TIRP enumeration tree to the lowest level (leaf level) of the branch, and also reach all the TIRPs in the tree. In Fig. 6b, the current TIRP path in the tree is displayed, which is also the sequence of its symbols, enabling a direct return to any of the previous levels in the path.

In Fig. 6c, the current 2-sized TIRP is illustrated visually using a timeline chart including the average duration of each of the TIRP's STIs. Each STI is presented in a different row and different color on the timeline. Fig. 6d displays more information about the current TIRP, including the number of entities in which the TIRP was found, its vS (Definition 3), the Mean Horizontal Support (Definition 5), and the Mean Mean Duration of the TIRP (Definition 7).

Fig. 6e displays a pie chart of the demographic variables' value distributions. For example, it shows how the supporting entities are distributed by gender to show the percentage of females or males, or the distribution of the supporting entities according to age groups, etc.

Fig. 7 shows the interface after the user "drilled down" to the third level of the tree. After the marked symbol in the table was selected, the following extended 3-sized TIRP was loaded. The changes can be seen in the other interface components as well: in the table in Fig. 7a, new optional symbol extensions are displayed,; in the table shown in Fig. 7b, the current path was updated and the selected pattern were changed, and so has the information displayed in Fig. 7c, d, and e. Particularly in 7c a 3-sized TIRP that is an extended TIRP of the one that is displayed in Fig. 6c, in which there is an additional symbolic time interval (yellow).

To show the supporting entities of a TIRP, the user can move to the "TIRP entities" option in the menu. Fig. 9 presents the supporting entities of the selected TIRP at KLW–Explore. Fig. 9b presents the mean presentation of the current TIRP. The table in Fig. 9a presents all the



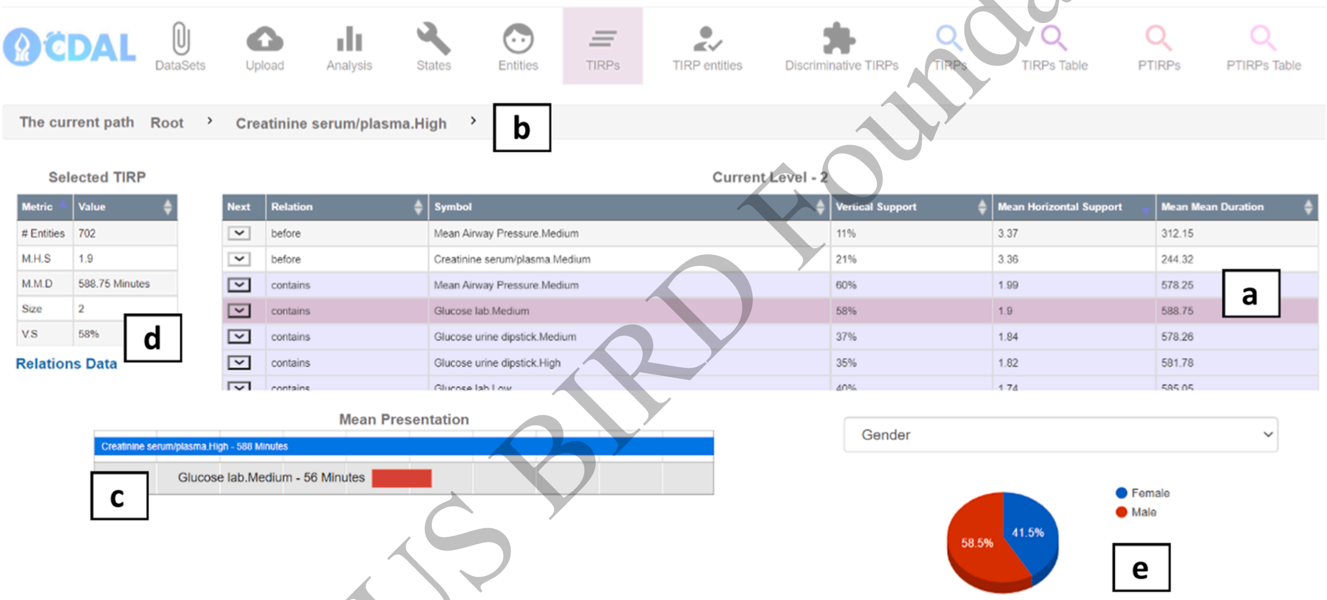
Search:

DataSet name	User name	Comments	Edit
AHE v.s 0.1 maxGap 30	Guy		✎
Sepsis - Cohorts v.s 0.4 maxGap 90	Guy		✎
Sepsis v.s 0.1 maxGap 90	Guy		✎

Fig. 5. The KLW Interface main page, including the menu at the top, and the currently available datasets.

Table 1
The main components of KLW, their main task, and their properties.

	KLW1 (A single population/class)			KLW2 (Two populations/classes)		
	KLW1-Explore	KLW1-Query		KLW2-Explore	KLW2-Query	
		Table	Graph		Table	Graph
Task	Explore the TIRPs' enumeration tree	Query for TIRPs	Query for TIRPs	Explore the TIRPs' enumeration tree	Query for TIRPs	Query for TIRPs
Query Results	–	Table	Graph-based bubble chart	–	Table	Graph-based bubble chart
Displayed Metrics	Current explored TIRP's metrics	–	–	Current explored TIRP's metrics in each of the populations	–	–
Graph Axes			X – Horizontal Support Y – Vertical Support			X – Vertical Support in Population 1 Y – Vertical Support in Population 2
Bubble Properties			Tone – Mean Mean Duration Size – not used			Tone – Delta Mean Mean Duration Size – Delta Horizontal Support



The current path Root › Creatinine serum/plasma.High ›

Selected TIRP

Metric	Value
# Entities	702
M.H.S	1.9
M.M.D	588.75 Minutes
Size	2
V.S	58%

Relations Data

Current Level - 2

Next	Relation	Symbol	Vertical Support	Mean Horizontal Support	Mean Mean Duration
	before	Mean Airway Pressure.Medium	11%	3.37	312.15
	before	Creatinine serum/plasma.Medium	21%	3.36	244.32
	contains	Mean Airway Pressure.Medium	60%	1.99	578.25
	contains	Glucose lab.Medium	58%	1.9	588.75
	contains	Glucose urine dipstick.Medium	37%	1.84	578.26
	contains	Glucose urine dipstick.High	35%	1.82	581.78
	contains	Glucose lab.Low	40%	1.74	585.05

Mean Presentation

Creatinine serum/plasma.High - 588 Minutes

Glucose lab.Medium - 56 Minutes

Fig. 6. KLW1-Explore: TIRPs tree exploration and visualization. The title indicates that the interface is at the second level of the tree (a 2-sized TIRP) in the enumeration tree. The dynamic table (a) presents the optional symbol extensions of the current TIRP. The path in the tree till the current TIRP (which is also its sequence of symbols) is shown at the top (b). A visual illustration of the current TIRP on a timeline diagram is displayed (c), along with some of its metrics (d). A filter of demographic data is applied to display a marginal distribution (gender in this example) of the supporting entities of the TIRP (e).

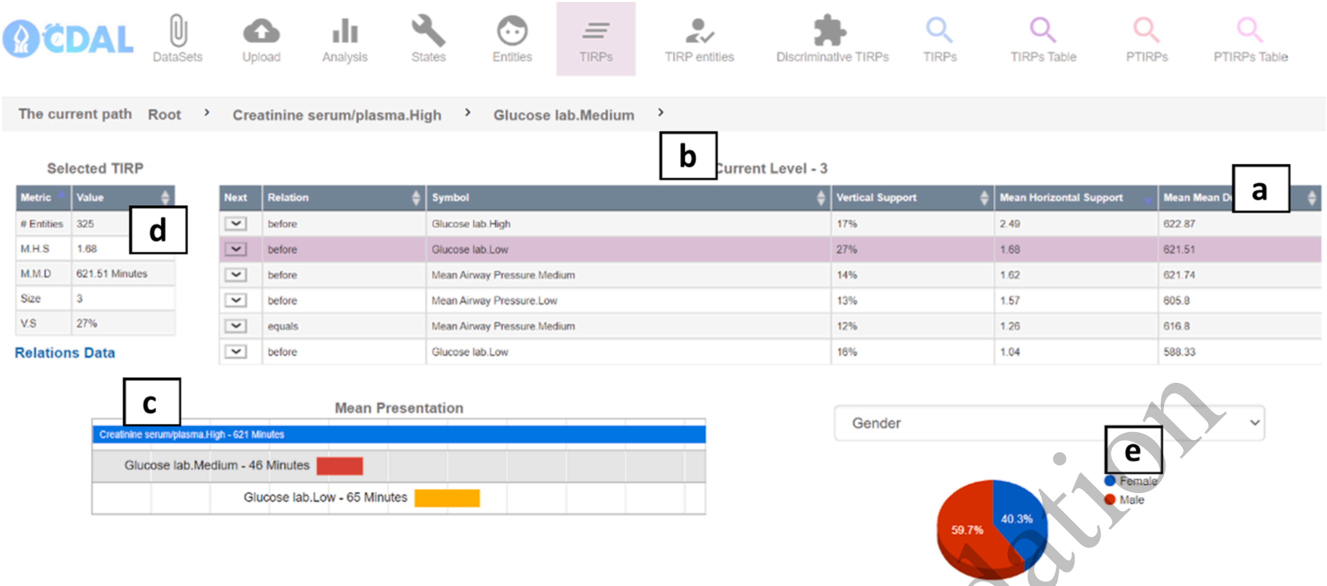
supporting entities of the TIRP with their HS, MD, and demographic data. Selecting a supporting entity will present on the right side information of the TIRP that is relevant for the selected entity. Fig. 9c presents the mean presentation of the TIRP's instances in the selected entity, and Fig. 9d presents each of the instances of the TIRP in the entity (its HS instances). As a result, the user can explore each entity and its TIRP instances' symbolic time interval durations.

3.1.2. KLW2–explore: visualization of two populations through discriminative TIRPs

To assist a domain expert or data scientist in exploring the patterns that are discriminative, which means that they appear differently in two populations, the following interface was designed. Identifying discriminative TIRPs can be useful for knowledge discovery and also for classification, when two populations are in fact two classes in a classification task. In this task, there are two TIRP enumeration trees. Each was discovered from one population, or class, of entities' data using a TIRP discovery method, which is the input to the interface. To present and explore the patterns to compare certain populations, the two TIRP enumeration trees are unified into a single tree, as shown in Fig. 10. Note that each of the trees includes TIRPs found to be frequent (above the minimal vertical support threshold) in the relevant population. While a meaningful part of the trees may intersect, some TIRPs may be found frequent in one population and in infrequent in the other. However, this does not mean that there are no instances of this TIRP in the other population, but only that it is not above the minimal vS.

The main requirements from this KLW2–Explore component are: 1. to explore each TIRP along with the TIRP's metrics in both populations; 2. to enable comparison of the TIRP's metrics in the populations; and 3. to assist a user in identifying the discrimination potential of the TIRPs among the populations according to their metrics' values and corresponding differences.

