DOI: 10.1111/cgf.14498

COMPUTER GRAPHICS *forum*
Volume 41 (2022), number 6 pp. 273–286

Visual Analytics of Multivariate Intensive Care Time Series Data

N. Brich,[1,2] C. Schulz,[3] J. Peter,[2,4] W. Klingert,[5] M. Schenk,[5] D. Weiskopf[3] and M. Krone[1,2]

[1]Big Data Visual Analytics in Life Sciences, University of Tübingen, Tübingen, Germany
[2]Institute for Bioinformatics and Medical Informatics (IBMI), University of Tübingen, Tübingen, Germany
[3]Visualization Research Center, University of Stuttgart, Stuttgart, Germany
[4]Medical Data Integration Center, University Hospital Tübingen, Tübingen, Germany
[5]Department of General, Visceral and Transplant Surgery, University Hospital Tübingen, Tübingen, Germany

Abstract

We present an approach for visual analysis of high-dimensional measurement data with varying sampling rates as routinely recorded in intensive care units. In intensive care, most assessments not only depend on one single measurement but a plethora of mixed measurements over time. Even for trained experts, efficient and accurate analysis of such multivariate data remains a challenging task. We present a linked-view post hoc visual analytics application that reduces data complexity by combining projection-based time curves for overview with small multiples for details on demand. Our approach supports not only the analysis of individual patients but also of ensembles by adapting existing techniques using non-parametric statistics. We evaluated the effectiveness and acceptance of our approach through expert feedback with domain scientists from the surgical department using real-world data: a post-surgery study performed on a porcine surrogate model to identify parameters suitable for diagnosing and prognosticating the volume state, and clinical data from a public database. The results show that our approach allows for detailed analysis of changes in patient state while also summarizing the temporal development of the overall condition.

Keywords: information visualization, visualization, methods and applications, visual analytics

CCS Concepts: • Applied computing → Health care information systems; • Mathematics of computing → Time series analysis; Dimensionality reduction; • Human-centered computing → Information visualization; Visual analytics

1. Introduction

In a modern clinical setting, patients in an intensive care unit (ICU) are closely monitored by multiple medical devices. This multitude of measurements has to be communicated for interpretation, either to medical personnel for progressive analysis (monitoring), or—as in our context—to medical researchers for post-hoc analysis (experiments). Humans have difficulties aggregating more than three quantities, especially under the presence of uncertainty [GASM17]. However, multiple parameters are often needed to provide an overall impression of the condition of a patient. Only considering a combination of changes in multiple parameters will hint toward important transitions of the whole homoeostatic system. For instance, individual parameters might seem uncritical if considered independently, but in conjunction, can indicate an imminent circulatory shock.

Ideally, an illustration of patient state would be based on the entirety of time-varying interacting variables, without having to rely on a set of specific, pre-defined parameters. Moreover, such an approach would reveal high-dimensional processes clearly without being overly specific and still provide fine-grained details on demand. In practice, one of the primary problems with multivariate time-varying data is the conceptual and physical limitation of straightforward visualisation techniques: most established approaches, such as scatter plot matrices, do not scale well enough to support the assessment of more than a few variables due to visual clutter and required mental effort to draw the right conclusions. In response to this observation, many expert users resort to univariate visualisation techniques and (over-)simplifying derived characteristic quantities that take little advantage of the human visual system.

In this work, which is an extension of an earlier workshop paper [BSP*20], we propose a visualisation approach that facilitates the exploration and analysis of multi-dimensional, time-dependent ICU data based on time curves [BSH*16, vdEHBvW16], small multiples [Tuf92] and comparative views for an ensemble of subjects. We showcase our approach using real-world data obtained from an experimental surgery study performed on a porcine surrogate model

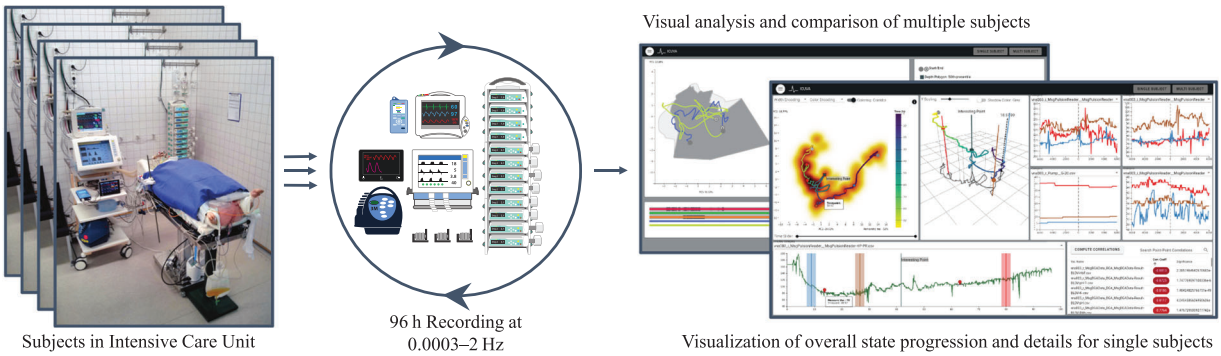

Figure 1: *Our visual analytics application for multivariate medical data applied to data obtained during an experiment with porcine subjects attached to multiple medical devices in an intensive care unit. The overall state progression over the set of measurements per animal is shown as time-curve-inspired plots. A line plot showing a selected measurement and small multiples composed of selection-dependent line plots allow for state comparison and in-depth analysis. For an ensemble of multiple subjects, a time curve boxplot allows to compare subjects.*

(Figure 1). To compute our time curves, we have analysed various projection techniques regarding their suitability for our application. Additionally, an unsupervised machine learning model—a Convolutional Variational Autoencoder (CVAE)—was implemented to better account for non-linearity and uncertainty in time curves. To disambiguate self-shadowing and resolve occlusion of the time curve, we implemented a linked 3D visualisation in which the time curve is visualised in a space–time cube. Different colouring modes for the time curve facilitate tracking single variables and connecting trends in the overall patient status with these variables. To compare patient states and drill down into details, our application allows the user to select multiple data points in time that, together with a user-defined temporal context, are then shown using small multiples. To quantify the similarity of changes of a specific variable at different time points, the point-to-point Pearson correlation can be shown. The contributions within our web-based visualisation application can be summarized as follows: the visual analysis approach to high-dimensional and time-varying ICU data, the adaption of various visualisation techniques using z-score and data-depth-based statistics to ease outlier analysis, a domain-driven generative data model to evaluate the applicability of various dimensionality reduction (DR) algorithms to ICU measurements and a discussion containing expert feedback as well as an outlook for clinical surveillance.

2. Related Work

Although the visualisation of high-dimensional and time-varying data has been investigated intensively and has been reviewed in many extensive surveys [AMST11, KH13, LMW*17], it is still an active area of research. Classical multivariate visualisations such as parallel coordinate plots and scatter plot matrices are either a poor fit for visualising multivariate time series, or only suitable for specific tasks like finding correlations between single variables. Thus, tailored task-driven visualisation methods are needed [Mun15] to facilitate an effective visual analysis of complex multivariate time-varying data. Since we are dealing with a number of time-varying input dimensions exceeding the capabilities of classical multivariate visualisation, we briefly review suitable approaches.

Dimensionality reduction. DR methods compute a low-dimensional representations from high-dimensional data, which are often easier to visualise and interpret than high-dimensional data. A typical example of this is principal component analysis (PCA) [Hot33, Pea01], which computes a linear projection based on variance. Another widespread technique is multi-dimensional scaling (MDS) [BG05], which optimizes for representing high-dimensional distances in a low-dimensional space. Generally, there is neither a perfect solution nor a uniquely defined global optimum for globally preserving distances. Thus, many techniques that use neighbourhood information instead of distance have been developed. A popular example is the computationally expensive t-Distributed Stochastic Neighbour Embedding (t-SNE) [vdMH08], which optimizes using entropy, generally leading to a misrepresentation of global distances in the embedding. However, t-SNE and other non-linear techniques excel at separating clusters. A more recent approach is Uniform Manifold Approximation and Projection (UMAP) [MHM18], which aims at reducing this misrepresentation of distant data points by using a Riemann manifold.

Vernier *et al*. [VGdS*20] investigated a wide variety of DRs regarding their feasibility towards multivariate temporal data. Our application guides the data-driven choice of DR technique, based on approaches from research on quantification and visualisation of projection error [SSK10, HAF13, WSA*16]. To gain a better understanding of the overall suitability of the DR techniques for our application, we use a domain-specific generative data model for evaluation purposes [SNEA*16] (Section 5.2). Vernier *et al*. found evidence that variational autoencoders [KW13] are a suitable choice for DR of multivariate temporal data. Since it depends on the data and the research question, which of these approaches is most suitable, our application offers the choice between linear (focusing on temporal change) and non-linear techniques (focusing on clusters).

