DPVis: Visual Analytics with Hidden Markov Models for Disease Progression Pathways

Bum Chul Kwon, Vibha Anand, Kristen A. Severson, Soumya Ghosh,
Zhaonan Sun, Brigitte I. Frohnert, Markus Lundgren, and Kenney Ng

Abstract—Clinical researchers use disease progression models to understand patient status and characterize progression patterns from longitudinal health records. One approach for disease progression modeling is to describe patient status using a small number of states that represent distinctive distributions over a set of observed measures. Hidden Markov models (HMMs) and its variants are a class of models that both discover these states and make inferences of health states for patients. Despite the advantages of using the algorithms for discovering interesting patterns, it still remains challenging for medical experts to interpret model outputs, understand complex modeling parameters, and clinically make sense of the patterns. To tackle these problems, we conducted a design study with clinical scientists, statisticians, and visualization experts, with the goal to investigate disease progression pathways of chronic diseases, namely type 1 diabetes (T1D), Huntington's disease, Parkinson's disease, and chronic obstructive pulmonary disease (COPD). As a result, we introduce DPVis which seamlessly integrates model parameters and outcomes of HMMs into interpretable and interactive visualizations. In this study, we demonstrate that DPVis is successful in evaluating disease progression models, visually summarizing disease states, interactively exploring disease progression patterns, and building, analyzing, and comparing clinically relevant patient subgroups.

Index Terms—Disease Progression, Hidden Markov Model, State Space Model, Diabetes, Huntington's, Parkinson's, Interpretability

◆

1 INTRODUCTION

CLINICAL researchers want to understand the progression of diseases. Improved understanding of disease progression can enable a variety of clinical tasks, including patient management, cohort selection, and drug discovery. To achieve this goal, clinical researchers often employ quantitative models, termed 'disease progression models' (DPMs), to characterize the course of disease progression from patients' data contained in longitudinal health records. Using the model, researchers aim to gain insights about how characteristics of patients interact with the evolution of disease. This knowledge can ultimately lead to early detection of diseases and precision care for each patient at appropriate points in time.

However, understanding disease progression trajectories is a non-trivial task. First, modeling progression of diseases is challenging because it is often manifested by multiple symptoms over time, which makes it difficult to summarize progression patterns. Second, clinical studies usually observe the symptoms of multiple subjects over discrete, irregular time points, which make modeling the patterns even more difficult. Third, clinical researchers not only want to summarize disease progression patterns of subjects, but also analyze the relationship between their trajectories and health outcomes. Likewise, understanding disease progression patterns involve modeling the distinct state transitions from complex, longitudinal patient records and analyzing the association between

learned transition patterns and various measures (e.g., genetic profiles) in order to derive clinically meaningful insights.

Research on type 1 diabetes is an illustrative example for the importance of DPMs. In the current understanding, type 1 diabetes development can be divided into three stages: i) Stage-1: normal blood sugar; ii) Stage-2: abnormal blood sugar; iii) Stage-3: clinical diagnosis of type 1 diabetes [28]. However, the three stages assume two or more autoantibodies, and it is not yet established how individuals progress from no autoantibodies to multiple autoantibodies. There may be differences in the progression speed between pre-diabetic individuals depending on the combinations of autoantibodies that they possess at different time points. Furthermore, some individuals with different genetic predispositions may follow different trajectories of the evolution of autoantibodies. To gain more insights, researchers want to discover transition patterns between high-level 'states' that are manifested by onsets of autoantibodies, to explore subjects' transition patterns between the states, and find association between various genetic variables and trajectories.

In this study, we conducted a design study with clinical researchers so that we build a tool that helps them to visually explore disease progression patterns. Throughout the design study, we derived our users' task hierarchy and explored design space to support the tasks using interactive visualization techniques. We chose Hidden Markov Models (HMMs), over other candidates because they can infer discrete latent states and transitions between the inferred states from time-varying multivariate data. In particular, they adequately handle longitudinal medical data from clinical studies, which are often incomplete, missing, and measured at irregular time points. In essence, a trained HMM infers a state (output) per patient visit based on a set of observed measures (input) at the visit. For example, T1D researchers can train a HMM with three autoantibodies variables (input) and use the trained model to infer states (output) of patients' records over time. If they

- *Bum Chul Kwon, Vibha Anand, Kristen A. Severson, Soumya Ghosh, Zhaonan Sun, and Kenney Ng are with IBM Research. BCK, VA, and KN are also with the T1DI Study Group. E-mail: {bumchul.kwon, anand, kristen.severson, ghoshso, zsun, kenney.ng}@us.ibm.com.*

- *Brigitte I. Frohnert is with University of Colorado Denver and the T1DI Study Group. E-mail: brigitte.frohnert@ucdenver.edu.*

- *Markus Lundgren is with Department of Clinical Sciences Malmö, Lund University, Sweden and the T1DI Study Group. E-mail: markus.lundgren@med.lu.se.*

choose a model with K states, then each visit will be assigned to one of K states. The number of states is a hyperparameter, which is typically chosen based on cross validation or prior knowledge. Typically the number of visits is far larger than the number of states thus causing multiple visits to be assigned to the same state. The characteristics of the state — a statistical summary of all visits assigned to the state, is learned automatically from the data. Using the states inferred (labeled) by the HMM, researchers can understand the evolution of autoantibodies of multiple patients by inspecting which states subjects tend to go through at different ages and which state transitions occur more frequently. As a result of our design study, we developed a visual analytics system called DPVis that incorporates HMMs with interactive visualizations. Our case study demonstrates the usefulness of DPVis for disease progression analysis. The main contributions of the paper are summarized as following:

1) We developed DPVis that incorporates HMMs with interactive visualizations for exploring disease progression patterns from longitudinal health records.

2) Our design study examines tasks of clinical researchers who aim to understand disease progression stages and provides various ways to visualize the statistical outcomes of HMMs for supporting these tasks.

3) We report a usage scenario and users' experiences that demonstrates the usefulness of DPVis for domain experts to interpret HMMs in a transparent manner, to detect interesting disease progression patterns, and to derive clinical insights.

2 RELATED WORK

In this section, we review disease progression studies in the field of medicine, modeling techniques in statistics and machine learning, and visual analytic approaches in the visualization domain.

2.1 Disease Progression Research in Clinical Studies

Well-designed longitudinal studies play an important role in clinical research because they provide rich information for tracking the progression of disease by collecting repeated measures from individuals of a target cohort over a long period of time. Such studies are designed and conducted with many present and future goals, such as understanding the etiology or pathogenesis of a disease. One of the most well-known longitudinal studies is the Framingham Heart Study, which has brought great insights into heart disease (e.g., [27], [38]). In particular, many studies call for research into applying suitable statistical modeling techniques that can explain the mechanism of disease progressions. For example, the Parkinson Progression Marker Initiative (PPMI) [45] is an international observational multi-center study of Parkinson's Disease. PPMI established protocols that were adopted at 21 clinical sites to provide a platform for researchers to gain common access to the data, which includes clinical, socio-demographic, and imaging variables as well as bio-specimens. In case of T1D, researchers seek to understand the diverse pathways from a healthy state to the onset of disease for both individual patients and the overall population [78]. The Type 1 Data Intelligence (T1DI) study group was established to accelerate disease progression research by combining decades of data from birth cohort studies conducted at various institutions around the world: 1) Diabetes Autoimmunity Study in the Young (DAISY) [59]; 2) Diabetes Prediction in Skåne (DiPiS) [29]; 3) Diabetes Prediction and Prevention (DIPP) [50];

and 4) Diabetes Evaluation in Washington State (DEW-IT) [73]. We conducted a design study with clinical researchers and statisticians who have been involved in the research groups above.