The KLW2-Explore interface, to some extent, consists of the KLW1-Explore interface shown in Fig. 6, but extends it to display the TIRPs' metrics' values in each of the two populations. The table in Fig. 11a enables navigation on the unified enumeration tree by the dynamic



The current path Root › Creatinine serum/plasma.High › Glucose lab.Medium ›

Selected TIRP

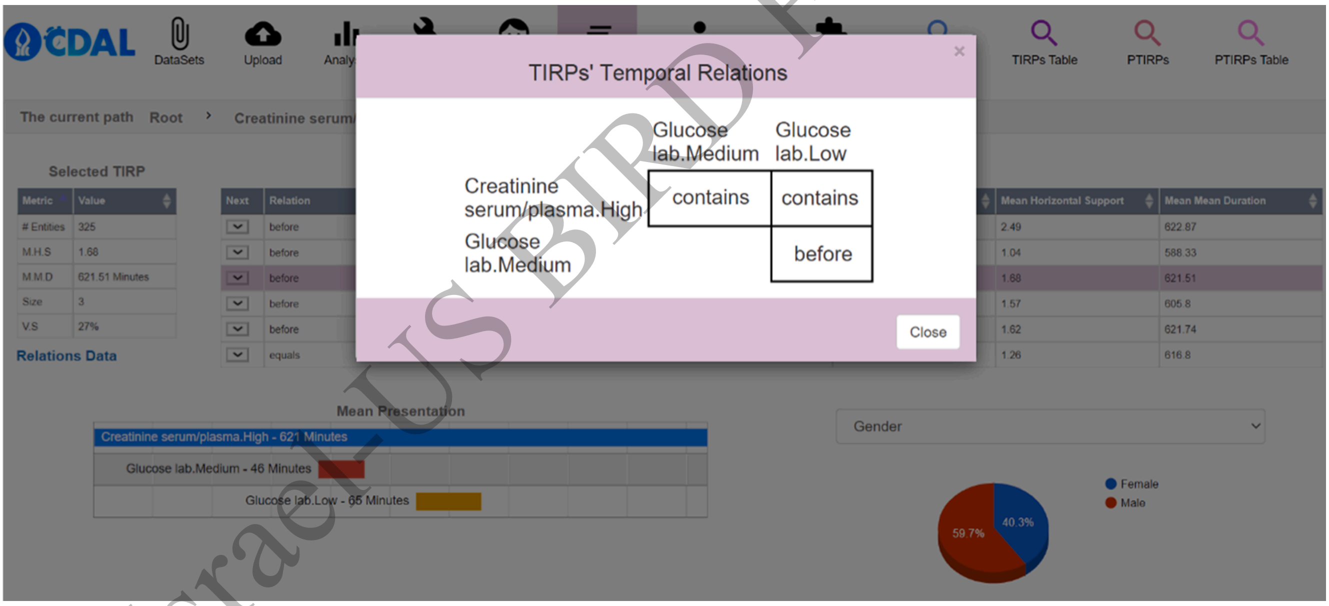
Metric	Value
# Entities	325
M.H.S	1.68
M.M.D	621.51 Minutes
Size	3
V.S	27%

Relations Data

Current Level - 3

Next	Relation	Symbol	Vertical Support	Mean Horizontal Support	Mean Mean D
	before	Glucose lab.High	17%	2.49	622.87
	before	Glucose lab.Low	27%	1.68	621.51
	before	Mean Airway Pressure.Medium	14%	1.62	621.74
	before	Mean Airway Pressure.Low	13%	1.57	605.8
	equals	Mean Airway Pressure.Medium	12%	1.26	616.8
	before	Glucose lab.Low	16%	1.04	588.33

Fig. 7. KLW1-Explore after the user selected the next symbol TIRP extension in Fig. 6, which is loaded as a 3-sized TIRP, as shown here (the third level of the TIRP enumeration tree), including the new optional symbol extensions in the table (a), and the rest of the information is updated according to the current chosen TIRP. If the user wants to see the entire set of temporal relations of a chosen TIRP, it can be viewed by clicking on the "Relations Data" at the bottom of Fig. 7d, which presents the conjunction of the temporal relations of the current chosen TIRP. Fig. 8 shows the page detailing the temporal relations, after clicking on the "Relations Data" link.


CDAL
DataSets
Upload
Analy
TIRPs Table
PTIRPs
PTIRPs Table
×
Close
Mean Presentation
Creatinine serum/plasma.High - 621 Minutes
Glucose lab.Medium - 46 Minutes
Glucose lab.Low - 65 Minutes
Gender
Female
Male
59.7%
40.3%

TIRPs' Temporal Relations

	Glucose lab.Medium	Glucose lab.Low
Creatinine serum/plasma.High	contains	contains
Glucose lab.Medium		before

The current path Root › Creatinine serum/

Selected TIRP

Metric	Value
# Entities	325
M.H.S	1.68
M.M.D	621.51 Minutes
Size	3
V.S	27%

Relations Data

Next	Relation	Mean Horizontal Support	Mean Mean Duration
	before	2.49	622.87
	before	1.04	588.33
	before	1.68	621.51
	before	1.57	605.8
	before	1.62	621.74
	equals	1.26	616.8

Fig. 8. KLW1-Explore after clicking on the "Relations Data" link. The current TIRP is presented as a half-matrix, which represents the conjunction of the temporal relations among its symbolic time intervals.

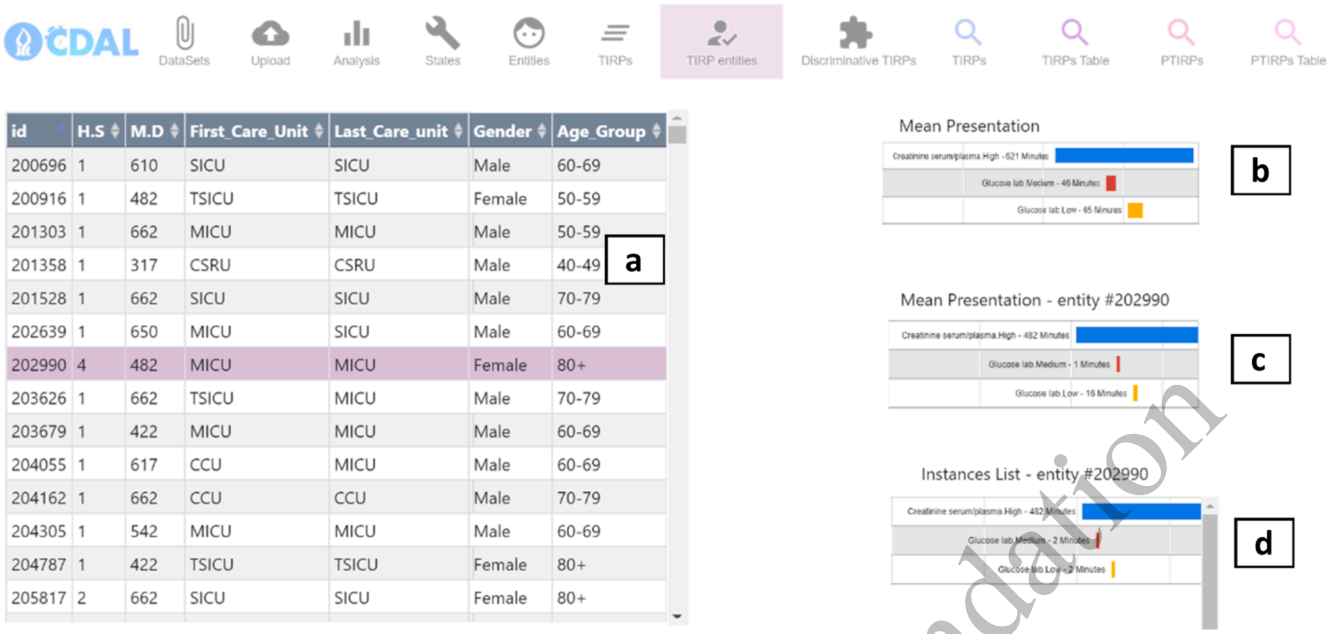
table, like in KLW1-Explore but, here, each TIRP is displayed with its metric values in each of the populations.

For comparison, the interface presents the TIRP metric values for each population in several informative ways. For instance, to show the MD of each symbolic time interval of the selected TIRP per population, Fig. 11c displays a timeline diagram. In addition, Fig. 11d displays a data table that provides the selected TIRP metric values per population.

To provide a visual display for each of the TIRPs' metrics and to allow for comparison within the populations, several comparison visualizations are provided. In Fig. 11e, the bar graph shows the vS of the selected TIRP in each of the populations, relative to the minimal vS threshold.

To demonstrate the MHS and MMD metric values of the selected TIRP, Fig. 11f and g, respectively, also display a confidence interval for each metric in each population to determine whether their difference is significant (when the ranges do not overlap). Finally, Fig. 11h shows a pie chart that presents the demographic distribution, similar to KLW1-Explore, but for each of the populations.

To present the patterns that are most discriminative, or different in the values of their metrics among the populations, the table in Fig. 11a includes a score that incorporates the differences, or deltas, of these values. Formula 1 describes the TIRP's discriminative score and its components. The components consist of the metric values' differences of the TIRPs in the populations, which are referred to as deltas, the


CDAL
DataSets
Upload
Analysis
States
Entities
TIRPs
TIRP entities
Discriminative TIRPs
TIRPs
TIRPs Table
PTIRPs
PTIRPs Table
a
b
c
d
Mean Presentation
Creatinine serum/plasma.High - 621 Minutes
Glucose lab.Medium - 46 Minutes
Glucose lab.Low - 65 Minutes
Mean Presentation - entity #202990
Creatinine serum/plasma.High - 482 Minutes
Glucose lab.Medium - 1 Minutes
Glucose lab.Low - 16 Minutes
Instances List - entity #202990
Creatinine serum/plasma.High - 482 Minutes
Glucose lab.Medium - 2 Minutes
Glucose lab.Low - 2 Minutes

id	H.S	M.D	First_Care_Unit	Last_Care_unit	Gender	Age_Group
200696	1	610	SICU	SICU	Male	60-69
200916	1	482	TSICU	TSICU	Female	50-59
201303	1	662	MICU	MICU	Male	50-59
201358	1	317	CSRU	CSRU	Male	40-49
201528	1	662	SICU	SICU	Male	70-79
202639	1	650	MICU	SICU	Male	60-69
202990	4	482	MICU	MICU	Female	80+
203626	1	662	TSICU	MICU	Male	70-79
203679	1	422	MICU	MICU	Male	60-69
204055	1	617	CCU	MICU	Male	60-69
204162	1	662	CCU	CCU	Male	70-79
204305	1	542	MICU	MICU	Male	60-69
204787	1	422	TSICU	TSICU	Female	80+
205817	2	662	SICU	SICU	Female	80+

Fig. 9. The supporting entities of the selected TIRP at KLW–Explore. For a selected TIRP, the interface presents its visual mean presentation (b) and its supporting entities (c). Selecting a supporting entity will present the mean presentation of the TIRP of the selected entity (c) and each of the entity's instances of that TIRP (d).

absolute subtraction of relevant metric values in the two populations, as we define clearly in Definitions 8, 9, and 10. Here, these components include the deltas for the vS (Definition 3), MHS (Definition 5), and MMD (Definition 7), but they can also be applied with other metrics in the future. The total score is calculated by the weighted sum of the metrics' deltas, based on the weights defined by the user in Fig. 11i. By default, an equal weight is given to each of the score components, and the sum of the weights is 1.