Visualisation of multivariate time series. Our medical visual analysis application was inspired by several techniques to visualise multivariate time series that leverage DR: Jaeckle *et al*. [JFSK16] use temporal MDS to project non-temporal dimensions into a single dimension while retaining one display dimension for the mapping of the time component. This results in an optimal depiction of the

progressing time, but impedes the distinction of different states, since they are plotted on only one axis. As we want to provide an intuitive overview of differences in patient states, the utilization of the two major axes for non-temporal components seems like a better choice. Bach *et al*. [BSH*16] concurred with this assessment and developed the time curve concept. They used metric MDS to project the high-dimensional data in two dimensions while illustrating the temporal succession by connecting consecutive points using a Bézier curve. A similar approach was previously presented by Bernard *et al*. [BWS*12], who used PCA. Concurrently to Bach *et al*., van den Elzen *et al*. [vdEHBvW16] developed a similar technique to visualise temporal evolution of graphs. These techniques can be classified as connected scatter plots, which follow the same idea of connecting temporally consecutive data points, but do not use DR and date back to 1790 [HKF16]. An example is the work of Grottel *et al*. [gro14], who map time to line density in traditional scatter plots and parallel coordinates. Recently, time curves have been used by Hinterreiter *et al*. [HSS*21] to analyse solutions in problem-solving tasks. We adopt the concept of time curves and integrate it into our exploratory data analysis application. Due to the oscillating nature of the ICU data, we extend the original concept of time curves with filtering. We use this filtering step to declutter the data—and, consequently, the resulting visualisation—and to remove outliers. Our application can handle not only single data sets but also multiple ones to compare the ensemble members.

For the analysis of multiple data sets, we extend the time curve visualisation with the notion of data depth and curve boxplots [MWK14]. Data depth acts as a measurement to describe how central a single data point is within the whole data set: the higher a data depth, the more central a given data point is [CANR08]. While data depth was already successfully applied to time series [LPR09, LPSLG14], we explore the application of this concept to time curves.

Medical visualisation. Ordonez *et al*. [OndL*10] explored the usage of star charts for visualising ICU data. While this works reasonably well for a small number of variables, visual scalability is limited. In the course of the increasing digitalization, the visualisation of electronic health records (EHR) has gained much attention. Zhang *et al*. [ZMG*10] proposed a visual analytics framework that unifies EHR information for use in emergency rooms. While it also shows time-varying multidimensional data, their approach is not scalable and assumes discrete patient states and transitions, not numerical measurements. Similarly, Rind *et al*. [RAM*11] presented a tool for the visual analysis of long-term medical records of patients with chronic diseases. Their application case is different from ours since the data are more sparsely sampled over much longer periods of time. Ten Caat *et al*. [CMR05] introduced the concept of tiled parallel coordinates for multi-channel EEG data. While their data are more similar to ours, they just have multiple sensors of the same type that all measure at the same frequency. They also use aggregation over the whole measurement time, which is not suitable in our case.

3. Medical Background

In this section, we summarize the experiment we used to evaluate our approach and the required clinical background knowledge [PKK*14, Pet18, KPT*18]. Preliminary research showed insufficient knowledge regarding the assessment of the volume state in clinical practice for hospitalized patients. The volumeneed-analysis (VNA) study was conducted to answer research questions regarding the reliable assessment of volume states in a clinical context. One such question is whether it is possible to predict states and potential problems by observing measurable quantities of the organism so that doctors can proactively intervene, instead of being forced to react. To this end, a set of ICU devices commonly found in a clinical context was used to obtain the measurements. However, the setup was far more comprehensive than a regular setup in a hospital, providing almost complete knowledge about the volume state of subjects.

The VNA study was conducted using 10 fully anaesthetised pig surrogates in accordance with ethical guidelines. In total, the experimental setup measured roughly 250 parameters over roughly 4 days each, with two experiments being ended prematurely. Recorded parameters include a variety of cardiovascular parameters such as blood pressure and heart rate, blood gas analysis results such as ion and carbohydrate concentrations, as well as general parameters like weight and diuresis volume. Initially, the animals were given defined amounts of fluid using intravenous infusions to reach an increased fluid state. Then, fluid supply was stopped, and the body fluid was reduced through diuretics. The fluid state was assessed at defined states by performing a variety of tests. All measurements were collected from a set of nine devices via manufacturer interfaces or wiretapping. While some devices sample adaptively, the average sampling rate varies between $^1/_{3600}$ Hz (blood gas analysis) and 1–2 Hz (bedside monitors). All devices are synchronized to an accuracy of approximately 1 s.

4. Requirements

The measurement data already defined the quantity structure (see Section 3); therefore, we can focus the requirement analysis on the functionality. The high-level goal of our collaboration is to provide *an orthogonal approach to visual analysis of individual measurements*. In line with our experts, we argue that a loss of information to gain insights into high-dimensional processes is justifiable: as our primary use case is the post-hoc analysis of ICU data, our system cannot cause harm to the subjects. Thus, the following goals and requirements were defined iteratively during multiple design sessions and on-site meetings together with our medical domain experts:

- R1 The main objective is to be able to tell whether the patient is in a stable state or not. Since a patient does not necessarily have to be within the ideal parameter ranges, communication of change rates and state comparison is more important than absolute values.
- R2 Both the aggregate of change and individual changes can be of interest. Therefore, the visualisation application should allow for both a top-down and a bottom-up exploration of the data.
- R3 Robust statistical methods should be used to guide the search for interesting patterns and relevant behaviour—some parameters are heavy-tailed distributed and oscillating.
- R4 Medical researchers have high expectations regarding sanity-checking and explanation of phenomena due to the nature of the medical domain—easy interpretation without a strong mathematical background is of major interest.

R5 It should be possible to put individual patients in the context of an ensemble to assess if a patient is representative or exceptional.

5. Visualisation of ICU Data

Our goal was to combine several visualisation approaches in one application tailored to medical experts. In this section, we describe the techniques we use as components and their adaption to our specific application. In accordance with the visualisation pipeline [Mor13], this includes data cleansing, an evaluation of DR algorithms, and a description of the visual mappings, which consist of filtered time curves and small multiples composed of line plots for in-depth analysis and comparison of patient states, as well as time curve boxplots for ensemble analysis (Figure 1).

5.1. Data cleansing and preprocessing

We cleanse the data beforehand to remove measurement errors. We remove erroneous samples using a list of user-defined error codes, filter inaccurate samples, and remove categorical time series and those that have a constant value. To identify and remove outliers, we have explored various techniques, such as the Savizky-Golay filter [SG64] that approximates a variable using a polynomial. While this filter removes severe outliers, it also introduces artificial values and thus conflicts with R4. We opted for a z-score-based filter that classifies values as outliers that are not contained in the hull spanned by the moving average and standard deviation.

All considered DR techniques require a uniform sampling, i.e. all components of a high-dimensional point have to be defined. Due to the high variance in sampling rates, a simple downsampling approach would lead to massive loss of data. In line with recommendations from our medical experts, we do not use interpolation to fill in missing values since this would introduce a possibly invalid assumption about the measured quantities. Therefore, we unify sampling rates across all time series. Given a user-specified sampling rate, we down-sample time series with a higher sampling rate than the chosen one and up-sample time series with a lower sampling rate, respectively. Missing values are filled via forward-filling (i.e. the last valid value is used until the next valid value is reached). Our medical experts approved of this strategy. If a measurement starts at a later point in time, we apply backward-filling, as there are no earlier samples available.

5.2. Dimensionality reduction

Time curves project a time-varying high-dimensional space into a two-dimensional space, where time is not part of the DR, i.e. $(t, \mathbb{R}^n) \mapsto (t, \mathbb{R}^2)$. Time is added again after the DR by connecting the points in chronological order using lines. Therefore, we had to choose a suitable DR technique. To evaluate the DR, we use a generative data model [SNEA*16], since the data from the VNA experiment contains unknown trends and patterns. Our data generator emulates the properties of the expected real-world multivariate time series using a combination of sinusoidal functions and piecewise linear ones. To justify the validity of our generator, we refer to the parameter range of the devices used during the experiment, as well as expert feedback. Thus, the resulting time series are representative of experimental measurements, as they contain variables with repeating patterns of varying frequency and amplitude (e.g. pulse, respiration) in conjunction with aperiodic, linearly changing quantities (e.g. temperature, weight). This approach allows us to establish a link between the parameters of the generative model and the output of the DR algorithm. This link is established through an evaluation metric: for PCA, it is the length of the first two components; for MDS, it is stress. For non-linear techniques, interpretation of the resulting dimensions is inappropriate. This form of sensitivity analysis would not be possible using only a few data sets with unknown properties. Moreover, the generative data model helps us to do sanity checking and allows users to familiarize themselves with the application using easy-to-understand data (R4).

For evaluation, we generated six synthetic time series. To model physiological processes, which often exhibit oscillating behaviour, the synthetic data sets are sinusoidal curve ensembles with varying frequency, amplitude and phase. Generally, we found that MDS using Euclidean distances represents long distances better than short ones, causing the beginning and end to be slightly deformed. Note that we used metric MDS with the SMACOF algorithm [dL77], which—in contrast to classical MDS [Tor52]—differs from PCA. Moreover, MDS tends to introduce loops that are not present in the data, which might be caused by non-optimal flipping and rotation while iterating (unfavourable local optima). PCA, on the other hand, shows the oscillating behaviour of the data while separating dissimilar clusters of time series. We investigated the influence of varying frequency, amplitude and phase (model parameters). We found that amplitude and phase had little impact on projection quality, while varying frequencies quickly degraded projection quality (see supplement for details). While both metric MDS and PCA are linear techniques, we suspect that suboptimal preconditioning leads to favoring PCA. Considering both aspects, we rate PCA as the suitable default and more defensive choice for DR in our application (R3). This is in line with van den Elzen *et al*. [vdEHBvW16], who used PCA to facilitate interpretation of the resulting dimensions, and t-SNE for the analysis of clusters. When a more cluster-oriented approach is desired, we offer UMAP as a non-linear DR method. It yields comparable results to t-SNE while exhibiting a more desirable performance profile.