2.2 Disease Progression Modeling

Disease Progression Models (DPM) [48] include a broad class of models that quantitatively predict disease status over time. DPM can be categorized by algorithm into systems biology, data-driven, and semi-mechanistic models [11]. There are many different approaches to model disease progression, including path models [69], tree based models [12], hierarchical latent variable models [61], Gaussian process models [43], and Bayesian networks [5]. In this study, we have chosen to focus on Hidden Markov Models (HMMs), a class of probabilistic generative models commonly used to model data with longitudinal dependencies. In an HMM, the target disease is represented by a fixed number of typical disease states, its progression is characterized by a Markov process, and the observed measures are the manifestations of the underlying disease states. By characterizing the progression of disease as a series of transitions between typical disease states, HMMs provide an inherently interpretable summary of the progression. Moreover, HMMs exhibit several desirable properties that make them well suited for disease progression modeling. First, they are probabilistic in nature and are able to represent uncertainties in clinical data organically. Second, they are able to handle missing data, a common issue in observational studies. Third, users can apply their domain knowledge into modeling disease progression dynamics by setting various constraints. For instance, users can prohibit a model from stepping backward in diseased states when they investigate chronic diseases, where the damage caused by the disease is irreversible and affected areas gradually lose their functions, and backward-enabled progression, where patients can recover from the diseases. Fourth, HMMs are unsupervised models that do not require a particular outcome task. Fifth, they provide interpretability by a small number of 'states' that are associated with multiple observed variables. Examples of variants of HMMs applied to disease progression include [23], [39], [66], [67], [72]. HMMs are described further in Section 3.3. Despite the advantages, it still remains challenging for clinical scientists to use the HMM-based models for exploring disease progression patterns without visual aids. To use the disease progression model, experts need (i) to understand the output probabilities of diverse features for individual states, (ii) to investigate differences in state transition patterns between patients, and (iii) to associate such patterns with patient outcome variables (e.g., onset, death, discharge).

2.3 Visualizations for Sequential Event Analysis and Disease Progression Analysis

In this section, we review previous studies that show inspiring examples to tackle our problem. Prior studies have investigated various visualization methods to represent temporal event sequences. Shneiderman and Plaisant [63] summarize the challenges of analyzing temporal sequence data. Several techniques were proposed to integrate sequence mining algorithms for visualizing the summary of event sequences, including LifeLines [55], LifeLines2 [71], EventFlow [47], and LifeFlow [75], which allow users to visually align and explore various patterns of multiple event sequences. Frequence [53], Care Pathway Explorer [54], and Peekquence [36] visualize frequent event sequences mined from the SPAM algorithm [3]. Outflow [74] and Sequence Synopsis [9] allow

users to interactively explore multiple pathways in event sequences. These techniques allow users to steer the algorithm with interaction so that they can find meaningful summaries of event sequences. However, these studies often deal with pre-defined event types that are often independent of feature values associated with each event, which is different from statistically learned states that are associated with various features in HMMs. Thus, the visualization approaches for pre-defined event types fall short in revealing the relationship between the HMM states and the associated variables.

Other techniques integrate automated results with user-defined queries or criteria. TimeStitch [56], Choronodes [57], and Co-quito [31] extract frequent mining patterns using algorithms like PrefixSpan [52], and then visualizes the sequences so that users can interactively provide relevant feedback to refine the search. Eventpad [7] and $(s|qu)eries$ [76] allow users to build search queries using the progressive visual analytics (PVA) paradigm [65]; the idea was further extended and implemented as the PPMT tool [58]. EventThread [22] visualizes clusters of event sequences using tensor analysis. Liu et al. [41], MAQUI [37], Cadence [19], and VASABI [51] allow users to recursively explore hierarchical patterns in event sequences. ET^2 [21], StageMap [8], and Mathisen and Grnbk [46] introduce a composite event sequence to aggregate patterns. CoreFlow [40] and Guo et al. [20] visualize branching alternative paths with uncertainties. IDMVis [77] allows users to fold and align records to derive event sequence patterns. Though query-based techniques are useful and inspiring, it requires more sophisticated operations to build queries based on various characteristics of HMMs, such as posterior distributions (uncertainties) over observed variables and time ranges of event occurrences.

Previous studies also investigate visual analytic methods to explore and discover patterns from longitudinal data in clinical studies and electronic medical records. DecisionFlow [18] allows users to explore multidimensional event sequences using multiple, coordinated visual analytics systems. Researchers also investigate methods to analyze heterogeneous cohorts using interactive visualizations [1], [30]. PhenoBlocks [16], PhenoStacks [15], and PhenoLines [17] allow users to explore and compare cohorts based on various measures. RetainVis, together with RetainEX, allows users to understand how individual visits and categorical variables contribute to making diagnostic risk predictions [32]. Clustervision [33] help users to find informative cohorts of patients based on their common diseases, treatments, diagnostic measures, and comorbidities. Bernard et al. developed and evaluated static dashboard network that can help users to observe longitudinal changes of multiple patients [6]. Other studies help users make sense of time-series health data, such as Stroscope [10], TimeS-pan [42], Marai et al. [44], and RegressionExplorer [13]. Previous studies show inspiring examples to coordinate multiple views for event sequence visualization. DPVis needs to tailor these approaches for showing the similarities and differences in state transition patterns between a variety of different subgroups of patients that can be defined from each view.

Though the previous approaches provide inspiring techniques, there is no unified approach that integrates state-space models such as HMMs for exploring disease progression patterns from longitudinal observational data. Our literature review shows that we need to adapt the approaches for our data, model, and tasks. The following section describes the adaptation process and result.

3 DESIGN STUDY METHOD

This section describes our design study by introducing characteristics of domain experts, data, HMMs, and tasks and requirements that experts intend to achieve through visualizations.

3.1 Target Users: Clinical Researchers

We joined a collaborative project established in 2017 by JDRF and their academic partners for computational modeling of T1D: the T1DI (Type 1 Data Intelligence) study group [26]. Other research groups with progression modeling studies for Parkinson's and Huntington's disease were invited to collaborate and provide feedback. Our primary target users are clinical researchers, whose goal is to investigate disease progression patterns in observational data collected from clinical studies.

3.2 Domain Goals and Tasks

In the *T1DI* study group, we organized more than ten conference calls and four workshops between October, 2018 and October, 2019 to discuss disease progression trajectories of type 1 diabetes. The participants were clinical researchers who conduct and investigate observational studies on children with type 1 diabetes. In the meetings, we focused on identifying clinically meaningful research questions and hypotheses domain experts would like to ask on the observational data through iterative discussion. Authors, as participants of the meetings, consolidated the questions and hypotheses into three high-level research goals below (**G1–3**).

A goal of clinical researchers is to develop optimal treatments tailored for individuals by estimating their disease progression patterns precisely. Thus, they want to be able to discover distinct states of patients that are associated with clinical diagnosis. Early and accurate detection of presymptomatic progression signals can lead to early intervention and precision medicine for target patients. To achieve this goal, clinical researchers want to characterize and summarize the disease progression patterns by a set of biomarkers, lab tests, and other measures in a data-driven, unsupervised manner. Ultimately, clinical researchers hope to gain clinically useful insights about the disease progression patterns through subjects' longitudinal data collected from observational studies.