The following definitions describe the metrics' delta components that are part of the TIRP Discrimination Score in Formula 1:

$$TIRPDiscrimationScore(t) = W_{\Delta VS}{}^{*}\Delta VS + W_{\Delta MHS}{}^{*}\Delta MHS + W_{\Delta MMD}{}^{*}\Delta MMD \qquad (1)$$

Definition 8. *The Delta Vertical Support (Δ vS t) is the absolute subtraction of the vS values of a pattern t in the two populations P and Q, such that Δ vS =* $|vs(t^{P}) - vs(t^{Q})|$*, in which* $vs(t^{P})$ *and* $vs(t^{Q})$ *are the vS values of pattern t in the two populations P and Q, respectively.*

Definition 9. *The Delta Mean Horizontal Support (Δ MHS, t) is the absolute subtraction of the MHS values of a pattern t in the two populations P and Q, such that Δ MHS =* $|MHS(t^{P}) - MHS(t^{Q})|$*, in which* $MHS(t^{P})$ *and* $MHS(t^{Q})$ *are the MHS values of pattern t in the two populations P and Q, respectively.*

Definition 10. *The Delta Mean Mean Duration (Δ MMD, t) is the absolute subtraction of the MMD of a pattern t in the two populations P and Q, such that Δ MMD =* $|MMD(t^{P}) - MMD(p^{Q})|$*, in which* $MMD(t^{s})$ *and* $MMD(t^{Q})$ *are the MMD values of pattern t in the two populations P and Q, respectively.*

3.1.3. KLW1-Query & KLW2–Query: Visualization of interactive TIRPs' Query and retrieval

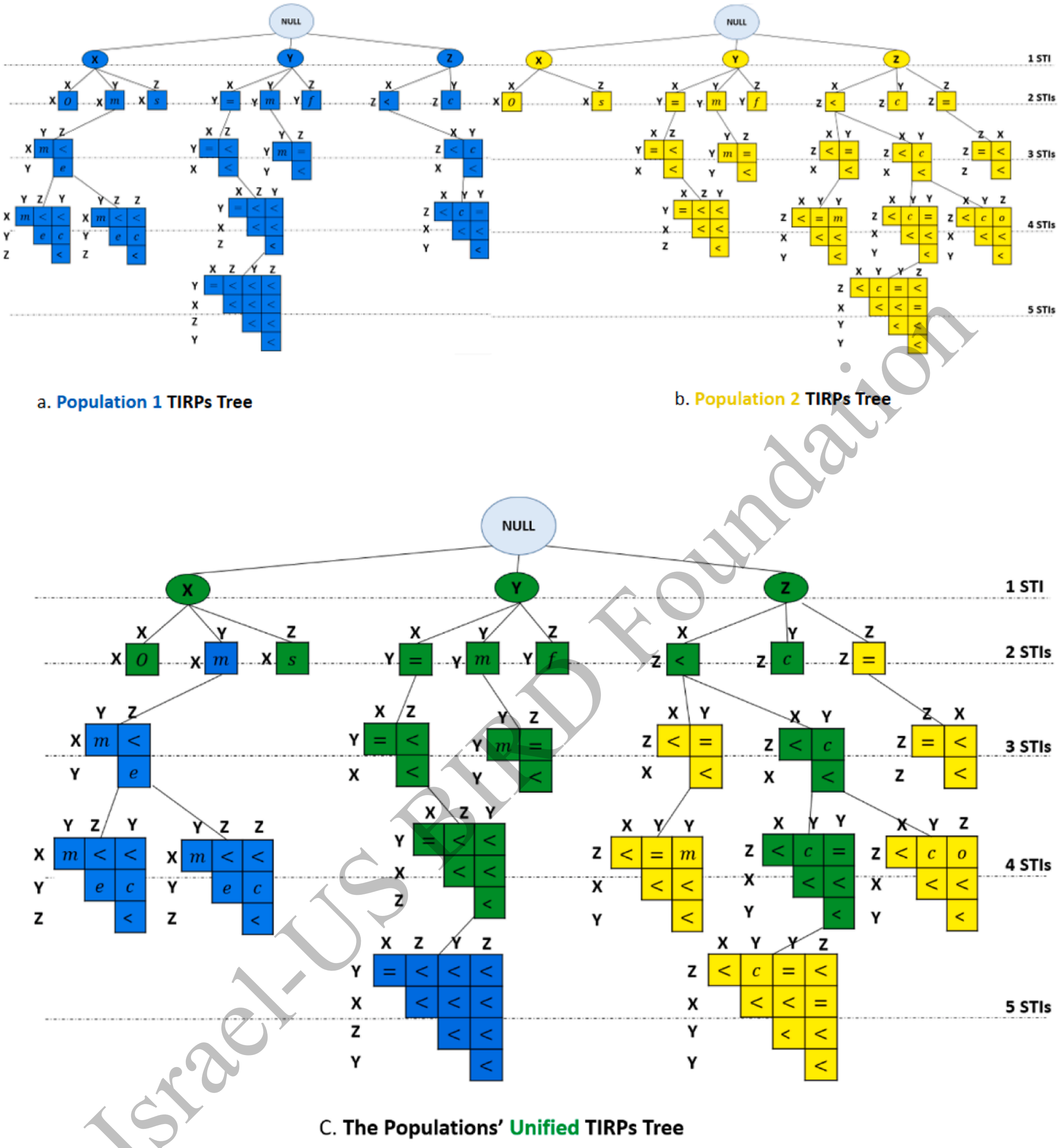
Instead of exploring the enumeration tree of TIRPs, an alternative way to explore them is by querying the enumeration TIRP tree based on some criteria, and retrieving the suitable TIRPs. In this section, two types of relevant query interfaces are introduced. Both types have the same query interface (see Fig. 12), but each queries a different enumeration tree (whether a single-population TIRP tree, or a two-population unified TIRP tree). For each query interface (KLW1-Query and KLW2–Query), two types of views were created in order to display the patterns that met the query conditions. The first type is a tabular view, and the second type is a graphical view. Fig. 12 presents the query interface that enables querying an enumeration TIRP tree, whether discovered from a single population, or a unified TIRP tree, for specific queries according to the following options.

3.1.3.1. Querying the enumeration TIRP tree. To query for TIRPs, the interface allows specification of the symbols of the TIRPs of interest according to its symbols, whether they be first (starts with), intermediate (contains), or last (ends with) as shown in Fig. 12 a, b, and c, respectively. For example, a user can search for all the TIRPs that end with the symbol "Low Heart Rate".

Moreover, the components in Fig. 12d allow a user to limit the query for patterns having metric values (VS MHS, and size) and retrieve only the patterns whose metric values are in the specified range. Alternatively, instead of presenting the user with all the symbols to choose from, one may suggest using a textbox, in which the user can enter the letters of the desired symbol. Typically, in such datasets, the user may not even remember the names of all the symbols (note, these are not the variables, but rather their states, or gradients), and based on some discussions with potential users, it seemed less convenient and would be more favorable to choose from a list.

It may be desirable to also query for the temporal relations between the symbols; however, this is challenging since the number of temporal relations grow exponentially with the size of a TIRP in general, and also within the TIRP, which means seeking multiple relations for each symbolic time interval. For example, even between the third and fourth STIs there are already-three relations to ask about. This is impractical and unnecessary. Moreover, even for the symbols, it is not optional to query for all of them—only on the first, any intermediate, and the last. Typically, the number of returned TIRPs is not large, and the user can choose the resulting TIRPs according to their temporal relations.

3.1.3.2. Tabular view of TIRP results. To display the retrieved TIRPs, we first present the tabular view in Fig. 13 that was used as a baseline in our user study, and then the novel graph-based view. The tabular view is shown in Fig. 13a, in which each retrieved TIRP is presented in a row that contains its size, symbols, the temporal relations among them in a list, and the values of its metrics: vS MHS, or MMD (when the unified



a. Population 1 TIRPs Tree

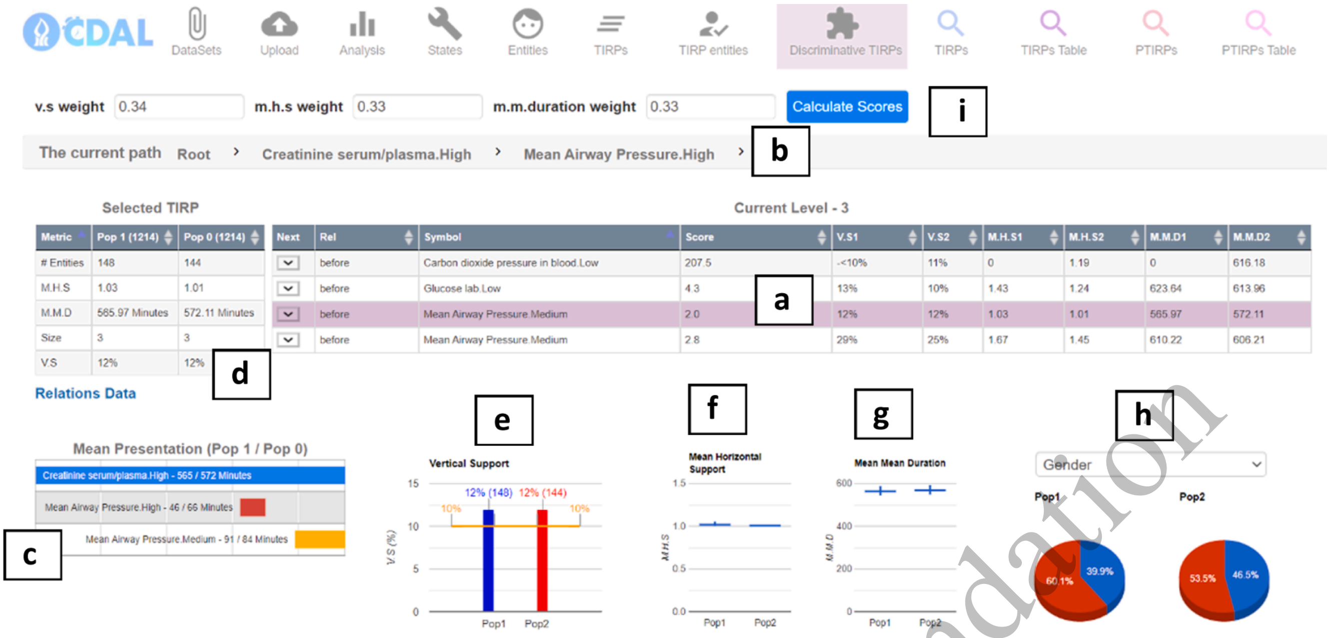
b. Population 2 TIRPs Tree

C. The Populations' Unified TIRPs Tree

Fig. 10. The two populations' TIRP enumeration trees, at the top in blue or yellow (a and b), are unified into the unified TIRP tree, shown at the bottom (c). In the unified TIRP tree, TIRPs that appear only in one of the populations are in blue or yellow, and those that appear in both are in green. Note, a TIRP that appears only in one population (blue or yellow) doe mean that there are no instances of this TIRP in the other population, but rather that it is not above the minimal vertical support in that population. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

TIRP tree of the two populations is queried, the metrics are shown for each population separately, and the discriminative score). The number of the retrieved TIRPs is displayed above the table, and the table can be sorted by any of its columns. Pressing on a row (TIRP) of interest, Fig. 13b displays a visual presentation of the specific TIRP. Moreover, the "Explore TIRP" button in Fig. 13c will move the user to the KLW1-Explore component at the location of the selected TIRP in the TIRP enumeration tree in order to get more information on the selected TIRP, as described in KLW1-Explore. This option works the same on both query interfaces.