CVAEs are an unsupervised machine learning method utilizing an encoder–decoder architecture, each consisting of several layers, and a variational constraint connecting the encoder and decoder. Our data preparation and layer design is inspired by Ali *et al*. [AJXW19] and the Keras documentation [C*15]. Similarly to Ali *et al*., we partition the multivariate data set into *snapshots*, i.e. fixed time intervals with a stride of one. The encoder component reduces the snapshot with each succeeding layer of nodes until a bottleneck is reached, which forms the latent space of the representation. Symmetrically to the encoder, the decoder takes input from the latent space and, with succeeding layers, reconstitutes the original amount of dimensions. The reconstituted snapshot is then compared with the original snapshot using the mean square error (MSE). The quality of the training is expressed by the *reconstruction loss* (i.e. MSE) and the *entropy loss* (here, the Kullback–Leibler divergence). Minimizing those losses using a gradient descent allows the autoencoder to learn the multi-dimensional structure of the data set, encoding it in the latent space. Due to the layer architecture, an autoencoder can

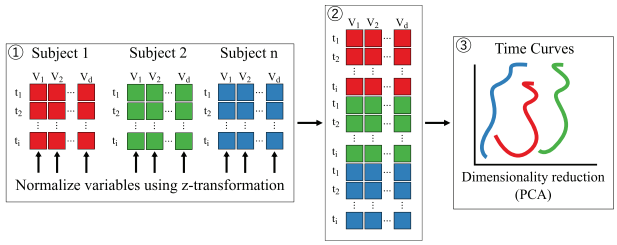

Figure 2: *Schematic of the normalization and principal component analysis (PCA) projection for multiple subjects. ① Subjects are normalized individually using a z-transformation. ② The data frames are concatenated, and ③ PCA is subsequently performed on this concatenated data frame ensuring identical loadings allowing comparison of individual subjects.*

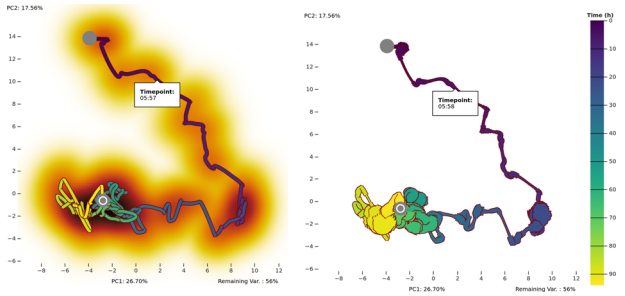

Figure 3: *Time curves for subject VNA004 with constant width and colour-coded density (left), and with data point density in proximity encoded as width (right). The curves are colour-coded according to the time since the start of the measurement. Both curves are filtered using the average of a moving window. The background is coloured by point density (, high to low).*

represent non-linear structures, which contrasts it to PCA. Training means fitting the entire data set in this context. Accordingly, the resulting latent space represents the entire data set as a set of variables that contain a notion of compressed and simplified (dis-)similarity of the original data points. Since our bottleneck is variational, our latent space models Gaussian distributions, i.e. mean and standard deviation. This allows compensating for a certain amount of fuzziness, making the training process more robust. To obtain a visualisation from this latent space, we employ PCA. Note that adding a CVAE does not improve results per se: convolution is a great tool for analysing oscillating signals but has to be taken with a grain of salt when applied to variables without meaningful spatial relationship [VGdS*20]—even though there can be correlation. Thus, we apply the convolution only on the temporal axis. In summary, our application offers three different DR methods to allow for three different analysis paths: PCA maintaining global distances and relations, CVAE in combination with PCA to leverage its non-linearity to accommodate for highly complex regions in the time curve, and UMAP for cluster analysis. The ensemble analysis (R5) requires a collective DR of multiple subjects. Here, we apply z-transform individually to each subject prior to DR. The DR is then performed for all subjects at once, allowing for comparisons of subjects (Figure 2). Consequently, and for a meaningful comparison, the same set of dimensions must be used for each subject.

5.3. Time curve visualisation

Filtering. The DR algorithms embed each multi-dimensional data point of the data set into a two-dimensional space. To obtain a time curve, these 2D data points are connected in chronological order.

Since all DR algorithms emit rather cluttered time curves for real-world as well as generated data, it became clear that we had to reduce the complexity of the curves to cope with visual clutter. Therefore, we smooth the emitted data using a Bartlett-window moving average. The width of the moving window determines the degree of smoothing, which can be adjusted by the user. Obviously, filtering introduces a loss of information (R4). Hence, we devised different ways to compensate for this loss of information.

One option is to encode the density of data for a given region before smoothing with colour (Figure 3, left). We use a Gaussian kernel in conjunction with kernel density estimation to generate a smoother density visualisation. This map allows assessing how many time points lie in the proximity of a time curve segment and, thus, enables the viewer to draw conclusions about the velocity of changes in the system. The other option is a mapping of variables onto the width of time curve itself (Figure 3, right).

Space-time cube. Time curves often self-intersect, which can make it difficult to follow the course of the curve. To remedy this problem, we also render the time curve in a space-time cube. Space-time cubes are a common representation in geospatial visualisation to show the temporal progression of a path on a map by using the third dimension to display time. However, while the space-time cube helps to alleviate this specific issue, it also has inherent problems [BDA*17]. Thus, we only use the space-time cube as an auxiliary view alongside the planar time curve to support getting a more complete understanding of the state progression through the 3D visualisation (see Figure 8). To improve spatial perception, we render the curve in the space-time cube as a shaded tube and add a grid as well as a shadow of the curve on the bottom plane. Temporally evenly spaced perpendiculars originating at the curve further add to the spatial perception.

Colour mapping. We use the colour channel to map additional data values in both planar and 3D time curve representation. Two ways of mapping the elapsed time to the curve were implemented: one showing the time using a sequential colour scale, and one using a discrete quantized scale. Two further colour modes map derived values to the time curves that enhance the analysis (see Figure 4). In the first mode, the user can supply a value range per variable (R2), which represents the expected—or desired—value range for a healthy subject. The time curve can be coloured per variable, with upward and downward deviations displayed in discrete percentage increments. Alternatively, a combined mode for all variables is available that colours each time point according to the variable that exhibits the strongest deviation. The second additional colour mode is a change-point-based colouring. At change points, the trend of a variable becomes different from earlier time segments. This allows for a direct comparison of the overall patient state and whether an observed change in overall patient state coincides with

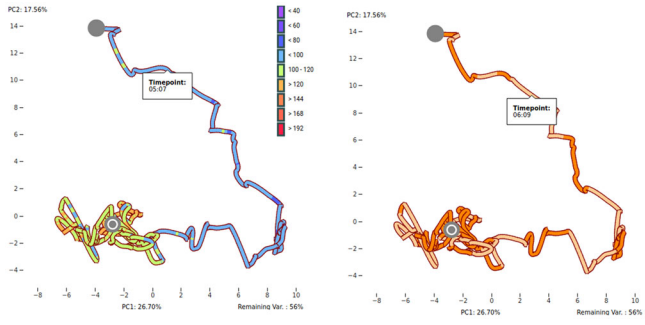

Figure 4: *Left: Expected value colouring shown on systolic blood pressure. User-supplied expected value ranges for the variables are used for colouring. Right: Change point colouring shown on heart rate. Colour change (■/■) illustrates that the preceding and succeeding intervals are substantially different for the chosen variable.*

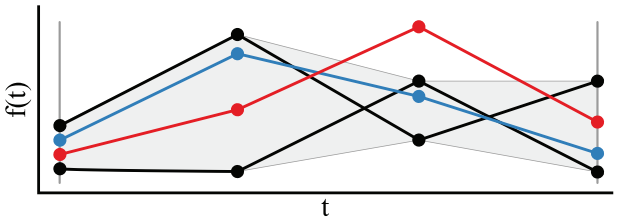

Figure 5: *Schematic of functional band depth. The band (grey), indicating the most representative functions (bounding functions, black). The blue function is included, i.e. belonging to the group of most representative function, whereas the red one is not included, and thus is less representative for the ensemble of functions.*

a change point in a given variable (R1, R2). The change points of each variable are calculated using the *Ruptures* package [TOV20].

5.4. Time curve boxplots and scarf boxplots

When plotting multiple time curves in a single visualisation, the result is often dominated by visual clutter even for relatively low numbers of curves. Especially the overlapping curves are challenging to assess, rendering it difficult to see which one of the curves is representative. Therefore, we extract a representative hull that encloses a centrally outward ordered share of curves. To establish this order, we resort to a class of order-statistical metrics called *data depth* [CANR08] that describe how central an element is within an ensemble. In this context, centrality can be understood as representativeness. We use a metric called *functional band depth* [LPR09], which builds on functional bands. A functional band B_k consist of $k \geq 2$ functions $f_{ir}(t \in I)$ with restriction interval I:

$$B_k\left(\begin{bmatrix} f_{i1} \\ \vdots \\ f_{ik} \end{bmatrix}\right) = \{(t, y) : t \in I,\ \min_{r\in[1,k]} f_{ir}(t) \leq y \leq \max_{r\in[1,k]} f_{ir}(t)\} \quad (1)$$

A function lies between other functional bands if it is enclosed by the respective minimum and maximum values, as shown in Figure 5.