Clinical researchers want to understand disease progression patterns using observation data. Understanding disease progression patterns can be divided into three analytic goals. First, researchers want to *explain disease progression* by summarizing the evolution of selected variables (**G1**). Explaining disease progression includes the discovery of distinctive states that are characterized by the selected variables. For example, researchers want to describe the evolution of autoantibodies that subjects possess and lose before they get diagnosed with type 1 diabetes. In doing so, they want to find the unique combination of multiple autoantibodies multiple subjects show before diagnosis. Second, they want to *discover heterogeneous trajectory groups* (**G2**). They want to find distinctive trajectory groups of patients whose records show common patterns in terms of the evolution of selected variables. For example, clinical researchers want to identify typical progression patterns in terms of the order of autoantibodies subjects gain before they are diagnosed with type 1 diabetes. They also want to discover how many subjects follow a specific progression pattern. Lastly, they want to *find associations between specific trajectories and variables* (**G3**). For example, they may hypothesize that subjects with a specific genetic profile may show rapid progression by gaining multiple autoantibodies simultaneously at an early age.
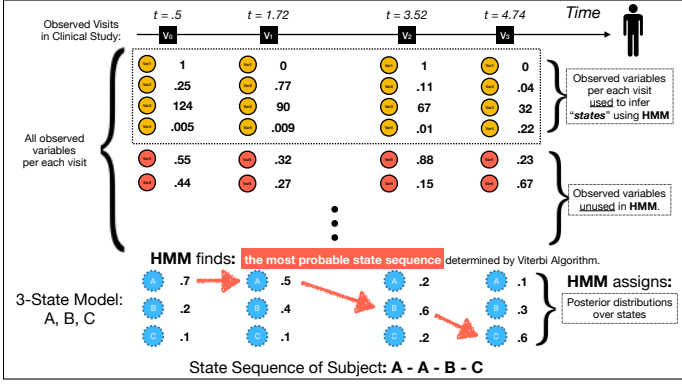
Fig. 1: HMMs learn disease states based on observed attribute values chosen by users for training the model. The trained model predicts the most probable state sequences for a patient and assigns posterior probabilities for each visit.

3.3 Observation Data and Hidden Markov Models

To expose distinct states of disease progression from the evolution of observed variables, we use HMMs to model sequences of hidden states over time on longitudinal observational data. As Fig. 1 shows, a patient's record includes multiple visits (black squares) spread over different time points (t). In each visit, a patient goes through multiple procedures, which generate observed variables (circles). Clinical researchers choose a set of variables (yellow circles) among them in order to discover hidden states using HMMs; we call the variables as output variables. Prior to training, in addition to the choice of output variables, they decide how many hidden states they are going to infer and what constraints they want to impose on progression dynamics. To determine the number of states, researchers conduct cross-validation analysis to measure the model fitness and evaluate the usefulness of the model instances based on clinical knowledge by exploring exposed patterns. Constraints can be interpreted as any restriction in state transitions. If we consider a 3-state model as Fig. 1 shows, we can set a constraint on transition probabilities so that a patient can only advance to the next step at a time without going backward or skipping. Researchers often choose the restrictions for more easily interpretable trajectory groups or based on prior knowledge about diseases (e.g., no-backward restriction for chronic diseases). Once the training is done, researchers used the trained model to infer hidden states for every visit. As part of inference, the model generates posterior probabilities (blue dotted circles) over hidden states discovered. Clinical researchers can investigate the state sequence of subjects (pink arrows). The HMM outcomes, hidden-state labels and posterior probabilities over them for every visit, combined with observed variables and other demographic variables are great sources to understand granular disease progression patterns of multiple subjects.

3.4 Visual Disease Progression Analysis Tasks

In this section, we translate domain tasks into visualization tasks by considering the outputs of HMMs discussed earlier. The first three tasks (T1, T2, T3) are drawn from the three domain tasks. Then, T4 and T5 were added to address researchers' need to investigate individual patients' records and to analyze complex subgroups, respectively. The tasks are used to guide our design decisions on various features of DPVis.

T1: Characteristics of states: To understand the distinct states discovered by HMM-based models, clinical researchers want

to view the distribution of multiple variables per each state. Understanding what each state means is also important for users to evaluate whether the discovered states clinically make sense. Understanding characteristics of states includes gaining an overview of distinct states, understanding the evolution of variables for each state, and understanding differences between the states.

T2: State transition patterns: Clinical researchers want to discover the state transition patterns from multiple subjects. For each state transition pattern, clinical researchers want to investigate details, such as the number of subjects who follow the pattern and when subjects make transition from one state to another state. Investigation of state transition patterns involve discovery of distinct state sequence patterns, understanding the heterogeneity of the patterns with respect to the time of transitions, and comparison between trajectories by the number of subjects, speed of progression, and how many states are involved.

T3: Relationship between trajectories and variables: Clinical researchers want to find the association between state transition patterns and variables like health outcomes and static variables. For example, they want to investigate whether certain trajectories correlate with the onset of the target disease. They also want to find association between other inherent characteristics of subjects, such as genetic profiles, and trajectories. Thus, they want to view distribution of static/outcome variables for trajectory groups, view the onsets of disease for trajectory groups, and compare the static/outcome variables between trajectory groups.

T4: Subject details: Clinical researchers want to investigate details by viewing state transitions, health outcomes, and other variables of a single subject. They also want to find similarities and differences between multiple subjects with respect to observation counts, times, and state transition patterns.

T5: Subgroup management: Clinical researchers build and refine subgroups based on transition patterns, health outcomes, and subject profiles. Then, they compare similarities and differences among them. Subgroups are important outcomes of their research, which need to be retained and transferable to other colleagues.

4 DESIGN OF DPVIS

Based on the derived tasks, we designed and assessed multiple views and operations and integrated them into a visual analytics application. In this section, we introduce the design of DPVis and how each view and interaction feature supports the user's analytic tasks. DPVis consists of seven view panels as Fig. 2 shows, and each plays a different role for one or more of five tasks we described in Section 3.4. We built DPVis using Python for the backend and Javascript for the frontend. We used Flask, Jinja, and Django for the web framework, and primarily used D3.js, jQuery, and Lodash to implement UI components, visualizations, and interactive features. In this section, we use type 1 diabetes as a running example to describe insights and clinical explanations from observational data. We explore disease progression patterns of 559 subjects, who were diagnosed with type 1 diabetes at the end of their observation, from birth cohort studies, using an 11-state HMM model. Find more details about the study, model, and data in Section 5.