3.1.3.3. Graphical view of TIRP results. In this study, we introduce a novel graph-based TIRP results view in which the TIRPs are presented on a two-dimensional bubble chart, and the axes represent their metric values as well as dot size and color depending on the use, as we introduce here. The use of the graph-based view is meaningfully different for single population querying, and when used for querying the unified TIRP tree of the two populations.



v.s weight 0.34 m.h.s weight 0.33 m.m.duration weight 0.33

The current path Root › Creatinine serum/plasma.High › Mean Airway Pressure.High ›

Selected TIRP

Metric	Pop 1 (1214)	Pop 0 (1214)
# Entities	148	144
M.H.S	1.03	1.01
M.M.D	565.97 Minutes	572.11 Minutes
Size	3	3
V.S	12%	12%

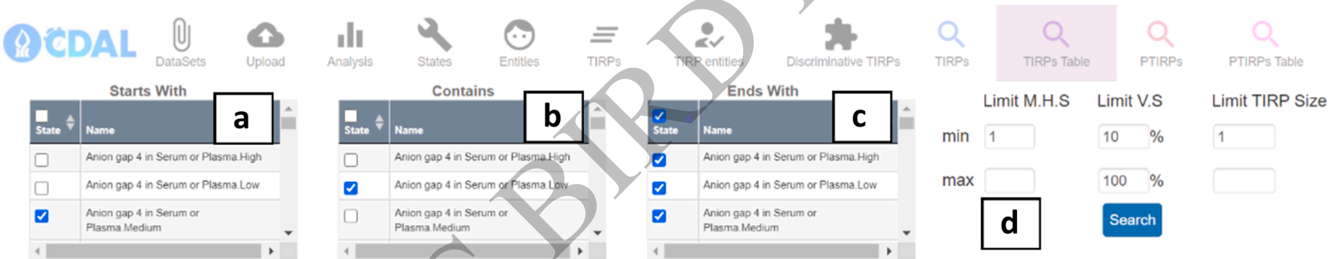
Current Level - 3

Next	Rel	Symbol	Score	V.S1	V.S2	M.H.S1	M.H.S2	M.M.D1	M.M.D2
⌄	before	Carbon dioxide pressure in blood.Low	207.5	-<10%	11%	0	1.19	0	616.18
⌄	before	Glucose lab.Low	4.3	13%	10%	1.43	1.24	623.64	613.96
⌄	before	Mean Airway Pressure.Medium	2.0	12%	12%	1.03	1.01	565.97	572.11
⌄	before	Mean Airway Pressure.Medium	2.8	29%	25%	1.67	1.45	610.22	606.21

Mean Presentation (Pop 1 / Pop 0)

Creatinine serum/plasma.High - 565 / 572 Minutes
Mean Airway Pressure.High - 46 / 66 Minutes
Mean Airway Pressure.Medium - 91 / 84 Minutes

Fig. 11. KLW2–Explore: two population-based discriminative TIRP exploration, comparison, and visualization page. In the interface at level three of the unified tree showing a 3-sized TIRP, the dynamic table (a) presents the optional fourth symbol extensions of the current TIRP. The path in the tree till the current TIRP (also its sequence of symbols) is shown at the top (b). A visual illustration of the current TIRP along a timeline diagram with its STIs' mean duration of each population is displayed (c). Its metrics' values per population are shown informatively (d) and using charts (e, f, g), as well as some demographic data distribution of its supporting entities (h).



Starts With

State	Name
☐	Anion gap 4 in Serum or Plasma.High
☐	Anion gap 4 in Serum or Plasma.Low
☑	Anion gap 4 in Serum or Plasma.Medium

Contains

State	Name
☐	Anion gap 4 in Serum or Plasma.High
☑	Anion gap 4 in Serum or Plasma.Low
☐	Anion gap 4 in Serum or Plasma.Medium

Ends With

State	Name
☑	Anion gap 4 in Serum or Plasma.High
☑	Anion gap 4 in Serum or Plasma.Low
☑	Anion gap 4 in Serum or Plasma.Medium

	Limit M.H.S	Limit V.S	Limit TIRP Size
min	1	10 %	1
max		100 %	

Fig. 12. The query interface enables querying of TIRPs for their symbols, whether for a symbol that starts (a), is contained (b), or is at the end (c). In addition, it enables limiting their metrics' values (d), including MHS, vS and TIRP size.

3.1.3.4. KLW1-query: one-population query interface. For the retrieval of TIRPs from a single population, the main requirement was to enable a good illustration of the TIRPs to be achieved based on their metric values. Fig. 14a presents the graphical view of the retrieved TIRPs presented by the bubble chart.

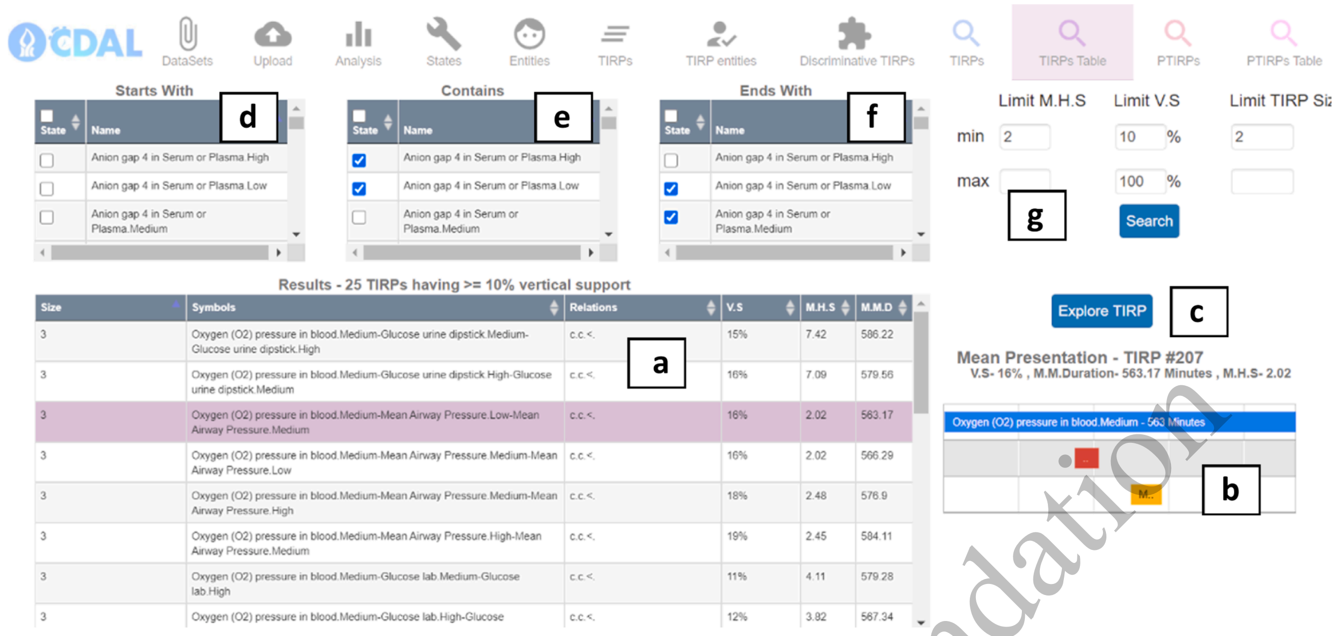
Each bubble on the graph represents a TIRP that was found suitable according to the search query constraints. The y-axis represents the vS values, while the x-axis represents the MHS values, and each TIRP is located on the chart according to its vS and MHS values. In addition, to represent the TIRP's MMD, each bubble (i.e., a TIRP) has a different color tone. Darker bubbles represent TIRPs that have longer MMDs.

In addition, the number of the retrieved TIRPs, given the minimal vS is mentioned in the title of the graph. As a result, the distribution of the TIRPs can be seen on the axes based on their metric values. For example, it can be seen in Fig. 14a, that for the presented population, most TIRPs have a relatively low MHS, whereas some of the TIRPs have a relatively high MHS, but the frequency of these patterns (their VS) is low. As a result, it can be concluded that, in this population, patterns that repeat many times are rare, and most of the patterns occur one to three times on average. In addition, most of the dark TIRPs are located on the left, and as a result, it can also be concluded that TIRPs with a long MMD usually repeat a small number of times.

Moreover, when the user detects a TIRP that he wants to explore, he can click on the bubble (TIRP) of interest, which will turn red upon selection. Fig. 14b displays a visual presentation of the TIRP with its metric values and STI MDs.

3.1.3.5. KLW2-query: two-populations query interface. In the case of two populations, the graph-based presentation of TIRPs is different, mainly intending to easily illustrate the differences in TIRP appearance in the two populations. Thus, here, the main idea and novelty is in using each axis to represent one of the populations and to locate the TIRPs, as bubbles on the graph axes according to their metric values in each of the populations. Note, this is a different use of the axes than in the KLW1-Query, in which each axis is used for a different metric. Here, each axis is used for a different population, while both axes show the *same* metric, typically the vS.

Fig. 15a presents the graphical bubble chart view for the results of the two populations' query interface which queries the unified TIRP tree, in which each bubble on the graph represents a TIRP that is relevant to the query. The number of retrieved patterns appears in the graph title as well. However, as mentioned, the axes describe the vS of the TIRPs in each of the populations; thus, the y-axis represents the vS of each TIRP (i.e., its frequency) in one of the populations, and the x-axis represents the vS of each TIRP in the other population. Consequentially, TIRPs located on the graph's diagonal have similar frequencies in both



Starts With

State	Name
□	Anion gap 4 in Serum or Plasma.High
□	Anion gap 4 in Serum or Plasma.Low
□	Anion gap 4 in Serum or Plasma.Medium

Contains

State	Name
☑	Anion gap 4 in Serum or Plasma.High
☑	Anion gap 4 in Serum or Plasma.Low
□	Anion gap 4 in Serum or Plasma.Medium

Ends With

State	Name
□	Anion gap 4 in Serum or Plasma.High
☑	Anion gap 4 in Serum or Plasma.Low
☑	Anion gap 4 in Serum or Plasma.Medium

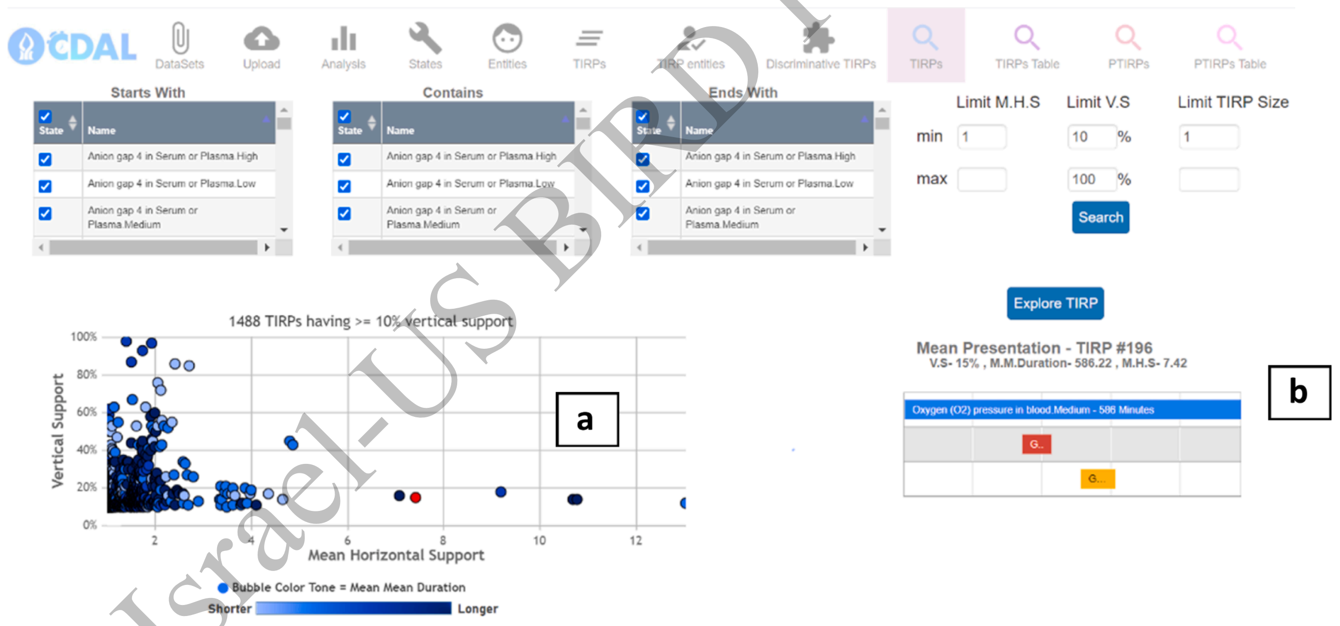
	Limit M.H.S	Limit V.S	Limit TIRP Si:
min	2	10 %	2
max		100 %	

Results - 25 TIRPs having >= 10% vertical support

Size	Symbols	Relations	V.S	M.H.S	M.M.D
3	Oxygen (O2) pressure in blood.Medium-Glucose urine dipstick.Medium-Glucose urine dipstick.High	c.c.<.	15%	7.42	586.22
3	Oxygen (O2) pressure in blood.Medium-Glucose urine dipstick.High-Glucose urine dipstick.Medium	c.c.<.	16%	7.09	579.56
3	Oxygen (O2) pressure in blood.Medium-Mean Airway Pressure.Low-Mean Airway Pressure.Medium	c.c.<.	16%	2.02	563.17
3	Oxygen (O2) pressure in blood.Medium-Mean Airway Pressure.Medium-Mean Airway Pressure.Low	c.c.<.	16%	2.02	566.29
3	Oxygen (O2) pressure in blood.Medium-Mean Airway Pressure.Medium-Mean Airway Pressure.High	c.c.<.	18%	2.48	576.9
3	Oxygen (O2) pressure in blood.Medium-Mean Airway Pressure.High-Mean Airway Pressure.Medium	c.c.<.	19%	2.45	584.11
3	Oxygen (O2) pressure in blood.Medium-Glucose lab.Medium-Glucose lab.High	c.c.<.	11%	4.11	579.28
3	Oxygen (O2) pressure in blood.Medium-Glucose lab.High-Glucose	c.c.<.	12%	3.82	567.34

Mean Presentation - TIRP #207
V.S- 16% , M.M.Duration- 563.17 Minutes , M.H.S- 2.02

Fig. 13. The tabular view of the KLW1-Query interface results presented in a table (a), where each row represents a retrieved TIRP. A visual presentation of the selected TIRP with information about it is displayed next to the table (b). The "Explore TIRP" button (c) will move the user to the KLW1-Explore interface at the location of the selected TIRP in the TIRP enumeration tree in order to allow more information to be retrieved about the selected TIRP. Here, we have limited the query to retrieve only patterns whose symbols match the constraints (d, e, f) and whose mean horizontal support and size are greater than or equal to 2 (g).