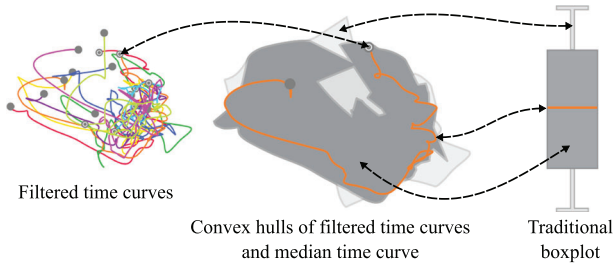

Figure 6: *Visual metaphor of the boxplot as applied to time curves. The cluttered view of eight time curves is simplified by representing curves as areas corresponding to different quantiles: The 50% quantile (dark grey) corresponds to the box of the boxplot, while the 80% quantile (light grey) corresponds to the whiskers. The purple time curve is the most representative one (by functional data depth), and thus corresponds to the median line in the boxplot.*

Formally, we can count this condition using an indicator function χ with $true \mapsto 1$ and $false \mapsto 0$ to define band depth $D_{n,k}$.

$$D_{n,k}(f_i) = \sum_{j=2}^{k} \binom{n}{j}^{-1} \sum_{l=1}^{j} \chi\left(\{(t, f_i(t)) \mid t \in I\} \subset B_l\left(\begin{bmatrix} f_{i1} \\ \vdots \\ f_{il} \end{bmatrix}\right)\right) \quad (2)$$

This formula first counts how often a function f_i is contained in bands of length l. The number of actual inclusions is then put in relation to the possible ones. The process is repeated from bands of length 2 to up to bands of length k. A smaller k is more sensitive to variation, whereas a larger k is the more resilient to small fluctuation. Note that Equation (1) can also be expressed as a convex combination, which emphasizes the enclosing nature of the isosurfaces that can be defined on this metric, i.e. our enclosing hull. While data depth has been used to reduce visual clutter and design boxplots for quite some time [SG11, MWK14], it has not been applied to dimension-reduced curves yet. In this application, the calculation of the functional band depth is performed on the high-dimensional, z-transformed data. The enclosing hulls are then generated, according to the order established by data depth, from smoothed time curves after DR. Figure 6 shows our time curve boxplot that covers the 50% and 80% quantile of most representative time curves, allowing the omission of all time curves included in this quantile. Consequently, recognizing which curves are outliers and which ones are representative for the whole ensemble is easily possible. Figure 6 shows this concept applied to an ensemble of eight curves. By showing the two quantiles (50% and 80%) and the most representative curve (which corresponds to the median), we follow the well-known design of a standard boxplot.

As the containment of a curve in the box area not only depends on being at the correct position, but on being at the correct position at a specific point in time, it becomes difficult to judge when a given curve is not included in the curve box area. Hence, we provide an accompanying scarf boxplot (Figure 7) that depicts box inclusion over time, with dark grey and light grey markers indicating the inclusion in the respective quantile. Since the scarfs are temporally

Figure 7: *Scarf boxplot illustrating representativeness over time from left to right for each subject that is not included in the inner quantile box plot area. Colours and labels correspond to the ones in the main time curve boxplot, while the dark grey and light grey markers indicate inclusion in the respective boxplot area.*

aligned, it is easy to judge when a curve is representative and or not in comparison to other curves.

5.5. Details using small multiples

The purpose of the time curve visualisation is to give an overview of the progression of patient state (R1). While it allows for identifying potentially interesting points in time, it is not intended for a detailed analysis of individual parameters (R2). Given the unavoidable projection error, loss of information, and the sanity checking requirements (R4), we show details on demand in small multiples of unprotected data dimensions. Thus, an auxiliary line plot that depicts values over time and several small line plots allows for easy superimposition of small time-spans. For each line plot, the user can select a different data variable in a drop-down menu. Users can select between 1 and N points on the time curve or auxiliary plot. Each selected point corresponds to an identically coloured polyline in each of the small line plots. Respective time series are centered to the selected time point, showing backward and forward temporal development of the data (Figure 8). The user can select the time delta that is used for the abscissa. The temporal interval is also drawn as a semi-transparent window in the auxiliary line plot (Figure 8). This allows the user to get an impression of temporal change in individual dimensions by comparing the values at different points in time. It also allows for comparing the progression of different data dimensions and, consequently, to infer interdependency and correlation. The line plots relate to requirement R2, since they facilitate the individual inspection of each line plot while the small multiples view also allows for investigating localized trends and patterns.

6. Description of the Visualisation System

We implemented our visualisations described in Section 5 as a client-server web-based application. Costly calculations are performed server side while the presentation of the results is facilitated client side. This enables access from different terminal devices—a scenario commonly found in a hospital setting.

The server was implemented using Python. The Flask framework [MRLU] was used to communicate with the HTML5 client page. For data filtering and sampling, we use Pandas data frames [Mo10]. Keras [C*15] and Tensorflow [ABC*16] were utilized for the CVAE, while scikit-learn [PVG*11] was used for PCA. The client-side visualisation and UI was built with the Vue.js framework, and the D3 library [BOH11]. For the WebGL-based rendering of the 3D space-time cube, Three.js was used.

The system features two distinct views: the first view is meant for analysing a single subject and offers features for an in-depth post-hoc analysis. The second view offers comparative visualisations for an ensemble of multiple subjects.

Figure 8 shows the single-subject mode. The collapsible menu to the left is used to select a data set and configure most of the visual-

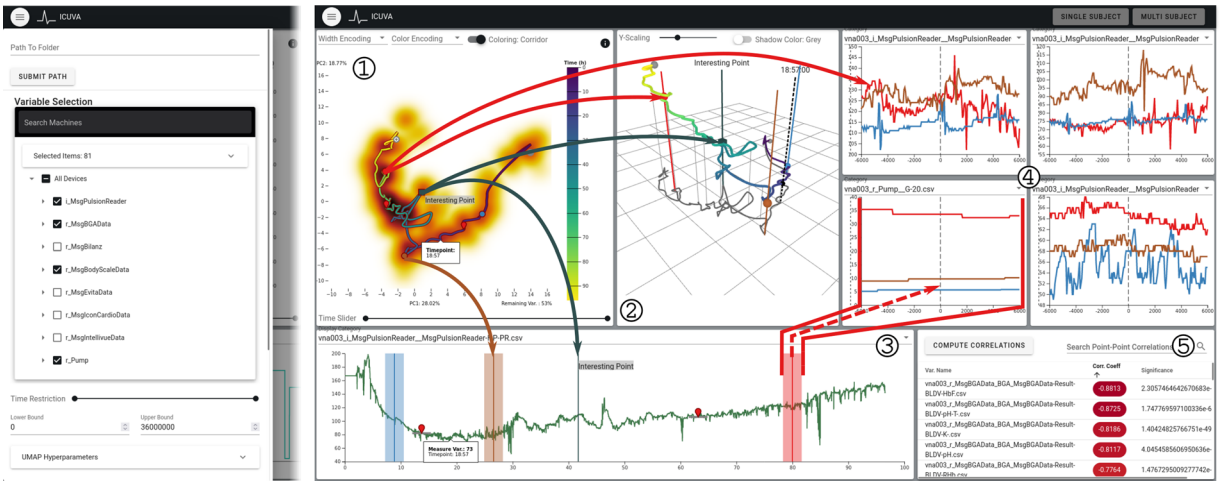

Figure 8: *Overview of the single-subject mode (right) and expandable menu (left). The time curve shows patient state progression (top left; 2D time curve plot ① and space-time cube ②). An auxiliary line plot shows a single time series, with time in hours ③. Three samples are selected (red, brown and blue) to showcase linking. Dark grey lines and markers with custom labels are used to annotate interesting point in time (e.g. peaks in specific measurement variables). The small line plots ④ show superimposed time intervals (in seconds) based on selected points, for one selected variable each; the width of the bar in the auxiliary line plot corresponds to the, identically coloured, time segment in the small line plots. The shown variables can be selected freely from the variables used for projection. The correlation table ⑤ compares two points in time that are indicated as red markers in the auxiliary view and the 2D time curve.*

isation pipeline including data selection, processing and choice of DR technique. This mode features a patient state progression visualised as a filtered time curve and space-time cube visualisation, a large auxiliary line plot, and four small selection-relative line plots. A drop-down menu above the time curve visualisation allows for the selection of a variable that is encoded onto the width of the time curve. Alternatively, constant width and spatial point density can be chosen. A second drop-down allows the user to choose between the colour modes explained in Section 5.3. The filled circle ● indicates the starting point of the curve, the dotted circle ⊙ the endpoint. To scrutinize the progression of individual variables, a large auxiliary line plot is shown below the filtered time curve. The visualised variable can be chosen in the drop-down menu above the plot. Right of the time curve visualisations, small selection-centric line plots are shown. All plots are linked views that support hovering and selection as well as annotating time points with labels. Selecting points on either the time curve or in the auxiliary line plot will determine the point in time that is shown in the small selection-relative line plots. These plots show the temporal surroundings of the selected time points over the chosen variables. In addition to the four small multiple line plots, a second way to compare segments of the curve was implemented: a pair of time points on the curve can be selected (red markers 📍 in Figure 8). The Pearson correlation coefficients are calculated for each variable of the two selected points over an equally sized window (R3). The coefficients are listed in a table, which can be sorted by correlation or significance.