4.1 State Summary Panel

State Summary Panel (Fig. 2 (B)) contains two views, Feature Matrix and Feature Distribution, which characterize states with respect to observed attributes. Using each view, users can gain an overview and details about states by viewing the distribution of
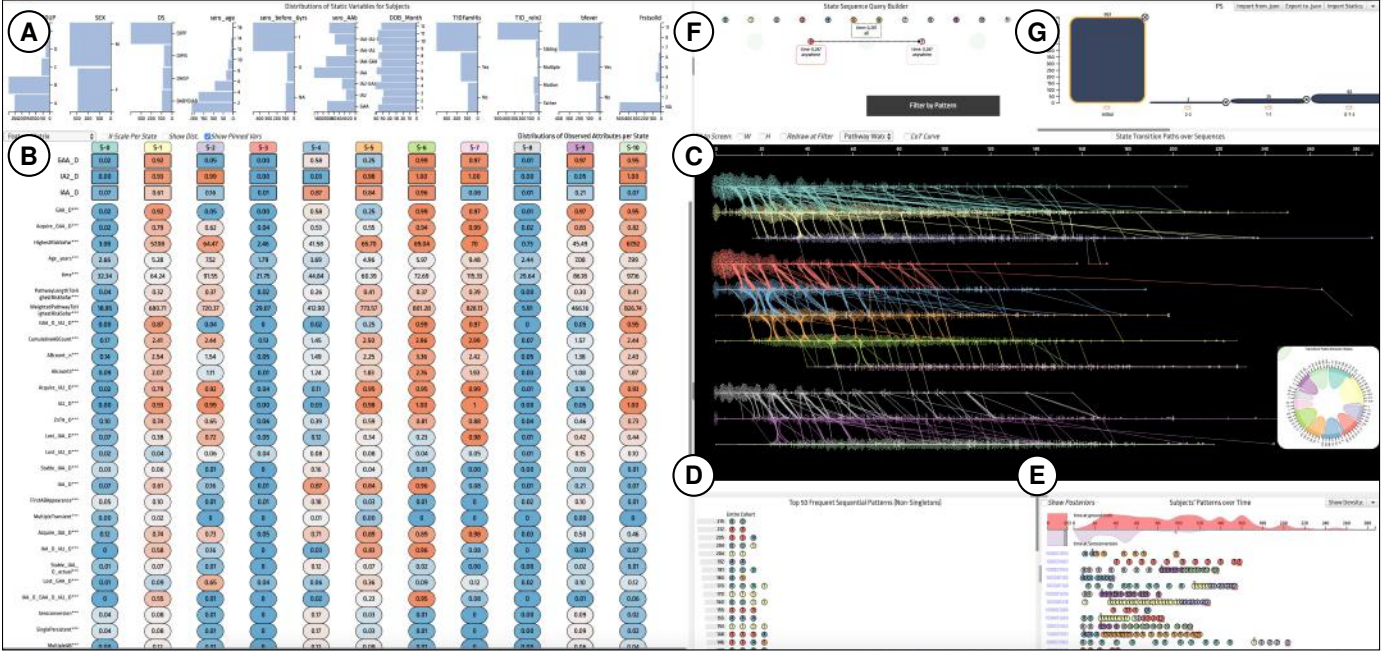
Fig. 2: Given a dataset of patients over time and the corresponding disease state assignment per visit as determined by an HMM, DPVis creates the following views: (A) Static Variables View shows distribution over static variables; (B) Feature Matrix summarizes states, discovered by HMM, with means of variables; (C) Pathway Waterfall shows state transition patterns; (D) State Transition Pattern List show frequently occurring state transition patterns; (E) Panel for Patterns and Outcomes shows subjects' visits in detail; (F) State Sequence Finder enables users to build and refine cohorts; (G) Subgroup List allows users to create and refine subgroups.

observed attributes per state (**T1**). Furthermore, users can make sense of differences between states that are discovered by HMMs.

4.1.1 Feature Matrix

Feature Matrix summarizes how each state is manifested by different observed attributes (**T1**). Feature Matrix shows a matrix layout where a row represents an observed attribute and a column represents a state; each cell in this layout represents the mean of the attribute (row) values of the visits that were inferred as the corresponding state (column). The diverging color map is used to map the range of minimum-maximum values of each measure to the blue-red scale across the states. Using this view, users can understand what each state (column) means by viewing the distribution over multiple attributes (row). In Fig. 2 (B), the leftmost column is 'State 0', which has cell values in mostly blue. This indicates that the 'State 0' indicates states, where subjects have low values for the majority of observed attributes. Thus, researchers can note that the state is likely to be before the onset of autoimmunity. In particular, we repeated and fixed attributes (rows) that are used to infer states at the top so that users can quickly refer to them while exploring state transitions in Fig. 2. At first, we considered a heatmap-style visualization where no space exists between rows and between columns. While the heatmap approach can be useful to detect correlation, users had difficulty in isolating column-wise and row-wise patterns due to the absence of borders. This can be potentially resolved by adding highlights when users hover particular state. However, users quickly glance at this view to be reminded about state characteristics overview, so it is cumbersome to hover over the view every time. Therefore, we decided to use the button-like approach for Feature Matrix.

4.1.2 Feature Distribution

The states are probabilistically inferred by HMMs, so users often need to understand standard deviation as well as mean
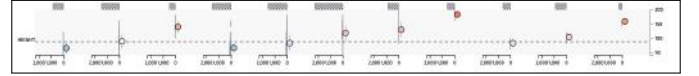

Fig. 3: Feature Distribution shows mean, standard deviation, and histogram of an attribute 'HEIGHT' across 11 states.

for each observed attribute. Greater uncertainties may indicate that it is possible to have sub-states within a single state, so users can expect more granular, diverse patterns with the state. In Feature Distribution, users can view mean, standard deviation, and histogram of each observed attribute per state in a single cell, replacing colored cells in Feature Matrix. In addition, users can check the number of missing values per cell with a shaded bar above each cell. Fig. 3 shows the distribution of an attribute 'HEIGHT' across 11 states. First, there are many missing values for the 'HEIGHT' measure, which can be observed by large, shaded bars above each state. We can conjecture that such growth variables (e.g., height, weight, BMI) were not regularly measured through the observational study. With the limitations in mind, the mean and standard deviation shows the steady increase of HEIGHT within each of three notable state sequences, namely 0–2, 3–7, and 8–10. Within each block, the mean height increases. Also, the last state of each block shows narrow intervals (standard deviation) and dense histogram, which may indicate that those last states are likely to be 'sink states' where subjects get diagnosed with Type 1 Diabetes. We implemented several ways to customize the view further. First, users can switch and toggle back to fit the scale for each cell. If users choose a consistent axis, users can compare the number of visits assigned to states. If users switch to fit the scale for each state, users can clearly observe distribution among individual state, which could be visually hidden on a common scale. Second, users can simplify the view by hiding bars. Users revealed that it is easier to compare mean and variance between selected subgroups and all subjects when there are no distracting bars.
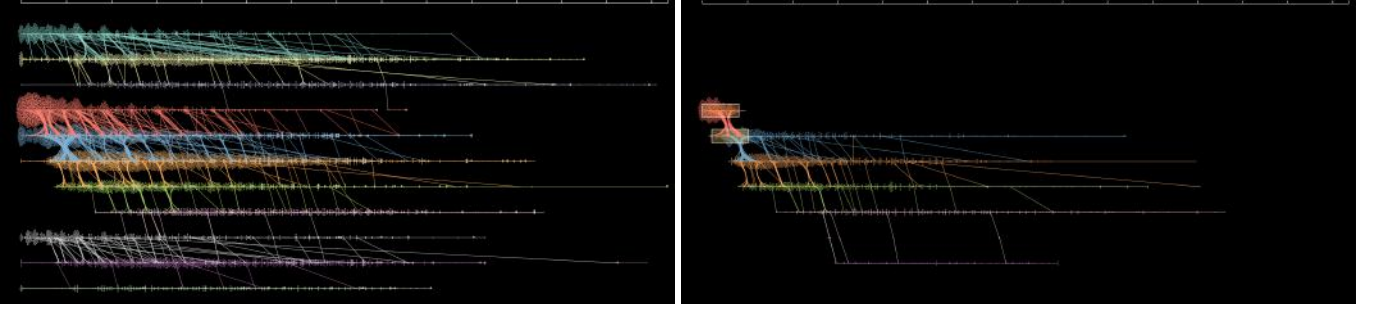
Fig. 4: Pathway Waterfall shows state transition pathways using parallel beeswarm plots and trajectory lines tied with force edge bundling: dot (visit), y-axis & color: state, x-axis: age. Vertical paths show transitions between states. **Left**: state transition patterns of all subjects over; **Right**: users filtered by state transition from state 3 to 4 before 20 months of age.