Starts With

State	Name
☑	Anion gap 4 in Serum or Plasma.High
☑	Anion gap 4 in Serum or Plasma.Low
☑	Anion gap 4 in Serum or Plasma.Medium

Contains

State	Name
☑	Anion gap 4 in Serum or Plasma.High
☑	Anion gap 4 in Serum or Plasma.Low
☑	Anion gap 4 in Serum or Plasma.Medium

Ends With

State	Name
☑	Anion gap 4 in Serum or Plasma.High
☑	Anion gap 4 in Serum or Plasma.Low
☑	Anion gap 4 in Serum or Plasma.Medium

	Limit M.H.S	Limit V.S	Limit TIRP Size
min	1	10 %	1
max		100 %	

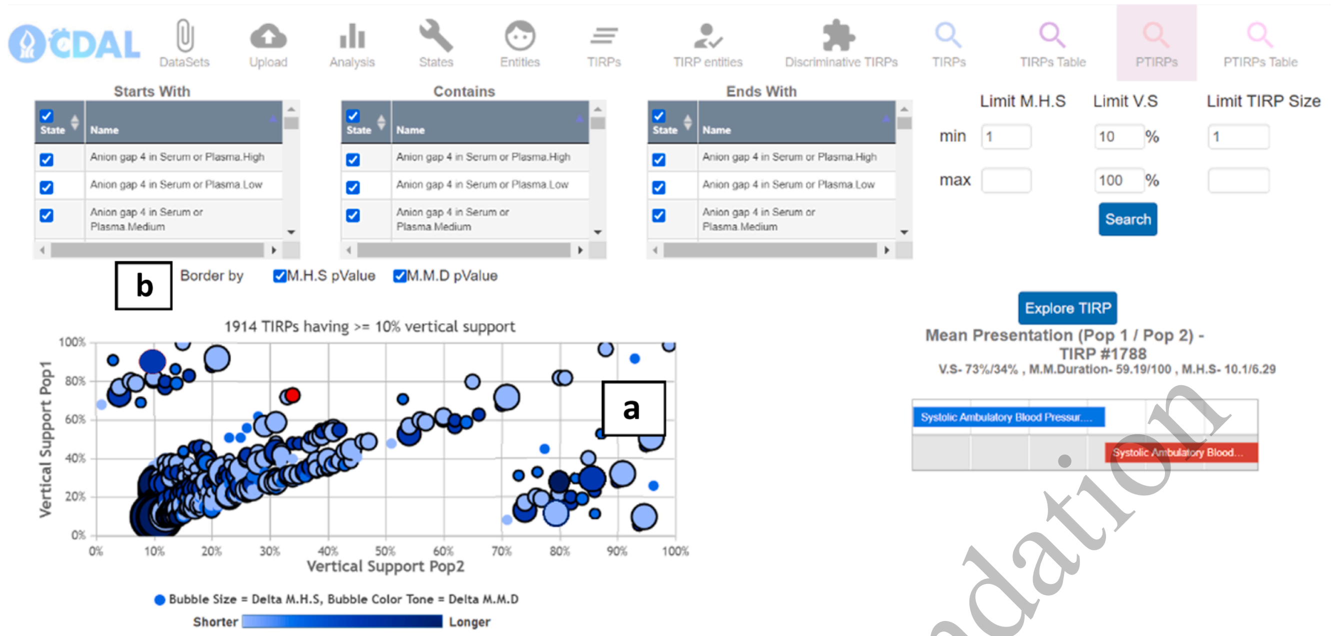
Mean Presentation - TIRP #196
V.S- 15% , M.M.Duration- 586.22 , M.H.S- 7.42

Fig. 14. The graphical view of the KLW1-Query results presented by a bubble chart (a), where each bubble on the graph represents a TIRP. The y-axis represents the vertical support of each TIRP (i.e., its frequency), and the x-axis represents its mean horizontal support. Each bubble (i.e., TIRP) has a different color tone representing its MMD. Selecting a TIRP on the graph will change the bubble color to red, and a visual presentation of the TIRP with information about it will be displayed next to the graph (b). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

populations, while the TIRPs that are unique to either of the populations are at the top left or bottom right, having high frequency in one population and low frequency in the other. In addition, each bubble (i.e., TIRP) has a different color tone that represents the delta MMD (Definition 10) of the TIRP among the populations. Darker TIRPs represent those whose difference in the MMD support among the populations is larger. Moreover, each bubble has a different size that represents the Δ MHS (Definition 9) of the TIRP among the populations. Larger bubbles represent TIRPs whose difference in the MHS among the populations is larger.

Another dimension of each bubble is its border thickness. As can be seen in Fig. 15a, some bubbles have a bold border, whereas others do not. The thickness of the border of each TIRP is determined by whether the difference in the MHS or in the MMD between the populations is significant.

In addition, the component in Fig. 15b allows a choice of which metrics (MHS or MMD) will affect the thickness of the border. For example, if the user wants only TIRPs with a significant difference in



Starts With

☑ State	Name
☑	Anion gap 4 in Serum or Plasma.High
☑	Anion gap 4 in Serum or Plasma.Low
☑	Anion gap 4 in Serum or Plasma.Medium

Contains

☑ State	Name
☑	Anion gap 4 in Serum or Plasma.High
☑	Anion gap 4 in Serum or Plasma.Low
☑	Anion gap 4 in Serum or Plasma.Medium

Ends With

☑ State	Name
☑	Anion gap 4 in Serum or Plasma.High
☑	Anion gap 4 in Serum or Plasma.Low
☑	Anion gap 4 in Serum or Plasma.Medium

	Limit M.H.S	Limit V.S	Limit TIRP Size
min	1	10 %	1
max		100 %	

☑M.H.S pValue ☑M.M.D pValue

Mean Presentation (Pop 1 / Pop 2) -
TIRP #1788
V.S- 73%/34% , M.M.Duration- 59.19/100 , M.H.S- 10.1/6.29

Fig. 15. The graph-based KLW2-query interface presents the query results in a bubble chart (a), where each bubble on the graph represents a TIRP. The x- and y-axes represent the vertical support of each of the populations of the TIRPs (i.e., their frequency). Each bubble is a TIRP, and beyond is its location determined by its vertical support in each of the populations. Its size represents its delta mean horizontal support, and its color tone its ΔMMD, both among the populations. Having a bold border means that the difference in the horizontal support values, or in the MD values (b), in the populations is significantly different.

MHS between the populations to affect the thickness of the border, he can select only the MHS *p*-value check box so that bubbles with a bold border will indicate only patterns whose difference in the MHS metric between the populations is significant. In order to determine if the difference in one of the metrics is significant, we performed a *t*-test that checked whether the difference in HS or MD values of the supporting entities of each TIRP in each population were significantly different. If the *t*-test resulted in a *p*-value that was lower than 0.05, the difference was considered significant.

As mentioned, the main goal of the graphical view in Fig. 15a was to find discriminative TIRPs that are meaningfully different among the populations in the most efficient way. The graphical view emphasizes the difference in each of the patterns' metric values among the populations.

TIRPs that vary in their frequency between the populations are located at the top left or bottom right of graph, while those having a similar frequency appear on the diagonal. Moreover, TIRPs that differ in their MHS and their MMD between populations are large and darker, and if these differences are significant, they have bold borders as well.

4. Evaluation

We conducted a user study to evaluate the usability of the KLW interface. Usability was assessed in terms of the main usability criteria: effectiveness, efficiency, and satisfactionDis, i. [12], and in terms of its learnability (IEEE Std. 1061, 1998). Consequently, we evaluated the following research questions:

1. Is the system usable for the KLW's main tasks: navigation, exploration, comparison, and retrieval of the various patterns that were discovered among the populations?
2. Will the graphical-based bubble chart view of the query interfaces improve system usability relative to the tabular view?

4.1. Evaluation method

4.1.1. Sample

Twenty-five engineering students who had at least some familiarity with the concept of temporal patterns were recruited for the evaluation. Seventeen of them were undergraduate students, and eight were graduate students. The study included five female participants, and their average age was 27 (age range 20–35). Note that the intended users of the system are data scientists who may also be domain experts (e.g., physicians, researchers), however, who are familiar with both the concept of TIRPs and with the clinical domain problem and data. Hence, it is assumed that users will be able to acquire meaningful temporal knowledge and insights, once they find it. Therefore, the evaluation focused on the usability of the KLW's interface in accessing the relevant information, and whether such users can effectively and efficiently use the system to their satisfaction.

4.1.2. Training

To acquaint the participants with the KLW interface, they watched two instructional videos lasting half an hour each, before interacting with the system. The first video[1] explained the concept of temporal data, temporal abstraction, and TIRP mining with its output – the frequent enumeration TIRP tree (see Fig. 4).

The second video[2] presented the main features of the KLW interface. These included navigation in the frequent enumeration TIRP tree to specific TIRPs, an explanation of the various components that showed the metrics of the TIRPs, and instruction on the various query interfaces. In addition to watching the instructional videos, the participants experimented with the system by performing three navigation tasks in which they looked for specific patterns.

We gave the participants the option to continue exploring the system beyond the three practice navigation tasks to gain more confidence before starting the questionnaire. However, all participants chose to

[1] https://www.ise.bgu.ac.il/cdalab/video/KL_and_TIRPs.mp4.

[2] https://www.ise.bgu.ac.il/cdalab/video/KLW_Introduction.mp4.

start the questionnaire directly after the three tasks. Following the practice, all participants rated the level of ease of use of the KLW system as 3 or 4 on a 0–4 scale (4 = very easy to use).

4.1.3. Datasets

In order to perform the user study and evaluate the use of the novel pattern visualization of the KLW, we needed datasets that included two populations for the KLW2, while for the KLW1, either one of the populations or the entire set of two populations was sufficient. For that, two datasets from the Intensive Care Unit domain were used, focusing on the prediction of two types of outcomes: Sepsis and Acute Hypertensive Episode (AHE). Both datasets were created based on the MIMIC database, and each was created using case controls relative to an earlier event. In both datasets, the earlier event, in our case, was the time of admission to the ICU. Two types of patients that constructed the classes existed in the datasets. The first class, the cohort, included patients who had the outcome in their records, whereas the second class, the controls, included patients who did not have the outcome in their data. The classification task was to predict, based on a fixed time duration relative to the admission, whether the patient would or would not have the outcome during their stay.