The multi-subject mode has an analogous menu for data loading and selection, with additional inputs for the quantiles of the time curve boxplot computation. The main view is a time curve plot showing an ensemble of (filtered) time curves as well as the time curve boxplot. It supports toggling the visibility of individual ensemble members and the enclosing curve boxplots (Section 5.4). To further inspect a specific subject, the user can switch to the single-subject view of this subject by selecting it in the legend. Since the superimposed time curves depict the similarity of the samples instead of a temporal alignment, we added a scarf plot below the time curves to illustrate box inclusion over time. It only shows subjects that are not included in the inner quantile range. Thus, representativeness can be analysed both in terms of values and time.

7. Results and Discussion

In this section, we detail several examples and investigations that showcase that our presented visual analysis approach is a promising tool for medical research concerned with ICU data. We also highlight how these use cases relate to the user requirements. For the single-subject mode, a set of 82 parameters from the data described in Section 3 was used. For the multi-subject mode, we present the results for nine animal subjects (one animal was omitted due to short experiment duration) using a set of nine relevant parameters each. We also evaluated the applicability of our approach to other ICU data sets using publicly available data from PhysioNet [GAG*00]: MIMIC-III-Waveform [JPS*16, MMV*20] and MGH/MF-Waveform [WFTR91]. We also investigated the potential of our visual analysis approach by gathering feedback from our domain experts during development and by conducting structured feedback interviews with independent medical professionals.

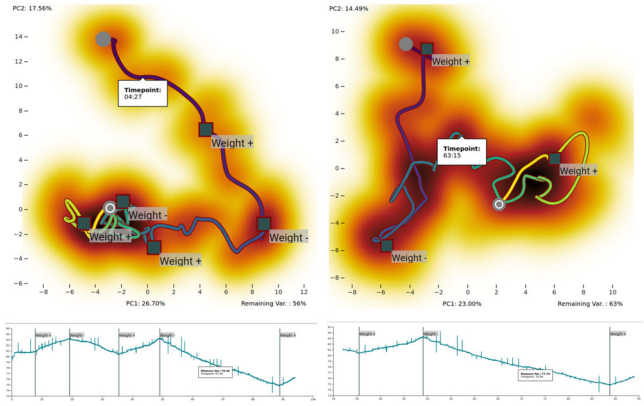

Figure 9: *Juxtaposition of the time curve and the net weight of the subject over the whole 94 h (left) and shortened to the last 62 h (right). Time curves smoothed were with a window size of 150 (Figure 3 shows the same curve with a windows size of 50).*

Single-subject mode. Figure 9 shows the filtered time curve of subject VNA004. The most prominent feature is the significant and rapid change in status during the first 20 h, indicated by the comparatively small temporal change as shown by the small change in colour and the large difference in state (patient stability: R1). While it is difficult to pinpoint, our medical experts have identified this segment of the curve as the onset of anaesthetics and the implantation of measuring devices that put a strain on the homoeostatic system of the animal. Using the auxiliary plot and the small multiples (R2), we could identify several variables (e.g., heart rate, glucose and blood pressure) that might explain this rapid change in the time curve. The dense region at the end of the time curve signifies that the patient state does not change as drastically anymore but only fluctuates slightly around a stable state.

During the VNA study, all subjects were infused with high amounts of liquids over an interval to be then depraved of liquids. This change can be seen in the continuous weight monitoring. The points at which the administration of liquids is either started or stopped leads to noticeable bends in the time curve. As shown in Figure 9, three of the five administration changes are in a hairball-like region, roughly between 32 and 94 h. In these 62 h, the observed subject reaches a stable state, indicating that the variables do not change considerably. This stabilization can also be seen in several variables (e.g. heart rate, glucose and blood pressure).

To disambiguate the hairball, the time curve was restricted to this temporal window of 62 h (see Figure 9, right). The aforementioned bends in the time curve can be narrowed down to the fact that the amount of liquids excreted from or added to the organism are substantial (2.7 and 8.3 kg), straining the organism. Thus, it is expected that the state as a whole—represented by the time curve—will show a noticeable reaction (R2).

To test our approach more systematically, we enriched the VNA data set with patterns corresponding to—exaggerated—scenarios encountered in real life. One such scenario is the sudden and repeated change of a few variables, e.g. a rapid increase in heart rate. A sudden change in some variables should lead to deviations of

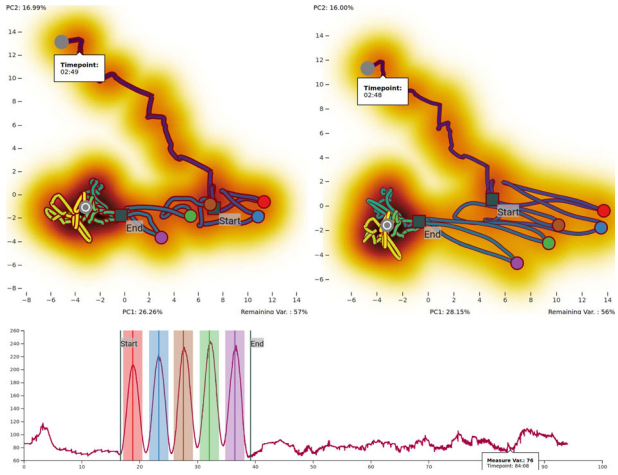

Figure 10: *Artificial oscillations induced in eight (top left) and 16 (top right) of the 82 variables of* VNA004 *lead to loop-like patterns in the resulting time curve, while the rest of the curve changes only marginally. This test shows that time curves are a robust and intuitive approach for the analysis of time-varying multivariate data.*

considerable magnitude at the corresponding positions. This is simulated by enriching an interval in a subset of values with sinusoidal functions (Figure 10 bottom). These should produce periodically similar patient states resulting in periodically similar positions in the time curve (i.e. loop-like structures). The peaks induced by the oscillations should indicate a substantially altered overall state, leading to noticeable deviations from the original time curve trajectory. This effect can be seen in the marked segments in Figure 10, which also shows that the effect is amplified when increasing the number of variables that are simultaneously enhanced with sinusoidal patterns.

Another scenario concerns repetitions in patient state, e.g. large-scale cyclic behaviour. To simulate such behaviour, a data set was replicated and concatenated, so that it contained multiple identical copies of itself. DR was performed as usual. As expected, identical patient states are plotted at identical position, and thus the expected time curve exhibits cyclic patterns (see supplementary material).

To demonstrate the applicability of our approach to other ICU data with different, more general characteristics, we applied the single patient analysis to data from PhysioNet [GAG*00]: the MIMIC-III waveform data set [MMV*20] and the MGH/MF waveform data set [WFTR91], which are openly available measurements of typical parameters recorded in daily clinical routine. While the MIMIC-III data set spans longer time-frames (long-term intensive care), the MGH data set spans only shorter time-frames. This displays the ability of our presented application to handle different types of data, especially with regard to varying length of the time series.

Figure 11 showcases an exemplary analysis of a ∼150-h patient record from the MIMIC-III waveform data set using our presented application. When all 15 variables are included in the projection, a prominent peak can be seen (Figure 11, ■). Using the auxiliary plots, we explored the variables and found possible reasons in two ECG records. After these records were removed, this peak is no longer present, supporting the hypothesis that these two were the reason for the peak. Several other features now become more prominent: a bi-partition of the whole structure, and a deflection of the time curve (■ and ■, respectively). We analysed these features in the same way as the first peak, revealing oxygen saturation (SpO2) and blood pressure as the main candidates for these features. Removing the oxygen saturation led to a less pronounced partition of the curve, while removing the blood pressure removes the loop. Afterwards, a last prominent loop-like feature can be seen (■), which, however, can not easily be related to a single variable but instead traced back to a combination of multiple variables. This could indicate a medically significant, complex process, warranting further exploration.

We also compared time curves generated using only PCA (Figure 3) to ones using CVAE coupled to PCA and UMAP (Figure 12). The most notable feature is the change in the representation of larger distances. While PCA correctly identifies the first few hours as a phase of drastic change, it does not show the smaller changes in the later phase particularly well. CVAE as well as UMAP also correctly identify this early trend, however, due to their non-linearity, scale it down. This allows for a higher resolution of the more fine-grained changes in the later stage. While this leads to a more pronounced visualisation of smaller changes when using the CVAE, the linearity of PCA might make it easier to judge global similarities. UMAP typically generates more pronounced local clusters along the time curve compared to CVAE. Hence, the users can choose between the methods in our application, based on their preference.

Multi-subject mode. As shown in Figure 6, the time curve box plot allows the user to assess the representativeness of the subjects (R5). Figure 13(a) shows the multi-subject view, where each time curve represents one of the nine animal subjects. Six of these nine subjects (Figure 13(b)) were part of a temperature control experiment [Pet18]: VNA001 ■ and VNA002 ■ had no active temperature control management, VNA003 ■ and VNA004 ■ had their temperature adjusted manually by ICU staff using a hot air blanket, VNA007 ■ and VNA008 ■ were subject to automated temperature control using a hot air blanket. As subject VNA002 ■ had a reduced experiment duration of 57 h [Pet18]—compared to 69 h for the remaining subjects—the analysis is also restricted to this 57-h window.