4.2 State Transitions Panel

Pathway views show various representations of state transition patterns over sequences of visits or over time (**T2**). Together with State Summary Panel, State Transitions Panel (Fig. 2 (C)) includes views that help users to understand subjects' state transition patterns. State transition patterns can be represented in many different ways as shown in previous studies in Section 2. State Transitions Panel includes five views that users can choose from. Each view has advantages and disadvantages. In the following sections, we discuss the design of views and their trade-offs.

4.2.1 Pathway Waterfall

Pathway Waterfall shows subjects' state transition patterns using parallel beeswarm visualizations overlaid with individual subjects' trajectory lines. As Fig. 4 shows, we first divide the vertical space by the number of states, each of which represents a state. A dot indicates a visit of a subject, which is colored corresponding to its state and placed on the timepoint (X) and the state space (Y). In this process, we show density of visits while preserving their timestamps using the beeswarm algorithm [14]. Then, we visualize trajectory lines for individual subjects that follow the state transition pathways over time. To reduce the clutter while preserving the transition points, we use a modified version of force-edge bundling technique [25]. To contrast the patterns more vividly, we used a dark background color. A participant commented on the design of Pathway Waterfall, "it looks like Mardi Gras beads hanging on the wall. Easy to see where (in which state) they hook and tangle up (congested departing/landing time periods per state)."

The overview in Fig. 4 (Left) shows three distinctive trajectory groups, showing state transitions: i) 0–2; ii) 3–7; iii) 8–10. Together with state characteristics found using State Summary Panel, we observe that each of trajectory group starts with no autoantibodies. The three trajectories show heterogeneous progression pathways

with respect to their first autoantibodies: i) multi-AAB First (0–2); ii) IAA First (3–7); iii) GADA First (8–10). We may want to create a subgroup among the IAA First (3–7) trajectory, who gains IAA early before the age of 20 months. To achieve the goal, we apply filters by specifying transition points, as shown in Fig. 4 (Right). The updated figure shows a onward progression pattern of the subgroups. Those who early progressed into IAA (4) seem to further possess more autoantibodies in early ages, as shown in thicker vertical paths drawn in the leftmost area. Users can also observe outliers who progressed beyond state 7 into states 8 and 9.

4.2.2 Pathway over Observation

Pathway over Observation allows users to gain an overview of transition patterns by comparing the height of stacks over visits first (**T2**). Pathway over Observation shows state transition patterns of multiple subjects using the sankey diagram approach. As Fig. 5 shows, the view is made of nodes and paths. Nodes are vertically stacked for each visit (x-axis), where each stack's height represents the number of subjects with the corresponding state at the visit. Paths are drawn horizontally between two consecutive visits, where each path connects a state (State-A) of the previous visit (Visit-0) to another state (State-B) of the current visit (Visit-1). The thickness of each path represents the number of subjects who transitioned from State-A to State-B over the visits between 0 and 1.

The main advantage of this view is summary of state transitions over a series of observations. Sankey diagrams allow users to compare the volume of transitions between consecutive visits. Alignment allows users to understand the state transition patterns with respect to specific states. Fig. 5 (Left) shows, by shifting the alignment to focus on the first trajectory group (Multi-AAB First), we observe that subjects in the trajectory tend to have zero autoantibody (green; state 0) in early age, many of them gain multiple autoantibodies (yellow; state 1) as they grow, but rarely
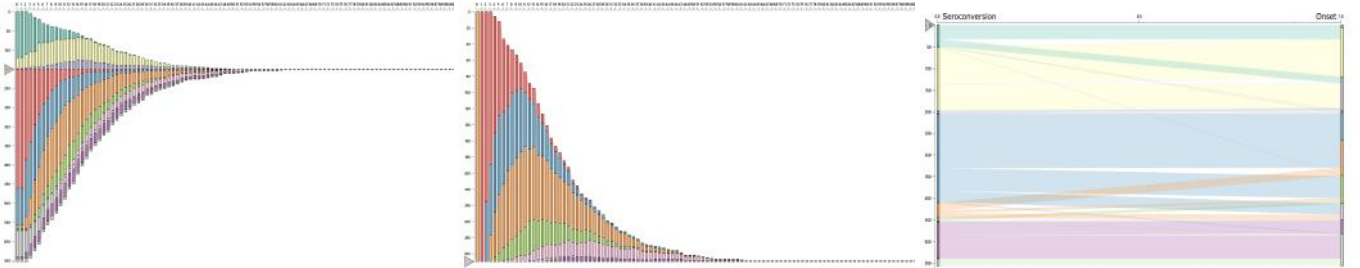

Fig. 5: Pathway over Observation summarizes state transition pathways for subjects over observed visits. **Left**: Alignment Handle rearranges nodes and paths with respect to a new anchor point (between states 2 and 3), showing the proportion of State 0 decrease over first 10 visits; **Middle**: Pathway over Observation shows pathways for subjects who started their visit with State 0 (green); the first node is highlighted in orange after users click to filter; **Right**: Users converted the view to only show start and end nodes and the path between the two nodes, showing in which state subjects seroconverted and get diagnosed (onset).
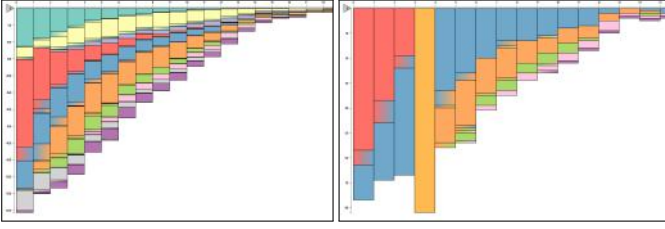
Fig. 6: Pathway by Time Unit shows state transition pathways. **Left**: Pathway by Time Unit shows stacked bars aligned at an anchor point between states 3 and 4; **Right**: Users set a filter by clicking on State 4 at subjects' age between 4 and 5 years old.

lose GADA and IAA (purple). Furthermore, users can simplify the state transitions using the biparite Sankey diagram so that they can focus on the relationship between state transitions and health outcomes. Fig. 5 (Right) shows sankey diagrams that summarize in which states subjects show seroconversion (left) and onset of type 1 diabetes (right). The first autoantibody state per each trajectory, namely state 1 (yellow), state 4 (blue), and state 9 (purple), are the most common states for seroconversion. On the other hand, subjects get diagnosed with type 1 diabetes in various states. The downside of this approach is that it aggregates individuals, so users cannot view or choose individual patients from this view.