The study data were acquired from the publicly available Medical Information Mart for Intensive Care (MIMIC)-III version 1.4 dataset [20]. It comprises over 53,423 ICU stays for more than 40,000 patients admitted to the Beth Israel Deaconess Medical Center (BIDMC) in Boston, MA, between 2001 and 2012. The database contains data from two different critical care information systems: Care-Vue (Philips) and Metavision (iMDSoft), from 2001 to 2008 and 2008 to 2012, respectively.

4.1.3.1. Sepsis. A repository was built for the analysis of ICU-acquired sepsis identification methods in an electronic health records database [20], that was used for the identification of the case-patients and outcome onset time within the MIMIC database. We used 1,240 entities from each class: cohort and controls. A single time-window from each entity was taken, where each time-window duration was 12 h. The median age of the cohort was 65, of whom 59.3 % were men with a median ICU length of stay of 11.6 days. The median age of the controls was 66, of whom 54.1 % were men with a median ICU length of stay of 3.3 days.

4.1.3.2. Acute Hypertensive Episode (AHE). AHE is a long-term medical condition in which severe elevations in blood pressure are likely to cause damage to one or more organ systems. In this dataset, we used 600 entities from each class: cohort and controls. From the cohort, a single time-window from each entity was taken, and from the controls, five time-windows from each entity were taken.

Here, each time-window duration was 6 h. The median age of the cohort was 67, of whom 57.66 % were men with a median ICU length of stay of 4.56 days. The median age of the controls was 66, of whom 57.21 % were men with a median ICU length of stay of 3.24 days.

Both datasets were mined using the KarmaLego algorithm to discover the frequent pattern tree of each population, i.e., class (cohort or controls). The minimum vS threshold that was used was 10 %. After the frequent pattern trees for each class were discovered, they were unified into a unified tree (see Fig. 10) to allow for the KLW2 components.

4.1.4. Procedure

Due to social-distancing restrictions during the coronavirus outbreak, the user study was conducted online via Zoom. Each participant was scheduled for an online session that lasted about an hour, in which they completed a series of tasks. The study had two parts of 20 tasks each, and the tasks in both parts were structurally equivalently. All users performed the tasks on the Sepsis dataset first and then on the AHE dataset. Following each task, participants rated the ease of performing it. As mentioned in Section 3.1.3, the results in each of the KLW's query interfaces (KLW1-Query and KLW2-Query) can be presented in two views: graphical (using a bubble chart) and tabular. To compare the usability of these views, the tasks for both views were the same. The order in which the views were presented was counterbalanced across participants. Half of the participants started with three tabular view tasks and then continued with three graphical view tasks, while the other half were presented with the opposite order.

4.1.5. Evaluation metrics and dependent measures

Per the conventional practice in usability evaluations, the dependent measures included three elements: *effectiveness*, measured as the percentage of correct answers (each answer was classified as 1 if it was answered correctly and 0 otherwise); *efficiency*, measured in terms of time to complete the tasks (in seconds); and a subjective component, operationalized as ratings of the ease in which each task was completed on a 5-point scale (0 = very difficult, 4 = very easy).

4.1.6. Study design

The study included four types of tasks (see task details in Appendix 3): general, navigation, tables, and graphs. General tasks were performed on two system components: KLW1-Explore and KLW2-Explore (Sections 3.2.1 and 3.2.2, respectively). The first four general tasks were performed on KLW1-Explore, while a single navigation task and additional three general tasks were performed on KLW2-Explore.

The latter 12 tasks involved presenting data in response to queries by either tabular or graphical formats. These tasks were performed on the system's query components (six one population tasks on the KLW1-Query and another six on the KLW2-Query).

Table tasks were conducted on the table view interface of KLW1-Query and KLW2-Query, and graph tasks were conducted on the graphical view interface of these components. The KLW1-Query included tasks that required finding patterns in one population according to their metrics and symbols. The KLW2-Query tasks required finding the patterns that had the greatest distinguishing power between two classes.

5. Results

We evaluated the participants' performance in terms of effectiveness (i.e., accuracy), efficiency (execution time), and subjective ease of performing the tasks. We performed separate analyses for the first eight tasks (general and navigation) and for the latter 12 tasks (one-population and two-population queries).

5.1. General and navigation tasks

For these tasks, we performed a two-way repeated measures ANOVA with dataset (first or second) and task type (general *vs* navigation) as the independent factors.

Accuracy. The answer to each task was classified as 1 if correct and 0 otherwise. Task accuracy was very high ($M = 0.98$, $S.D. = 0.11$). None of the independent variables including their interaction had a significant effect on task accuracy (see ANOVA Table 1.1 in Appendix 1).

Task duration. Average task duration was 36.1 s ($S.D. = 33.3$). Only task type had a significant main effect on task duration ($F = 4.09$, $df = 1{,}24$, p = 0.05, see ANOVA Table 1.2 in Appendix 1).

The navigation task ($M = 37.8$ s, $S.D. = 34.7$) was on average significantly slower than the general tasks ($M = 31.2$ s, $S.D. = 20.1$ s). There was no main effect for dataset ($p = 0.51$) or for the interaction of task type with the dataset ($p = 0.08$).

Subjective ease of task performance. Subjective reports of ease of task completion were high ($M = 3.57$, $S.D. = 0.77$, on a 0–4 scale). The only statistically significant effect on the ease of task execution was by dataset ($F = 11.76$, $df = 1{,}24$, $p < 0.001$, see ANOVA Table 1.3 in Appendix 1). Completing tasks of the second dataset ($M = 3.7$, $S.D. = 0.66$)

were perceived as easier than tasks of the first dataset ($M = 3.55$, *S.D.* $= 0.01$). This finding is to be expected due to the experience participants gained with the system. However, it should be noted that even the first dataset's tasks were perceived as easy to complete. There was no main effect of task type and no interaction effect ($p = 0.06$ and 0.07, respectively).

In conclusion, we found that performance on the general and navigation tasks was effective and quite efficient. A small improvement was felt for subjective ease of task performance due to experience in system usage, albeit initial ratings on this measure were high to begin with.

5.2. Query tasks using tabular and graphical-based bubble chart views

For each KLW1-Query and KLW2-Query, we performed a two-way repeated measures ANOVA on the query tasks, with view (table *vs* graph) and dataset (first or second) as within-subjects factors.

5.2.1. KLW1-Query results

Accuracy. Task accuracy was very high ($M = 0.96$, *S.D.* $= 0.17$). We found no main or interaction effect on accuracy, perhaps due to a ceiling effect (see ANOVA Table 2.1.1 in Appendix 2 and Tables 2.1.4–2.1.6 in Appendix 2).

Task duration. Average task duration was 44.29 s (*S.D.* $= 28$). Fig. 16 shows the mean and the 95 % confidence interval of task duration broken down by dataset and view.

The dataset had a significant ($F = 5.16$, $df = 1{,}24$, $p = 0.03$) effect on task duration (see ANOVA Table 2.1.2 in Appendix 2 and Tables 2.1.4–2.1.6 in Appendix 2). The mean duration of the first dataset ($M = 46.8$ s, *S.D.* $= 30.6$) was significantly higher than the MD of the second dataset ($M = 41.8$ s, *S.D.* $= 25.0$). View type did not have a significant effect, nor was there an interaction effect ($p = 0.79$ and 0.89, respectively) on task duration.

Subjective ease of task performance. Mean perceived ease of task execution was high ($M = 3.42$, *S.D.* $= 0.79$, see Fig. 17), although somewhat lower than that of the first eight tasks. Here, again, the only significant effect was associated with dataset ($F = 9.06$, $df = 1{,}24$, $p = 0.01$; see ANOVA Table 2.1.3 in Appendix 2 and Tables 2.1.4–2.1.6 in Appendix 2). The 12 queries on the second dataset ($M = 3.53$, *S.D.* $= 0.72$) were significantly easier to perform than those on the first dataset ($M = 3.31$, *S.D.* $= 0.86$). There was no main effect of view type and no interaction effect ($p = 0.74$ and 0.52, respectively).

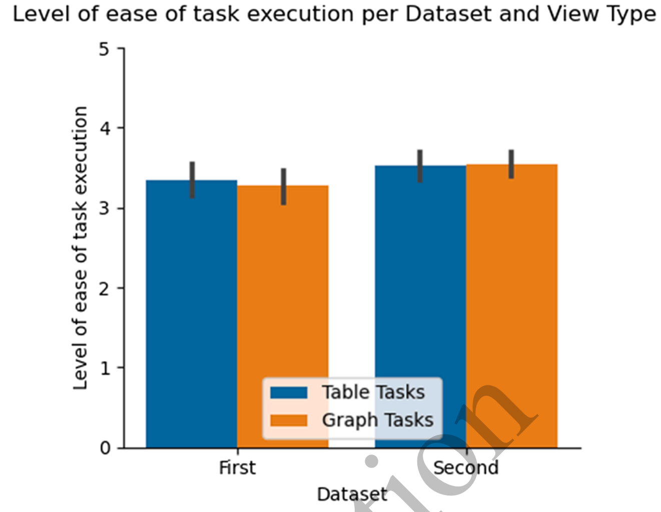


Fig. 17. The mean and confidence interval task level of ease of the factors: dataset and view type. The tasks on the second dataset were significantly easier to perform than those on the first dataset.

5.2.2. KLW2-Query results

Accuracy. The participants were less accurate in this part of the system relative to the one-population part ($M = 0.86$, *S.D.* $= 0.34$, see Fig. 18). View type had a significant effect on task accuracy ($F = 12.83$, $df = 1{,}24$, $p < 0.001$; see ANOVA Table 2.2.1 in Appendix 2 and Tables 2.2.4–2.2.6 in Appendix 2). The accuracy of the graph view ($M = 0.96$, *S.D.* $= 0.2$) was significantly higher than the accuracy of the table view ($M = 0.77$, *S.D.* $= 0.42$). There was no main effect for dataset ($p = 0.43$) or for the interaction of view type with dataset ($p = 0.6$).

Task duration. On average, the participants completed each task in 52.6 s (*S.D.* $= 48.5$, see Fig. 19). Tasks took somewhat longer to complete than tasks in the KLW1-Query module, a result that reflects the higher complexity of comparing two populations versus finding information in one population. The dataset had a significant ($F = 56.53$, $df = 1{,}24$, $p < 0.001$) effect on task duration (see ANOVA Table 2.2.2 in

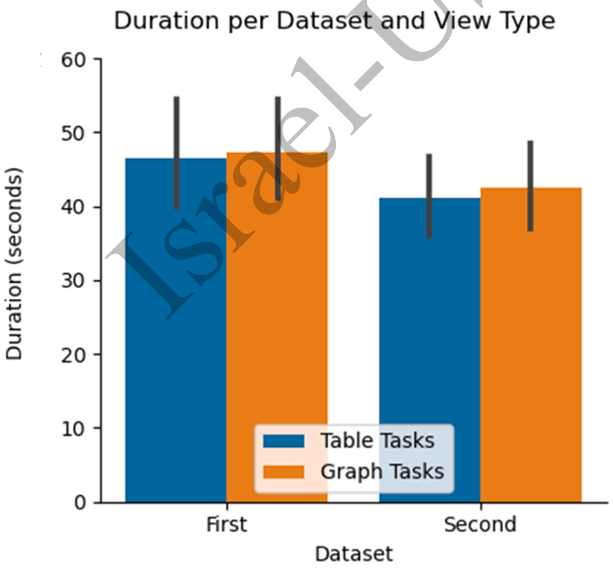


Fig. 16. The mean and confidence interval task duration of the factors: dataset and view type. The mean duration of the first dataset was significantly higher than the MD of the second dataset and View type did not have a significant effect.