The group for which the temperature was not actively adjusted (VNA001 ■ and VNA002 ■) exhibits rather different progressions of their respective time curves (Figure 13(c)). The time curves of the manually adjusted group (VNA003 ■ and VNA004 ■) show a more similar trend (Figure 13(d)). The progression of the subjects with automatically adjusted temperature (VNA007 ■ and VNA008 ■) is the most similar (Figure 13(e)). Although temperature is only one of the nine included variables, it seems that the subjects can be grouped by the applied temperature adjustment. Actively adjusting the temperature may improve homoeostasis management, which is in turn visible in the time curve representing the organism state.

The feasibility of the multi-subject mode was also tested on a subset of the MGH/MF waveform data set [WFTR91] consisting of 16 patients (*mgh002–mgh021*, with *mgh012*, *mgh014*, *mgh017* and *mgh018* being omitted due to missing variables). As the data contained events assumed to be application and removal of the measurement devices, heavily skewing the projection, the tool's func-

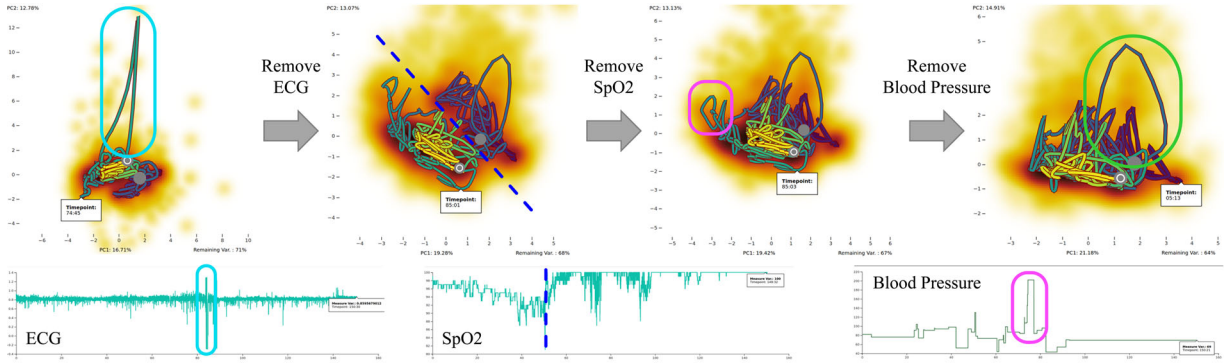

Figure 11: *Analysis of a patient record from the MIMIC-III data set using our application (patient: p042075, 15 variables). The time curve exhibits a strong peak (■) caused by a peak in two out of three ECG recordings (one is shown below). After these two variables are excluded, the bifurcated structure is revealed more clearly (■), which is at least partially caused by the oxygen saturation, as the bifurcation becomes less prominent after removing the* SpO2 *variable. Lastly, two loop structures are visible (■, ■). While one (■) can be removed by excluding two variables containing blood pressure data, the last one (■) seems to be induced by multiple variables, warranting further inspection.*

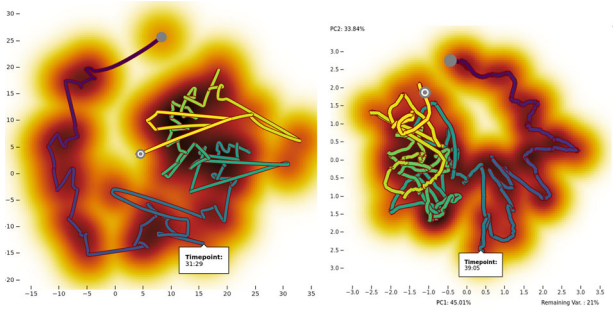

Figure 12: *Time curves for VNA004 created using UMAP (left) and our CVAE model coupled to PCA (right). See Figure 3 for PCA.*

tionality to restrict the time series was used to cut-off beginning and end of the time series. When displaying the complete time curve ensemble, cluttering is prohibiting any analysis (Figure 14, left). Our proposed time curve boxplot visualisation can be used to declutter the visualisation significantly, while still retaining some information, like the identity and shape of the most representative curve (Figure 14, centre). Inspection of the outliers—time curves that are not wholly contained either of the percentile ranges—shows that not only spatial but also temporal inclusion are of relevance when calculating the functional data depth (Figure 14, right). The scarf boxplot below visualises the true extents of inclusion in the respective hulls, both temporal and spatial. It reveals that the yellow curve showing subject *mgh020* is quite often within the 50th percentile range, but only for very short durations, which might be an interesting aspect suggesting a more detailed inspection of the individual variables. The orange curve representing subject *mgh013* is always in the 80th percentile range but only rarely in the 50th one, suggesting that this patient might be overall more dissimilar to the most representative patient, but has no strong occasional deviations like *mgh020*.

Expert feedback. To estimate the utility in clinical practice, we collected feedback from our expert co-authors as well as four medical doctors from the local university hospital. The latter were neither involved in the development of our application, nor had any other contribution to this work. They saw our tool for the first time in the feedback session. The feedback from these experts was collected in structured feedback sessions of about 1 h each, in which they were given introductions, performed three tasks and were asked eight questions. During the tasks, we used a think-aloud protocol. In the first task, the users were presented with a projected time curve with two distinctive points selected and were asked to find possible reasons for this behaviour, essentially performing a top-down workflow (R2). The second task was set up for a bottom-up workflow, with a distinctive point being selected in a single variable using the auxiliary line plot. Finally, the users were presented with the multi-subject setup described above and were asked to evaluate if the temperature groups defined by the experimental setup can be validated by the time curves course. The eight questions were designed with the requirements in mind to get a measure of the general usefulness of the tool (see supplementary material for screenshots of the prepared starting points for each task and the questions).

Overall, the users clearly understood the presented time curve as a representation of the overall patient state and its progression. All users mentioned its primary benefit in the use as a patient overview, as the tools and workflows mentioned as current clinical practice (e.g. Excel, SPSS) mainly utilize single parameters, often reduced to a few statistical key figures (i.e. mean and median), making it difficult to judge the overall patient state. One of the independent experts explicitly noted that such a visual analysis tool is '*exactly what we need*'. They rated the combination of detail plots (auxiliary and small multiples) and time curve as highly usable to judge the overall stability of a patient (R1). While the users understood the general concepts of the DR methods, several of them pointed out that—in part due to the complexity and unfamiliarity of DR—some training is needed to fully exploit the potential of the system, especially if it was to be used in a live environment.

All users were also successful in judging the stability of patients when used in a multi-subject scenario (R5). One expert noted that, due to time constraints and the large amount of generated data,

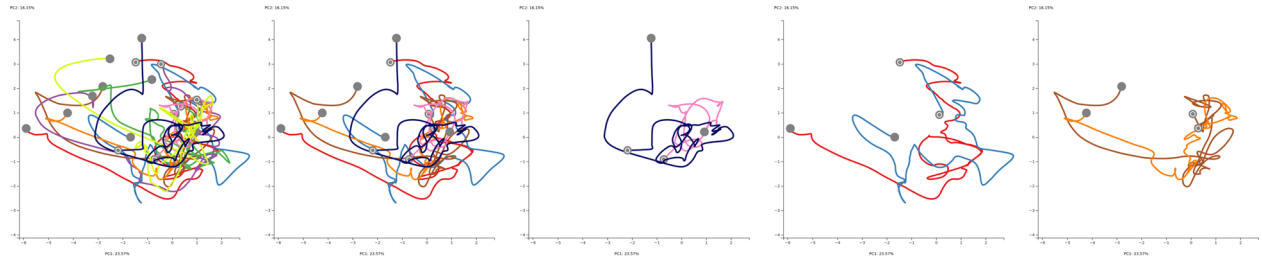

Figure 13: *(a) An ensemble of nine animal subjects shown as filtered time curves. (b) The six subjects that were part of the temperature control experiment. (c) Subjects VNA001 and VNA002: no active temperature control. (d) VNA003 and VNA004: manual temperature control. (e) VNA007 and VNA008: automated temperature control. Each time curve results from the same nine parameters of each subject.*

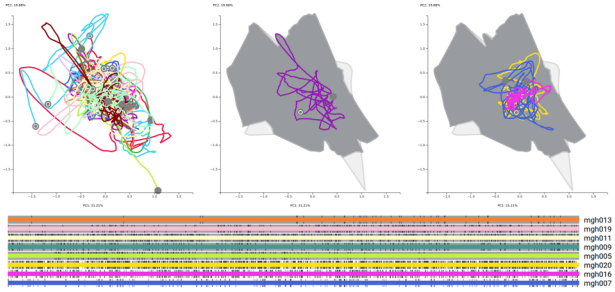

Figure 14: *Ensemble of 16 patients from the MGH/MF data set. A naïve overview of all time curves is quite cluttered (left). Our time curve boxplot (centre: combined with the most central time curve; right: only outliers) allows for detailed analysis based on functional data depth. Bottom: Scarf boxplot showing detailed information about inclusion of the curves for the respective quantile hulls.*

retrospective analysis rarely happens in clinical practice, although it would be important. They reckoned that the fast overview provided by our application—especially the time curves—could enable more frequent retrospective case analyses, which they expect to be highly useful. The top-down approach was favoured by most users. However, all users were also successful in using the bottom-up approach. One user even preferred this way of working, indicating the functionality of both approaches (R2).