4.2.3 Pathway by Time Unit

Pathway by Time Unit aggregates state transitions with respect to a time unit of users' choice. Thus, the patterns that one can observe are similar to Pathway over Observation, but it shows the temporal patterns of state transitions. In Fig. 6 (Left), the overview shows state transition patterns aggregated per subjects' age in 'years'. State 3 (red), which represents no autoantibody, is the most prevalent state for subjects. It is interesting to observe that some subjects show early transition from no autoantibody (State 3) to IAA (State 4) before one year old, shown as a stack with red-to-blue gradient on the leftmost bar. We can set filters by showing subjects who went through State 4 (blue) at their age between 4 and 5 years old as Fig. 6 (Right) shows. The majority of subjects possessed IAA at State 4 a year before and many of those did not lose or gain a new autoantibody afterward, as seen from bars in blue throughout the observed periods. The advantages and disadvantages of this approach are similar to those of Pathway over Observation. One distinguished advantage is to be able to show state transition with respect to ages rather than visits, which is often more relevant to clinical observational research.

4.2.4 State Transition Chord Diagram

State Transition Chord Diagram shows the overall state-to-state transition patterns over all consecutive pairs of observations from subjects using a radial network diagram. The view can be used to understand the overall frequency of state transitions (**T2**). This view serves as an explanatory summary of the state transition learned

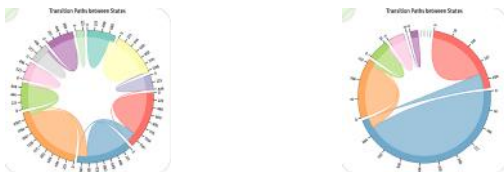Fig. 7: State Transition Chord Diagram shows a summary of state transitions between all pairs of consecutive states.**Left**: It summarizes state transitions for all subjects; **Right**: It shows subjects who had State 4 between 3 and 4 years old.
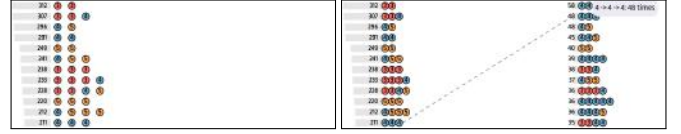

Fig. 8: State Transition Pattern List shows the result of pattern mining algorithm (BIDE) [70]. **Left**: The top-50 most occurring state transition patterns; **Right**: A new pattern list for the cohort who were breastfed.

from the particular setting of HMMs. It allows users to compare the number of visits per state by viewing the size of the nodes (states) on the circular layout and to compare the transitions between two states by the size of arcs connecting two nodes inside the circular layout. As Figure 7 (Left) shows, state transitions occur rarely. On the other hand, a cohort with subjects who had IAA with State 4 (blue) between ages of 3 and 4, as shown in Fig. 7 (Right) shows a greater proportion of transitions among States 3 (red), 4 (blue), and 5 (orange) than the rest of states. Users can filter subjects who contain specific subpatterns by clicking an edge of interest **T5**. State Transition Chord Diagram summarizes state transitions by frequency. The summary is useful for users to determine the overall prevalence of states and their transitions. Such patterns cannot be easily extracted in other views. On the other hand, the view does not provide any temporal context (age) of state transitions. Thus, the view is useful along with other views like Pathway Waterfall.

4.2.5 State Transition Pattern List

State Transition Pattern List in Fig. 8 shows a list of frequently occurring state transition patterns mined from all consecutive observations (**T2**). By running the BIDE algorithm [70] on subjects' state sequences, we first generate the most frequently occurring state sequence patterns. To avoid filling up the list with fine granular patterns, we filtered top 50 patterns in terms of the number of unique subjects with the corresponding pattern; we also do not take into account any single-state patterns. Then, we list all patterns, where each pattern is a row of state circles, in a vertical order, as shown in Fig. 8 (Left). Inspired by Peekquence [36], we also visualized the frequency of the patterns as a bar chart next to each pattern. As Fig. 8 (Right) shows, once users set filters for those who possessed IAA (state 4 in blue) between 4 and 5 years old, the view expands and introduces yet another list of patterns. By doing so, users can check the ranked list of patterns for those who possess and maintain IAA at early age.

As discussed, each view provides a unique perspective of state transition patterns. It is impossible to show the different perspectives in a single view, which motivates the options we have provided. Furthermore, different views enable opportunities for refining subgroups. For instance, users can create a subgroup by setting filters on Pathway Waterfall, State Transition Chord Diagram, and State Transition Pattern List. In the design process, we learned that supporting subgroup building activity is facilitated by multiple, different views, which we will revisit in Section 6.

4.3 Panel for Patterns and Outcomes

Panel for Patterns and Outcomes (Fig. 2 (E)) shows individual subjects' observations over time (**T3**). The panel includes two views: 1) Dual Kernel Densities View; 2) Subject List View.

4.3.1 Dual Kernel Densities View

Dual Kernel Densities View shows the two independent 1-D Kernel Density (KD) diagrams, each with increasing y-axis (density) in the
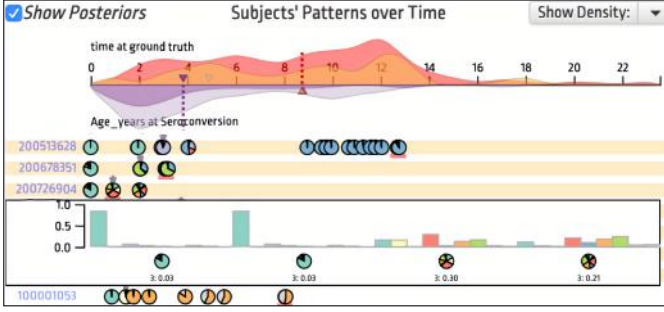
Fig. 9: Panel for Patterns and Outcomes shows 1) Dual Kernel Densities View and 2) Subject List View. Dual Kernel Densities View overlays additional dual Kernel Density Diagrams for a selected cohort, and Subject List View highlights subjects in the cohort. Users can open a popup to view posterior distributions.

opposite direction, upward and downward, respectively, as Fig. 9 shows. The KD above the timeline shows the density of time at the onset (e.g., diagnosis of type 1 diabetes) (**T3**). Subsequently, the KD below the timeline can be freely chosen by users, with the dropdown menu on the top-right corner, among the secondary onset variables (e.g., seroconversion for type 1 diabetes) (**T3**). Each KD includes an arrow indicating the mean age of its onset for subjects. Once users set filters, the KDs introduce and overlay brighter KDs inside, which indicate KDs for the selected cohort (**T5**). Fig. 9 shows the cohort who were never brestfed. The density chart in orange above the timeline suggests that the cohort' mean diagnosis age is at 8-9 years old with a peak around age 12 and a long tail towards older ages. On the other hand, the purple density below the timeline indicates that the cohort has earlier seroconversion than the overall population (light purple) (**T3**).

4.3.2 Subject List View

By viewing the state sequences of individual subjects in Subject List View, users can inspect the details at the patient level (**T3**). Subject List View provides a list of subjects' observations as dots (visits), color-coded based on the state, over a timeline. For each subject, we draw a narrow, red rectangle below a dot for a visit with the onset of the disease and an asterisk symbol above a visit with the onset of the secondary onset variable (**T3**). In Subject List View, users can choose to show the posterior distribution over all states as a pie chart. By observing how slices of pies change their values over time, users can learn states of visits, inferred by the HMM (**T1**). Fig. 9 shows the cohort of patients who were never breastfed. Most dots present full circles (dominant with a slice), which show model's confidence in the state assignments. However, a subject on the popup shows evenly-spread posterior distrbutions over 11 states for the last two visits, which signify uncertainties of state assignment for those visits. The subject shows seroconversion (purple star above pie) and diagnosis (red underline below pie) at the same visit before 1 year old. Both onsets rarely occur at the same visit, which is why the model thinks of the visits as 'uncertain' with posterior values spread over multiple states.