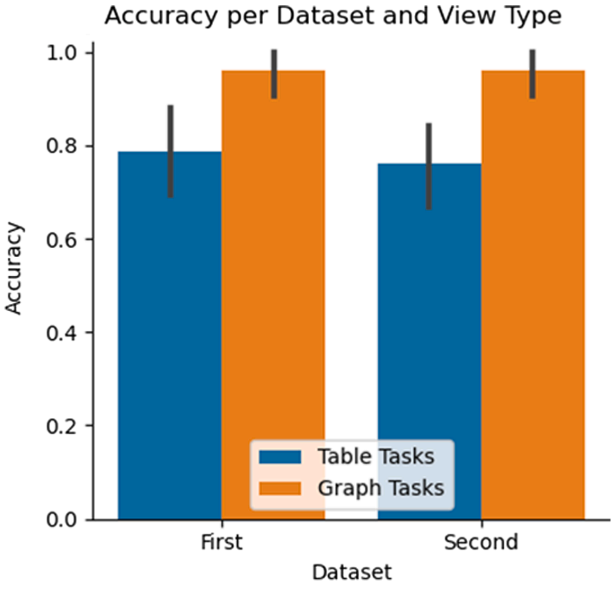


Fig. 18. The mean and confidence intervals of the task accuracy of the factors: dataset and view type. The accuracy of the graph view was significantly higher than the accuracy of the table view and Dataset did not have a significant effect.

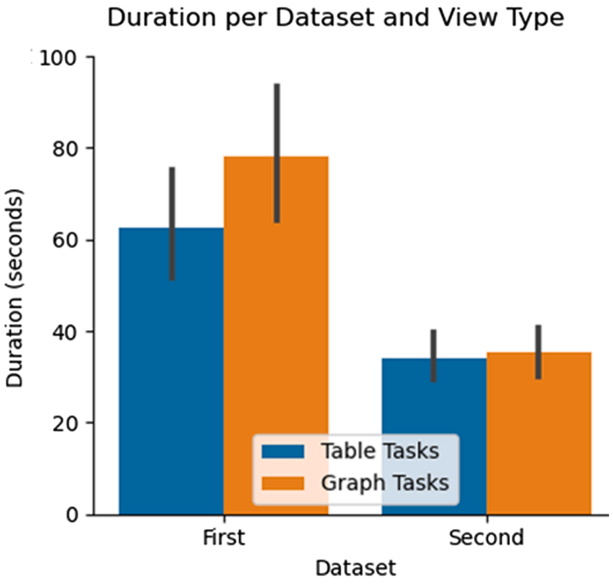


Fig. 19. The mean and confidence intervals of the task duration of the factors: dataset and view type. The tasks on the first dataset took considerably longer than the durations on the second dataset and there were no significant effects of view type.

Appendix 2 and Tables 2.2.4–2.2.6 in Appendix 2). The query tasks on the first dataset ($M = 70.4$ s, $S.D. = 59.2$) took considerably longer than the durations on the second dataset ($M = 34.8$ s, $S.D. = 23.89$). There were no significant effects of view type or of the interaction on task duration ($p = 0.08$ and 0.07, respectively).

Subjective ease of task performance. Participants felt that the KLW2-Query tasks were not as easy as the other tasks ($M = 2.62$, $S. D. = 3.0$, see Fig. 20). There were statistically significant main effects of dataset and type of view (F = 19.27, df = 1,24, $p < 0.001$, and F = 11.03, df = 1,24, $p < 0.001$, respectively) on perceived ease of task performance (see ANOVA Table 2.2.3 in Appendix 2 and Tables 2.2.4–2.2.6 in Appendix 2). Again, the users felt that it was more difficult to carry out tasks in the first dataset ($M = 2.28$, $S.D. = 1.33$) than in the second dataset ($M = 2.97$, $S.D. = 1.27$). Graph-view tasks were perceived as more difficult than table-view tasks ($M = 2.39$, $S.D. = 1.34$, compared to $M = 2.85$, $S.D. = 1.31$, respectively). This result may be explained by the fact that a bubble chart is not a familiar view and, therefore, it is harder to use. There was no interaction effect ($p = 0.42$).

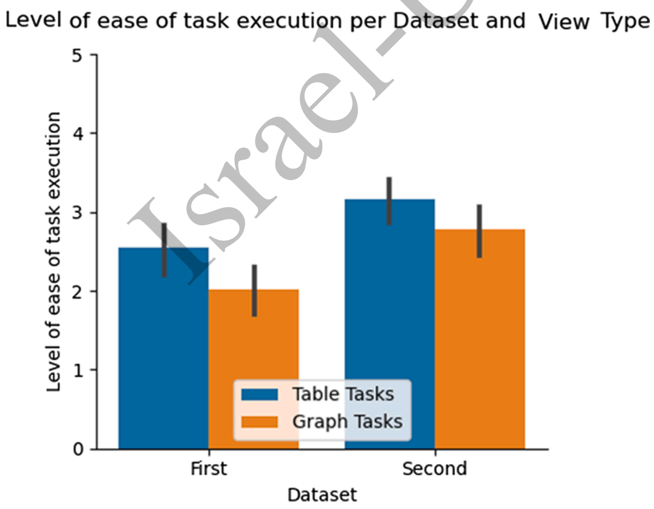


Fig. 20. The mean and confidence interval task level of ease of the factors: dataset and view type. There were significant main effects of dataset and type of view. The tasks on the second dataset were significantly easier to perform than those on the first dataset and the Graph-view tasks were perceived as more difficult than table-view tasks.

6. Discussion and conclusions

The use and advantages of temporal abstraction and TIRPs mining for longitudinal data analysis, and specifically in EHR data, has already been demonstrated for several tasks, such as classification and outcomes prediction [35]. Among its advantages is the potential explainability through the TIRPs. Explainability is an important capability in data science, specifically in the biomedical domain, since it is important to explain the models to domain experts, and also for knowledge acquisition. However, the exploration of TIRPs is challenging, and requires efficient and effective visualization of the discovered patterns, which is a relatively understudied research field.

This paper introduced KLW, a novel visual interface designed to support the exploration of TIRPs that were discovered within a single or two populations of entities (i.e., patients). The interface introduces several main visualizations: KLW1-Explore, which focuses on exploration of TIRPs that were discovered from a single population of entities, while the second visualization, KLW2-Explore, focuses on TIRPs that were discovered in two populations of entities. Similarly, another type of visualization focuses on query and retrieval of TIRPs, KLW1-Query and KLW2-Query, which enable an efficient query and retrieval of TIRPs that were discovered given a criteria defined by a user, and the results are displayed using a bubble chart which is one of the most novel contributions of the paper.

The main contributions of the paper can be summed in three. First deals with the problem of visualization of an enumeration tree of patterns. While we refer to TIRPs, the approaches introduced in this paper can be generalized and used also for other types of patterns, whether temporal patterns, such as sequential patterns, or even generally patterns, such as item sets. The main challenge is the ability to navigate in a tree structure, while presenting the information about a TIRP, and the various metrics, so the user can browse the patterns and decide where to delve in according to various aspects. Moreover, there is an illustration of the current TIRP and what are its extensions. Second is the use of a bubble chart to display efficiently the group of relevant TIRPs for a given search criteria, in a multidimensional view, which by customizing the meaning of the bubble chart axes, and the bubbles' color tone and size, to the various metrics of a TIRP (i.e., vS MHS, MMD) enables to focus fast on the TIRPs of interest. Moreover, it gives a larger picture of the discovered TIRPs distribution according to the TIRPs' metrics. For example, whether most of the TIRPs have high vS or low MHS and cet. Again, such visualization can be used also for other types of temporal patterns, and even generally patterns. The third is the visualization of TIRPs in two populations. This is a completely novel topic that was not touched in the existing literature. It allows to compare two populations, whether coming from a classification problem, or just two populations of patients according to different criteria, based on their TIRPs. Thus, it allows to identify which TIRPs have different presence in the two populations, both looking at the TIRPs as a whole and their distribution along the various metrics, and also accessing efficiently specific TIRPs of interest and getting their information in both populations. These three novel topics have meaningful implications in the ability to explore and investigate TIRPs in a single or two populations, and even used in any other types of temporal patterns, or even generally pattenrs, and in that respect the contributions here are much more meaningful.

More specifically, KLW1-Explore enables a user to browse and navigate in the TIRPs tree by navigation according to the TIRPs' properties or metrics, in a single population. This allows the patterns, along with their metrics, to be visualized, facilitating their exploration, and demonstrating demographic data distribution of the supporting entities of each TIRP in order to gain more knowledge about the discovered patterns. Similarly, KLW2-Explore enables exploration of each TIRP, along with that TIRP's metrics in two populations KLW1-Query and KLW2-Query support the ability to find TIRPs by querying for specific patterns that were discovered in a single population or in two populations. However, the current proposed interfaces are limited to the comparison of two populations, and not more, which may be needed when there are multiple populations that are investigated which we intend to explore and address in our future work. Additionally, we would like to enable to visually navigate through the TIRPs, based on the symbolic time intervals sequences to explore what leads to an event of interest, or what are the paths that can be followed by an event of interest. This requires a hole new thinking that we intend to do in the future.

To support the presentation of the queries' results (i.e., the retrieved patterns) both tabular and graphical views were implemented. to enable a data scientist or domain expert to analyze the data through the TIRPs and acquire knowledge Moreover, an additional contribution of the tool enables to characterize the TIRP's supporting population by their demographical variables' distribution, such as for example gender, age group, and other relevant properties, depending on the domain problem. This creates the ability to identify accordingly a TIRP that describes a "behavior" of a sub population along time, such as reaction to a treatment, or the sequence of their variables along time.

Finding patterns that are discriminative in two populations may explain differences in two populations, for example, of patients who have a specific outcome, or not, or the differences between a population that recovers, or not, based on their "behavior" along time. This can be achieved by the two-population exploration tools. Thus, if scientists are interested in TIRPs that have a higher frequency in one population, compared to another, they can look at the chart's bottom right or top left. Additionally, based on the metrics, especially the delta metrics, it is easy to explain why a TIRP is of interest. Once such TIRPs are identified, they can then be used for knowledge discovery – to understand how the data evolves along time, and in order to use these TIRPs in an actionable way to classify or to predict outcomes.

We conducted a user study to evaluate the visual interfaces, in which participants performed various pattern-finding tasks. The interface was evaluated in terms of three central usability dimensions: effectiveness (i.e., accuracy), efficiency (i.e., performance time), and subjective evaluation (perceived ease of task performance). We found that the interface enabled users to perform tasks in the context of exploring patterns and comparing them between populations, navigating between the different patterns, and searching for discriminative patterns. Most of the tasks were performed accurately, quickly, and easily. However, the population of users in the study were engineering students, and not clinicians, which is a limitation of the study in some extent, since they represent well potential data scientists, but not clinicians, who may be interested in using the tool. Note that such interfaces are expected to be used by data scientists, epidemiologists, or physicians who are scientists, and in practice are data scientists as well. Nevertheless, these interfaces can be used by clinicians without formal data science experience.

Users were able to learn how to use the system by watching two video tutorials. They reached satisfactory accuracy levels beginning with their first experience with the system (the first dataset), and performance time and their satisfaction was improved with the second experience (the second data set).

Query results were returned to the participants in a familiar tabular format and an unfamiliar (novel) bubble chart format. We found no outcome differences between the tabular and graphical interfaces in the one-population queries. However, interesting differences were observed for the more complex two-population queries. In that condition, the use of the graphical interface improved the accuracy relative to the use of the tabular interface, yet time duration was shorter and satisfaction was higher in the use of the tabular interface. These findings suggest a speed–accuracy tradeoff and incongruence between users' effectiveness and their self-reported ease of use. Evidently, the users either did not appreciate the contribution of the bubble chart to their performance, or were not aware of not being accurate, when using the tabular representation. Consequently, future implementations of KLW should come with clear indications to users about the advantages of the bubble-chart visualization and with recommendations regarding the benefits of using it instead of quickly coming to conclusions based on information presented in the tabular format. We intend to incorporate the KLW interfaces for visualization of TIRPs discovered from a single population, or two populations within a framework that will enable a data scientist to upload and abstract a dataset, and to mine and visualize TIRPs.