When asked about the handling of missing data, the users generally either had a strong preference for the chosen method (Section 5.1, R4), or did not have a clear preference, mentioning pros and cons for either: our approach and smooth interpolation. Furthermore, clearly communicated and user-controllable filtering was strongly preferred as well, as even values perceived as 'faulty' could yield helpful medical insights.

One user noted that curve visualisations, especially as space-time cubes, are not commonly used for ICU data in the medical domain. However, they were shown to be beneficial for solving complex tasks in other domains [KDA*09]. In general, the additional auxiliary functions (i.e. space-time cube, correlation view and colour/width modes) were less frequently used. However, some of the experts explicitly noted that they consider them interesting and useful and presume that they would apply them more often after more extended use. Nevertheless, only two of them rated the space-time cube as not useful. Regarding possible additions, the inclusion of the value corridor in the auxiliary view (as already implemented by colour-coding for the time curve) was mentioned multiple times. Multiple users reckon that it would be beneficial to allow for the display of multiple variables in the auxiliary plot. Some users expressed the need for a guide to more easily find variables that exhibit significant features for time points selected in the time curve.

In conclusion, our application shows high potential and acceptance by medical experts, not only for the post-hoc analysis of experimental data, but also for fast and convenient retrospective analysis. Based on the expert feedback we collected, it seems to satisfy the increasing demand for analysis tools to cope with the increasing data output and complexity of medical surveillance devices.

8. Conclusion and Outlook

We presented a web-based visual analytics application for multivariate, time-varying ICU surveillance data. Due to the large amount of data, a comprehensive view on the patient state progression is quite difficult and traditional visualisations like line charts fail to adequately present the data. Our application uses DR to visualise these high-dimensional data as time curves to provide an overview of this progression of the patient state over time. The time curves are shown as a 2D plot and in a space-time cube. Since ICU data can have missing or erroneous values and parameters with varying sampling rates, we extended the time curves using filtering and smoothing. To prevent misinterpretation of the smoothed time curve and to show the density of the input data, the underlying data can be shown as colour-coded density in the background. A small multiples view shows details on demand about the individual dimensions as line plots and allows users to select and compare different points in time. Our visual analysis application also provides methods to facilitate the comparison of multiple patients by projecting their individual filtered time curves into one two-dimensional space. Furthermore, the system can depict an ensemble as time curve boxplots that summarize typical time curves while highlighting outliers. The time curve boxplots are complemented by a scarf boxplot that shows at which points in time the outlier curves deviate from the box. We evaluated our application with real-world data and structured feedback

session with medical experts, who consistently rated our application as useful and see a definite potential for further developments.

In the future, we also want to extend the scope of our application to live data, which would allow for monitoring of patients in the hospital. To this end, our application has to be improved and retargeted in several ways: first, it has to be ensured that the time curve does not change too drastically when adding new samples and recalculating the DR. One possible avenue is the adaption of projection techniques to allow for the progressive addition of data points. Several interesting approaches to avoid frequent, global recalculations have been proposed for streaming data (e.g. [FCS*20]). Improving projection quality, e.g. through landmark-based guidance as proposed by Vernier *et al*. [VCT21], is an active field of research. Its transferability to our application requires further investigation.

Acknowledgements

N.B. and M.K. were funded by Carl-Zeiss-Stiftung. C.S. and D.W. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG)—Project-ID 251654672—TRR 161.

Open access funding enabled and organized by Projekt DEAL.

References

[ABC*16] Abadi M., Barham P., Chen J., Chen Z., Davis A., Dean J., Devin M., Ghemawat S., Irving G., Isard M., Kudlur M., Levenberg J., Monga R., Moore S., Murray D. G., Steiner B., Tucker P., Vasudevan V., Warden P., Wicke M., Yu Y., Zheng X.: TensorFlow: A system for large-scale machine learning. In *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)* (2016), USENIX Association, pp. 265–283. https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi.

[AJXW19] Ali Mohammed, Jones Mark W., Xie Xianghua, Williams Mark (2019) TimeCluster: dimension reduction applied to temporal data for visual analytics. *The Visual Computer*, *35*, (6-8), 1013–1026. https://doi.org/10.1007/s00371-019-01673-y

[AMST11] Aigner W., Miksch S., Schumann H. & Tominski C.: Visualization of Time-Oriented Data. Human-Computer Interaction Series. Springer, London, 2011. https://doi.org/10.1007/978-0-85729-079-3.

[BDA*17] Bach B., Dragicevic P., Archambault D., Hurter C., Carpendale S.: A descriptive framework for temporal data visualizations based on generalized space-time cubes. *Computer Graphics Forum 36*, 6 (2017), 36–61. https://doi.org/10.1111/cgf.12804.

[BG05] Borg I., Groenen P. J. F.: *Modern Multidimensional Scaling: Theory and Applications. Springer Series in Statistics* (2nd edition). Springer, New York, 2005. https://doi.org/10.1007/0-387-28981-X.

[BOH11] Bostock M., Ogievetsky V., Heer J.: D^3 data-driven documents. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2301–2309. https://doi.org/10.1109/TVCG.2011.185.

[BSH*16] Bach B., Shi C., Heulot N., Madhyastha T., Grabowski T., Dragicevic P.: Time curves: Folding time to visualize patterns of temporal evolution in data. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 559–568. https://doi.org/10.1109/TVCG.2015.2467851.

[BSP*20] Brich N., Schulz C., Peter J., Klingert W., Schenk M., Weiskopf D., Krone M.: Visual analysis of multivariate intensive care surveillance data. In *Proceedings of the VCBM 2020: Eurographics Workshop on Visual Computing for Biology and Medicine* (2020), B. Kozlíková, M. Krone, N. Smit, K. Nieselt and R. G. Raidou (Eds.), The Eurographics Association. https://doi.org/10.2312/vcbm.20201174.

[BWS*12] Bernard J., Wilhelm N., Scherer M., May T., Schreck T.: Timeseriespaths: Projection-based explorative analysis of multivariate time series data. *Journal of WSCG20* (2012), 97–106.

[C*15] Chollet F., et al.: *Keras: Deep learning library for theano and tensorflow*. https://keras.io [Accessed 21st February 2022].

[CANR08] Cuesta-Albertos J., Nieto-Reyes A.: The random Tukey depth. *Computational Statistics & Data Analysis 52*, 11 (2008), 4979–4988. https://doi.org/10.1016/j.csda.2008.04.021.

[CMR05] Caat M. t., Maurits N. M., Roerdink J. B. T. M.: Tiled parallel coordinates for the visualization of time-varying multichannel EEG data. In *Proceedings of the EUROVIS 2005: Eurographics/IEEE VGTC Symposium on Visualization* (2005), K. Brodlie, D. Duke and K. Joy (Eds.), The Eurographics Association. https://doi.org/10.2312/VisSym/EuroVis05/061-068.

[dL77] de Leeuw J.: Applications of convex analysis to multidimensional scaling. In *Recent Developments in Statistics*. J. Barra, F. Brodeau, G. Romier and B. V. Cutsem (Eds.). North Holland Publishing Company, Amsterdam (1977), pp. 133–146.

[FCS*20] Fujiwara T., Chou J.-K., Shilpika S., Xu P., Ren L., Ma K.-L.: An incremental dimensionality reduction method for visualizing streaming multidimensional data. *IEEE Transactions on Visualization and Computer Graphics 26*, 1 (2020), 418–428. https://doi.org/10.1109/TVCG.2019.2934433.

[GAG*00] Goldberger A. L., Amaral L. A., Glass L., Hausdorff J. M., Ivanov P. C., Mark R. G., Mietus J. E., Moody G. B., Peng C.-K., Stanley H. E.: Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation 101*, 23 (2000), e215–e220. https://doi.org/10.1161/01.CIR.101.23.e215.

[GASM17] Greis M., Avci E., Schmidt A., Machulla T.: Increasing users' confidence in uncertain data by aggregating data from multiple sources. In *CHI'17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2017), Association for Computing Machinery, p. 828–840. https://doi.org/10.1145/3025453.3025998.

[gro14] Grottel S., Heinrich J., Weiskopf D., Gumhold S. (2014) Visual Analysis of Trajectories in Multi-Dimensional State Spaces. *Computer Graphics Forum*, *33*, (6), 310–321. https://doi.org/10.1111/cgf.12352

[HAF13] Heulot N., Aupetit M., Fekete J.-D.: ProxiLens: Interactive exploration of high-dimensional data using projections. In *Proceedings of the EuroVis 2013 Workshop on Visual Analytics using Multidimensional Projections* (2013), The Eurographics Association. https://doi.org/10.2312/PE.VAMP.VAMP2013.011-015.

[HKF16] Haroz S., Kosara R., Franconeri S.: The connected scatterplot for presenting paired time series. *IEEE Transactions on Visualization and Computer Graphics* (2016). https://doi.org/10.1109/TVCG.2015.2502587.