4.4 Static Variables View

Static Variables View (Fig. 2 (A)) shows a list of selected measures that do not change over time in horizontal bar charts. Users can use this view to show the general summary of subjects in terms of their static measures, such as genetic profile, gender, race, and family history status. Using this view, users can build cohorts based on factors related to subjects' genetic characteristics and biological
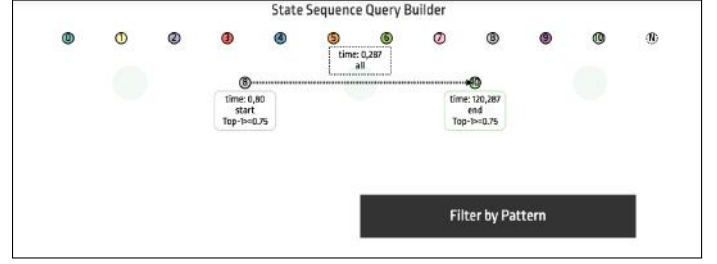

Fig. 10: State Sequence Finder allows users to build and refine state transition patterns. Users can adjust filtering criteria (e.g., time of visit) for each node and edge on a context menu.

attributes (**T4**). For example, users can click on 'Y' on the 'bfever' bar chart and 'F' on the 'SEX' bar chart to create a cohort that includes only female subjects who were breastfed.

4.5 Subgroup Builder: Build, Refine, Compare, Retain

In this view group, users build, refine, and maintain subject groups using a variety of operations.

4.5.1 Subgroup List

Subgroup List (Fig. 2 (G)) shows a list of subject cohorts that have been built and kept by the user. As mentioned in previous subsections, users can build cohorts using any of the views described above; whenever users make updates on the cohorts, the changes are automatically saved in this view (**T5**). The height of each bar corresponds the relative number of subjects within each cohort. The title can be modified with free text so that users can keep short descriptions about the cohort. As users hover over each bar, a popup view shows the description of filters, namely a list of pairs of an attribute name and its value range, used to create the corresponding cohort. Users can also export and import cohorts from existing static variables by choosing one variable on the dropdown menu at the top-right corner (**T5**).

4.5.2 State Sequence Finder

State Sequence Finder is a way users can freely define and refine queries to create a cohort based on state transition patterns (**T5**). We built this view based on users' feedback. Users want to build cohorts by setting sophisticated filters using state transition patterns. We initially considered a command-line interface, where users can 'type' to build a query using some syntax language like SQL. We later found that our users can face difficulties in learning how to use the syntax language. Thus, we decided to build a graphic user interface that allows users to interactively build state sequence query. As Fig. 10 shows, the view includes a canvas, where users can graphically create transitions of interest by using state nodes. The view initially shows an empty canvas with horizontally arranged state nodes with an additional dummy node (in total, N+1 nodes). Using a drag-and-drop interaction, users can express the state transition patterns that they are looking for. Though the node-link representation is intuitive to understand and learn how to build, it may sacrifice expressivity. Users may want to refine their search based on other attributes related to states (nodes) or transitions (edges). For example, instead of just state 4, users may want to filter state 4, which occurred in specific ages. Thus, we added a popup menu for users to express more sophisticated constraints on states and transitions. Every node and edge has its own properties shown with labels, which can also be edited on a popup window. Each node has three properties: 1) time, 2)

node_at, and 3) posteriors, and each edge between two nodes has two properties: 1) the maximum time between nodes and 2) the sequential order of the two nodes. Once users configure patterns, they hit the "Filter by Pattern" to execute the query. Fig. 10 shows that a user built a query using two nodes: state 8 (no autoantibody) and state 10 (GADA and IA2). The pattern includes the following configurations: 1) State 8 occurs between age 0 and 80 months at the beginning visit of the subject; 2) State 10 occurs between age after 120 months at the final visit; 3) both states are assigned with posterior probabilities greater than or equal to 0.75.

5 CLINICAL STUDIES OF TYPE 1 DIABETES

In this section, we introduce 1) a usage scenario that demonstrates how DPVis help users explore disease progression pathways and 2) users' experiences of using DPVis.

5.1 Long-Term Collaboration with Study Group

This section briefly explains the long-term collaboration with clinical researchers in type 1 diabetes, who contributed to the user experiences and usage scenarios. As part of the *T1DI* study group, we organized four quarterly workshops in four different cities in the US and Europe between October, 2018 and October, 2019. The goal of each workshop was to gain clinically meaningful insights into subjects' trajectories prior to diagnosis. In each workshop, clinical researchers and two leading authors of the paper used DPVis T1D to explore T1D data collected from multi-site birth cohort studies. During the workshop, we conducted Pair Analytics [2], where the two leading authors drive the tool and clinical researchers interpret the patterns, formulate hypotheses, and derive clinically meaningful insights. As a result, we derive the usage scenario that leads to key insights, which will be presented to an upcoming Diabetes conference, in Section 5.2. We recruited nine clinical researchers, who actively participated in the workshop sessions, for the interview session to ask questions about their experiences. In the open-ended discussion with each, we focus our discussion on the cost and value of DPVis for clinical research. We recorded the conversation, which took an hour on average, and coded the transcribed scripts using grounded theory [49]. We report the results in Section 5.3. Though the study is part of large study consortium, the process and outcome reflect the goal of Multi-dimensional In-depth Long-term Case study (MILC) [62].

5.2 Usage Scenario

Our users investigated disease progression patterns from 559 subjects who were diagnosed with type 1 diabetes (T1D), using observational data from three prospective clinical studies: DiPiS [29], DAISY [59], and DIPP [50]. Researchers decided to model the subjects' progression data using Continuous time HMM (CT-HMM), a variant of HMM. They trained models with a range of parameters, primarily the number of states, from 2 to 20. They focused on autoantibodies (AAB) variable due to recent interest and work in the type and number of AAB development in the T1D community. For training the autoantibody (AAB) model, they used three observation variables, all markers of disease progression: glutamate decarboxylase autoantibody (GADA), insulin autoantibody (IAA), and islet autoantigen-2 autoantibody (IA2A) (More studies on autoantibodies can be found in [68]). To test their hypothesis, they trained the models with a range of states ($N = [2, 20]$). After testing the models with multiple rounds of cross-validations, they narrowed down to the best performing model based on the log-likelihood scores on the validation dataset: an 11-state autoantibody (AAB) model.

What are the states discovered by the 11-state model (**T1**)? Users open Feature Matrix and State Summary Panel to interpret the states with respect to the evolution of AABs. They found that there are at least three low-AAB states, which are states 0, 3, and 8, highlighted in red in Fig. 11a. They are characterized as buttons in blue for the majority of rows in Feature Matrix, including the three AABs, which indicate that the three states represent young, early states prior to onsets of any AAB. The remaining states can be categorized into groups with respect to the number of AAB: (i) 1-AAB state: 2, 4, and 9; (ii) 2-AAB state: 4, 7, and 10; (iii) 3-AAB state: 1 and 6. Users found that state 4 can be categorized into 1-AAB, 2-AAB, or both because it has .56 probability in GADA; state 2 can be categorized into 2-AAB, 3-AAB, or both because it has .59 probability in IAA, as Fig. 11a highlights in blue. This indicates that transitions including state 2 or 4 can also be subdivided into more granular trajectory groups. After inspecting the state characteristics, users confirmed that the 11-state model expose diverse states represented by various combinations of AABs.