CRediT authorship contribution statement

Guy Shitrit: Formal analysis, Validation, Software, Methodology, Conceptualization. **Noam Tractinsky:** Visualization, Investigation, Formal analysis, Validation, Conceptualization. **Robert Moskovitch:** Conceptualization, Validation, Resources, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors wish to thank the Israeli Ministry of Science and Technology for assistance in funding this project with grant 8760521 and 877023, and a grant from the Israeli Prime Minister Office Cyber Bureau.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbi.2022.104169.

References

[1] W. Aigner, S. Miksch, H. Schumann, C. Tominski, Visualization of time-oriented data, Springer Science & Business Media, 2011.

[2] J.F. Allen, Maintaining knowledge about temporal intervals, Commun. ACM 26 (11) (1983) 832–843.

[3] Ayres, J., Flannick, J., Gehrke, J., Yiu, T., 2002. Sequential pattern mining using a bitmap representation. In: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, pages 429–435. ACM.

[4] Bade, R., Schlechtweg, S., Miksch, S., 2004. Connecting time-oriented data and information to a coherent interactive visualization. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pages 105–112. ACM.

[5] I. Batal, H. Valizadegan, G.F. Cooper, M. Hauskrecht, A temporal pattern mining approach for classifying electronic health record data, ACM Trans. Intell. Syst. Technol. (TIST) 4 (4) (2013) 1–22.

[6] C. Billiet, N. Van de Weghe, J. Deploige, G. De Tré, Visualizing and reasoning with imperfect time intervals in 2-D, IEEE Trans. Fuzzy Syst. 25 (6) (2016) 1698–1713.

[7] L. Chittaro, C. Combi, Visualizing queries on databases of temporal histories: new metaphors and their evaluation, Data Knowl. Eng. 44 (2) (2003) 239–264.

[8] L. Chittaro, C. Combi, G. Trapasso, Data mining on temporal data: a visual approach and its clinical application to hemodialysis, J. Visual Languages Comput. 14 (6) (2003) 591–620.

[9] C. Combi, A. Sabaini, in: Extraction, analysis, and visualization of temporal association rules from interval-based clinical data, Springer, Berlin, Heidelberg, 2013, pp. 238–247.

[10] C. Combi, B. Oliboni, Visually defining and querying consistent multi-granular clinical temporal abstractions, Artif. Intell. Med. 54 (2) (2012) 75–101.

[12] DIS, I. (2010). 9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems (formerly known as 13407). International Standardization Organization (ISO). Switzerland.

[13] Dvir O., Wolfson, P., Lovat, L., Moskovitch, R., Falls Prediction in Care Homes Using Mobile App Data Collection, Artificial Intelligence in Medicine, Minneapolis, USA, 2020.

[14] Froehlich, P., 2005. Interactive sankey diagrams. In: IEEE Symp. on Information Visualization, page 233.

[15] Gotz, D., Stavropoulos, H., 2014. Decisionflow: Visual analytics for high-dimensional temporal event sequence data. IEEE transactions on visualization and computer graphics 20(12), 1783–1792.

[16] D. Gotz, F. Wang, A. Perer, A methodology for interactive mining and visual analysis of clinical event patterns using electronic health record data, J. Biomed. Inform. 48 (2014) 148–159.

[17] N. Itzhak, A. Nagori, E. Lior, M. Schvetz, R. Lodha, T. Sethi, R. Moskovitch, Acute Hypertensive Episodes Prediction, Artificial Intelligence in Medicine, Minneapolis, USA, 2020.

[18] Harel, O., Moskovitch, R., 2021. Complete closed time intervals-related patterns mining. In: Proceedings of the AAAI Conference on Artificial Intelligence, 35(5), 4098-4105.

[19] F. Hoppner, Time series abstraction methods - a survey. In Workshop on Knowledge Discovery in Databases, 2002.

[20] A.E. Johnson, J. Aboab, J.D. Raffa, T.J. Pollard, R.O. Deliberato, L.A. Celi, D. J. Stone, A comparative analysis of sepsis identification methods in an electronic database, Crit. Care Med. 46 (4) (2018) 494.

[21] J.M. Juarez, J.M. Ochotorena, M. Campos, C. Combi, Spatiotemporal data visualisation for homecare monitoring of elderly people, Artif. Intell. Med. 65 (2) (2015) 97–111.

[22] P.S. Kam, A.W.C. Fu, in: Discovering temporal patterns for interval-based events, Springer, Berlin, Heidelberg, 2000, pp. 317–326.

[23] M. Karpefors, J. Weatherall, The tendril plot—a novel visual summary of the incidence, significance and temporal aspects of adverse events in clinical trials, J. Am. Med. Inform. Assoc. 25 (8) (2018) 1069–1073.

[24] D. Klimov, Y. Shahar, M. Taieb-Maimon, Intelligent interactive visual exploration of temporal associations among multiple time-oriented patient records, Methods Inf. Med. 48 (03) (2009) 254–262.

[25] D. Klimov, Y. Shahar, M. Taieb-Maimon, Intelligent selection and retrieval of multiple time-oriented records, J. Intell. Inform. Syst. 35 (2) (2010) 261–300.

[26] D. Klimov, Y. Shahar, M. Taieb-Maimon, Intelligent visualization and exploration of time-oriented data of multiple patients, Artif. Intell. Med. 49 (1) (2010) 11–31.

[27] J. Lin, E. Keogh, L. Wei, S. Lonardi, Experiencing sax: a novel symbolic representation of time series, Data Min. Knowl. Disc. 15 (2) (2007) 107–144.

[28] L. Lins, M. Heilbrun, J. Freire, C. Silva, Viscaretrails: Visualizing trails in theelectronic health record with timed word trees, a pancreas cancer use case. InWorkshop onVisual Analytics in Healthcare (VAHC), 2011.

[29] Martins, S. B., Shahar, Y., Galperin, M., Kaizer, H., Goren-Bar, D., McNaughton, D., et al., 2004. Evaluation of KNAVE-II: a tool for intelligent query and exploration of patient data. In Medinfo (pp. 648-652).

[30] F. Mörchen, A. Ultsch, Efficient mining of understandable patterns from multivariate interval time series, Data Min. Knowl. Disc. 15 (2) (2007) 181–215.

[31] R. Moskovitch, Y. Shahar, Classification-driven temporal discretization of multivariate time series, Data Min. Knowl. Disc. 29 (4) (2015) 871–913.

[32] R. Moskovitch, Y. Shahar, Classification-driven temporal discretization of multivariate time series, Data Min. Knowl. Disc. 29 (4) (2015) 871–913.

[33] R. Moskovitch, C. Walsh, F. Wang, G. Hripcsak, N. Tatonetti, Outcomes prediction via time intervals related patterns, in: In 2015 IEEE International Conference on Data Mining, 2015, pp. 919–924.

[34] R. Moskovitch, H. Choi, G. Hripsack, N. Tatonetti, Prognosis of clinical outcomes with temporal patterns and experiences with one class feature selection, ACM/IEEE Trans. Computat. Biol. Bioinform. (2016).

[35] R. Moskovitch, Multivariate Time Series Mining, Wiley's, Data Min. Knowl. Disc. (2021).

[36] P.V. Nickerson, R. Baharloo, A.A. Wanigatunga, T.M. Manini, P.J. Tighe, P. Rashidi, Transition icons for time-series visualization and exploratory analysis, IEEE J. Biomed. Health. Inf. 22 (2) (2017) 623–630.

[37] P. Novitski, C. Melzer Cohen, A. Karasik, V. Shalev, G. Hodik, R. Moskovitch, All Cause Mortality Prediction in T2D Patients, Artificial Intelligence in Medicine, Minneapolis, USA, 2020.

[38] P. Novitski, C. Melzer Cohen, A. Karasik, V. Shalev, G. Hodik, R. Moskovitch, Artificial Intelligence in Medicine, In Press, 2022.

[40] D. Patel, W. Hsu, M.L. Lee, Mining relationships among interval-based events for classification, in: In: Proceedings of the 2008 ACM SIGMOD international conference on Management of data, 2008, pp. 393–404.

[41] J. Pei, J. Han, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal, M.-C. Hsu, Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth, in: In: Proceedings 17th international conference on data engineering, 2001, pp. 215–224.

[42] A. Perer, F. Wang, Frequence: Interactive mining and visualization of temporal frequent event sequences, in: In Proceedings of the 19th international conference on Intelligent User Interfaces, 2014, pp. 153–162.

[43] Plaisant, C., 2018. Visualization of temporal patterns in patient record data. Fundament. Clin. Pharmacol. 32(1), 85–87.

[44] Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D., Shneiderman, B., 1998. Lifelines:Using visualization to enhance navigation and analysis. Inof Patient Records. Proceedingsof the American Medical Informatic Association Annual Fall Symposium. Citeseer.

[45] M. Pohl, S. Wiltner, A. Rind, W. Aigner, S. Miksch, T. Turic, F. Drexler, Patientdevelopment at a glance: an evaluation of a medical data visualization. InIFIP Conferenceon Human-Computer Interaction, Springer, 2011, pp. 292–299.

[46] L.K. Poon, S.-C. Kong, T.S. Yau, M. Wong, M.H. Ling, Learning analytics for monitoring students participation online: visualizing navigational patterns on learning management system, International Conference on Blended Learning (2017) 166–176.

[47] Shabtai, A., Klimov, D., Shahar, Y., Elovici, Y., 2006. An intelligent, interactive tool for exploration and visualization of time-oriented security data. In Proceedings of the 3rd international workshop on Visualization for computer security, pages 15–22. ACM.

[48] Shahar, Y., Cheng, C., 1999. Intelligent visualization and exploration of time-oriented clinical data. In: Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers, pages 12–pp. IEEE.

[49] A. Rind, T.D. Wang, W. Aigner, S. Miksch, K. Wongsuphasawat, C. Plaisant, B. Shneiderman, Interactive information visualization to explore and query electronic health records, Foundat. Trends Human-Comput. Interact. 5 (3) (2013) 207–298.

[50] K. Rodden, Applying a sunburst visualization to summarize user navigation sequences, IEEE Comput. Graphics Appl. 34 (5) (2014) 36–40.

[51] B. Shneiderman, The eyes have it: a task by data type taxonomy for information visualizations, Craft Inform. Visual. (2003) 364–371.

[52] K. Vrotsou, A. Nordman, Interactive visual sequence mining based on pattern growth, in: In 2014 IEEE Conference on Visual Analytics Science and Technology (VAST), 2014, pp. 285–286.

[53] E. Winarko, J.F. Roddick, Armada–an algorithm for discovering richer relative temporal association rules from interval-based data, Data Knowl. Eng. 63 (1) (2007) 76–90.

[54] Wongsuphasawat, K., Gotz, D., 2011. Outflow: Visualizing patient flow by symptoms andoutcome. In: IEEE VisWeek Workshop on Visual Analytics in Healthcare, Providence, RhodeIsland, USA, pages 25–28. American Medical Informatics Association.

[55] S.-Y. Wu, Y.-L. Chen, Mining nonambiguous temporal patterns for interval-based events, IEEE Trans. Knowl. Data Eng. 19 (6) (2007) 742–775.

[56] Maya Shwetz, Lior Fuchs, Victor Novack, Robert Moskovitch, Outcomes prediction in longitudinal data: study designs evaluation, use case in ICU acquired sepsis, J. Biomed. Informatics (2021).