[Hot33] Hotelling H.: Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology 24*, 6 (1933), 417–441. https://doi.org/10.1037/h0071325.

[HSS*21] Hinterreiter A., Steinparz C., Schöfl M., Stitz H., Streit M.: Projection path explorer: Exploring visual patterns in projected decision-making paths. *ACM Transactions on Interactive Intelligent Systems 11*, 3–4 (Aug. 2021). https://doi.org/10.1145/3387165.

[JFSK16] Jackle D., Fischer F., Schreck T., Keim D.: Temporal MDS plots for analysis of multivariate data. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 141–150. https://doi.org/10.1109/TVCG.2015.2467553.

[JPS*16] Johnson A. E. W., Pollard T. J., Shen L., Lehman L.-w. H., Feng M., Ghassemi M., Moody B., Szolovits P., Anthony Celi L., Mark R. G.: MIMIC-III, a freely accessible critical care database. *Scientific Data 3*, 1 (2016), 1–9. https://doi.org/10.1038/sdata.2016.35.

[KDA*09] Kristensson P. O., Dahlback N., Anundi D., Bjornstad M., Gillberg H., Haraldsson J., Martensson I., Nordvall M., Stahl J.: An evaluation of space time cube representation of spatiotemporal patterns. *IEEE Transactions on Visualization and Computer Graphics 15*, 4 (2009), 696–702. https://doi.org/10.1109/TVCG.2008.194.

[KH13] Kehrer J., Hauser H.: Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE Transactions on Visualization and Computer Graphics 19*, 3 (2013), 495–513. https://doi.org/10.1109/TVCG.2012.110.

[KPT*18] Klingert W., Peter J., Thiel C., Thiel K., Rosenstiel W., Klingert K., Grasshoff C., Königsrainer A., Schenk M.: Fully automated life support: An implementation and feasibility pilot study in healthy pigs. *Intensive Care Medicine Experimental 6*, 1 (2018). https://doi.org/10.1186/s40635-018-0168-3.

[KW13] Kingma D. P. & Welling M.: Auto-encoding variational Bayes. arXiv preprint arXiv:1312.6114 (2013).

[LMW*17] Liu S., Maljovec D., Wang B., Bremer P.-T., Pascucci V.: Visualizing high-dimensional data: Advances in the past decade. *IEEE Transactions on Visualization and Computer Graphics 23*, 3 (2017), 1249–1268. https://doi.org/10.1109/TVCG.2016.2640960.

[LPR09] López-Pintado S., Romo J.: On the concept of depth for functional data. *Journal of the American Statistical Association 104*, 486 (2009), 718–734. https://doi.org/10.1198/jasa.2009.0108.

[LPSLG14] López-Pintado S., Sun Y., Lin J., Genton M.: Simplicial band depth for multivariate functional data. *Advances in Data Analysis and Classification 8*, 3 (2014), 321–338. https://doi.org/10.1007/s11634-014-0166-6.

[MHM18] McInnes L., Healy J., Melville J.: UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv:1802.03426 [cs, stat]* (2018).

[MMV*20] Moody B., Moody G., Villarroel M., Clifford G. & Silva I.: Mimic-III waveform database matched subset (version 1.0). https://doi.org/10.13026/c2294b

[Mo10] McKinney W.: Data structures for statistical computing in Python. In *Proceedings of the 9th Python in Science Conference* (2010), vol. 445, pp. 51–56. https://doi.org/10.25080/Majora-92bf1922-012.

[Mor13] Moreland K.: A survey of visualization pipelines. *IEEE Transactions on Visualization and Computer Graphics 19*, 3 (2013), 367–378. https://doi.org/10.1109/TVCG.2012.133.

[MRLU] Mönnich A., Ronacher A., Lord D., Unterwaditzer M.: *Flask*. https://palletsprojects.com/p/flask/ [Accessed 21st February 2022].

[Mun15] Munzner T.: *Visualization Analysis and Design. AK Peters Visualization Series*. New York: CRC Press, 2015. https://doi.org/10.1201/b17511.

[MWK14] Mirzargar M., Whitaker R. T., Kirby R. M.: Curve Boxplot: Generalization of Boxplot for ensembles of curves. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 2654–2663. https://doi.org/10.1109/TVCG.2014.2346455.

[OndL*10] Ordóñez P., desJardins M., Lombardi M., Lehmann C. U., Fackler J.: An animated multivariate visualization for physiological and clinical data in the ICU. In *IHI'10: Proceedings of the 1st ACM International Health Informatics Symposium* (New York, NY, USA, 2010), Association for Computing Machinery, pp. 771–779. https://doi.org/10.1145/1882992.1883109.

[Pea01] Pearson K.: LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2*, 11 (1901), 559–572. https://doi.org/10.1080/14786440109462720.

[Pet18] Peter J.: Assessment of Goal-directed Closed-loop Management in Intensive Care Medicine. Dissertation, Tübin-

gen, 2018. https://publikationen.uni-tuebingen.de/xmlui/handle/10900/84091.

[PKK*14] Peter J., Klingert W., Königsrainer A., Rosenstiel W., Bogdan M., Schenk M.: TICoMS—a modular and message-based framework for monitoring and control of medical devices. In *Proceedings of the IEEE 27th International Symposium on Computer-Based Medical Systems* (2014), pp. 473–474. https://doi.org/10.1109/CBMS.2014.96.

[PVG*11] Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E.: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research 12* (2011), 2825–2830.

[RAM*11] Rind A., Aigner W., Miksch S., Wiltner S., Pohl M., Turic T. & Drexler F.: Visual exploration of time-oriented patient data for chronic diseases: Design study and evaluation. In Information Quality in e-Health, Lecture Notes in Computer Science (2011), Springer, Berlin, Heidelberg, pp. 301–320. https://doi.org/10.1007/978-3-642-25364-5_22.

[SG64] Savitzky A., Golay M. J. E.: Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry 36*, 8 (1964), 1627–1639. https://doi.org/10.1021/ac60214a047.

[SG11] Sun Y., Genton M. G.: Functional boxplots. *Journal of Computational and Graphical Statistics 20*, 2 (2011), 316–334. https://doi.org/10.1198/jcgs.2011.09224.

[SNEA*16] Schulz C., Nocaj A., El-Assady M., Frey S., Hlawatsch M., Hund M., Karch G., Netzel R., Schätzle C., Butt M., Keim D. A., Ertl T., Brandes U., Weiskopf D.: Generative data models for validation and evaluation of visualization techniques. In *BELIV'16: Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization* (New York, NY, USA, 2016), Association for Computing Machinery, pp. 112–124. https://doi.org/10.1145/2993901.2993907.

[SSK10] Seifert C., Sabol V., Kienreich W.: Stress maps: Analysing local phenomena in dimensionality reduction based visualisations. In *EuroVAST 2010: Proceedings of the International Symposium on Visual Analytics Science and Technology* (2010), J. Kohlhammer and D. Keim (Eds.), The Eurographics Association. https://doi.org/10.2312/PE/EuroVAST/EuroVAST10/013-018.

[Tor52] Torgerson W. S.: Multidimensional scaling: I. Theory and method. *Psychometrika 17*, 4 (1952), 401–419. https://doi.org/10.1007/BF02288916.

[TOV20] Truong C., Oudre L., Vayatis N.: Selective review of offline change point detection methods. *Signal Processing 167* (2020). https://doi.org/10.1016/j.sigpro.2019.107299.

[Tuf92] Tufte E. R.: *The Visual Display of Quantitative Information*. Cheshire, Ct., USA: Graphics Press, 1992.

[VCT21] Vernier E. F., Comba J. L. D., Telea A. C.: Guided stable dynamic projections. *Computer Graphics Forum 40*, 3 (2021), 87–98. https://doi.org/10.1111/cgf.14291.

[vdEHBvW16] van den Elzen S., Holten D., Blaas J., van Wijk J. J.: Reducing snapshots to points: A visual analytics approach to dynamic network exploration. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 1–10. https://doi.org/10.1109/TVCG.2015.2468078.

[vdMH08] van der Maaten L., Hinton G.: Visualizing data using t-SNE. *Journal of Machine Learning Research 9*, (Nov. 2008), 2579–2605.

[VGdS*20] Vernier E. F., Garcia R., da Silva I. P., Comba J. L. D., Telea A. C.: Quantitative evaluation of time-dependent multidimensional projection techniques. *Computer Graphics Forum 39*, 3 (2020), 241–252. https://doi.org/10.1111/cgf.13977.

[WFTR91] Welch J., Ford P., Teplick R., Rubsamen R.: The Massachusetts general hospital–marquette foundation hemodynamic and electrocardiographic database—comprehensive collection of critical care waveforms. *Clinical Monitoring 7*, 1 (1991), 96–97.

[WSA*16] Wang Y., Shen Q., Archambault D., Zhou Z., Zhu M., Yang S., Qu H.: AmbiguityVis: Visualization of ambiguity in graph layouts. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 359–368. https://doi.org/10.1109/TVCG.2015.2467691.

[ZMG*10] Zhang Z., Mittal A., Garg S., Dimitriyadi A., Ramakrishnan I., Zhao R., Viccellio A., Mueller K.: A visual analytics framework for emergency room clinical encounters. In *Proceedings of the IEEE Workshop on Visual Analytics in Health Care* (2010).

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1

Video S1