Together with the characteristics of individual states, users set out to understand disease progression patterns (**T2**) and their relationship with onsets of diagnosis (**T3**). The overview shows that there are three large distinctive trajectory groups, each of which follow state transitions: i) 0-2, ii) 3-7, and iii) 8-10, which are named as "Multiple-AABs First", "IAA First", and "GAA First" groups, respectively, as Fig. 11b shows. The names are given as above by clinicians because the very first AAB(s) subjects gained provide clinically meaningful implications. Using State Sequence Finder, users create subgroups in Subgroup List for detailed investigation (**T5**). The number of subjects for the three trajectory groups are 221, 224, and 104. There are 10 subjects who do not belong to one of three because they show transitions across multiple trajectory groups: 7 subjects belong to Multiple-AABs First and IAA First, and 3 subjects belong to IAA First and GAA First. Looking at the overview of transitions, users notice key differences between three trajectory groups with respect to transition ages. Pathway Waterfall shows that the vertical paths drawn in the "IAA First" group are very dense especially early in their ages in comparison to the other trajectory groups. To investigate the association between average age for diagnosis and seroconversion and different trajectory groups, users view Dual Kernel Densities View that shows density charts for diagnosis (red) and seroconversion (purple). Fig. 11c shows that the "IAA First" group has earlier seroconversion and diagnosis than the population.

Following the trajectories, they found that some subjects were assigned State 0 (no autoantibodies) for all of their visits, yet they were diagnosed with T1D. To investigate this cohort further, they queried those who started and ended their observations with State 0 using State Sequence Finder (**T4**), as Fig. 11d shows. The query returned 11 subjects, who were diagnosed around 119 months (10 years) old, according to Dual Kernel Densities View. Among the 11 subjects, 6 of them seroconverted with IAA, and the rest had no seroconversion, according to Static Variables View. Subject List View shows posterior probabilities over 11 states for each visit, and there are many visits that show nontrivial posterior values for states other than state 0 (**T4**). Based on the patterns, users speculated two interpretations. First, the negative autoantibody at the last visit may be measurement error. Second, the subjects may represent an endotype characterized by factors that are not captured in our data [4]. They were intrigued and requested more information.

(a) State characteristics for 11-state model are summarized.

(b) Three distinct trajectory subgroups are revealed.

(c) Average ages for diagnosis and seroconversion are shown.

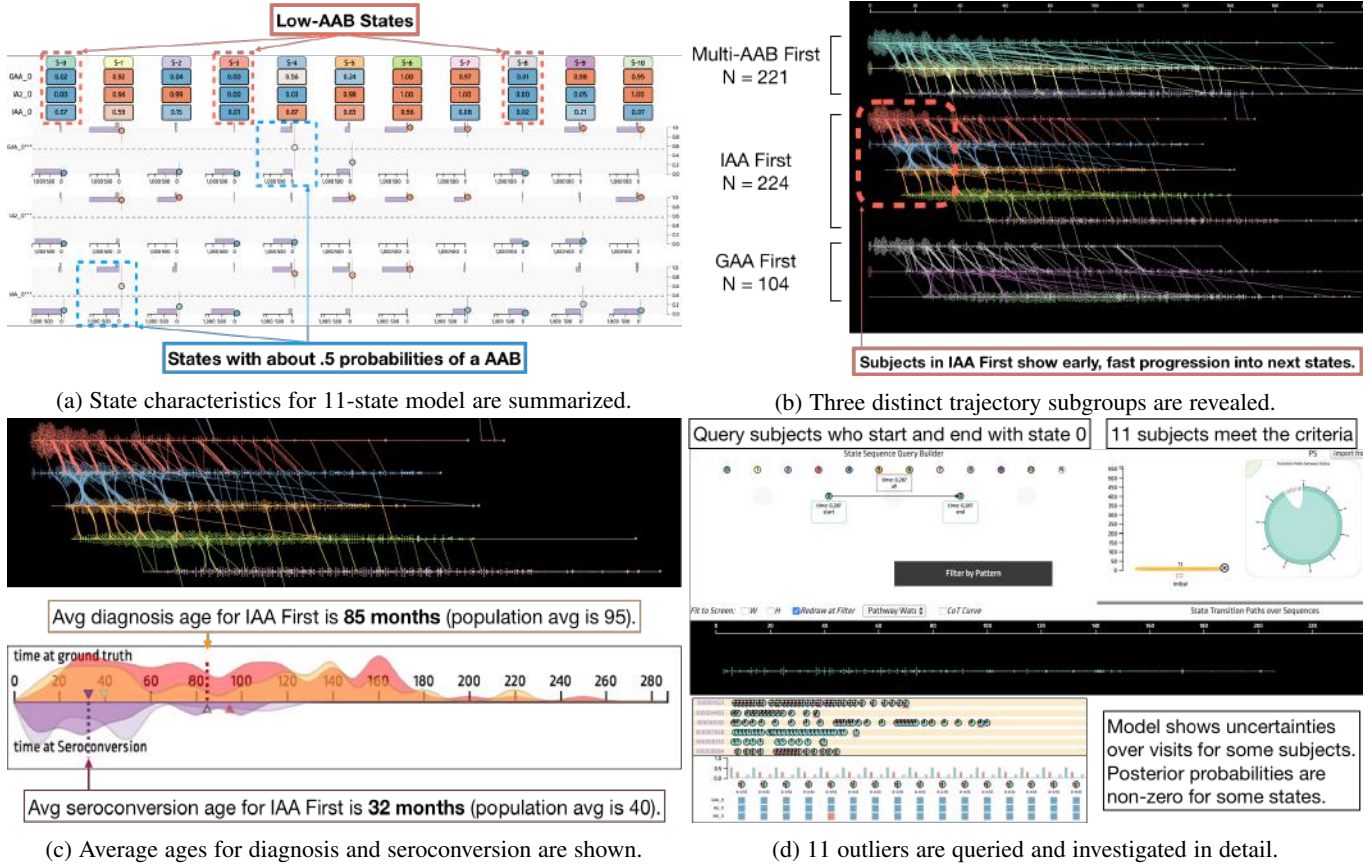(d) 11 outliers are queried and investigated in detail.

Fig. 11: The figure shows how users explore disease progression patterns using DPVis following the scenario in Section 5.2.(a) Users view and compare characteristics of states (**T1**); (b) users gain an overview of state transition patterns (**T2**); (c) users probe the relationship between health outcomes and disease progression trajectories (**T3**); (d) users check details of subjects and manage subgroups (**T4, T5**).

5.3 User Experiences on DPVis

This section reports the summary of interviews with nine clinical researchers who participated in the workshop sessions. Eight out of nine researchers had 10 or more years of experience (min: 3 years; max: 40 years) with clinical or healthcare research. Five out of nine researchers were practicing medical doctors and clinical researchers, who had Doctor of Medicine in various disciplines related to diabetes research; The rest had degrees in statistics, computer science, and industrial engineering and had experiences with clinical and medical informatics. They had some knowledge and experiences with HMMs and visualizations in their research.

DPVis reduces the effort required to analyze observational data by providing interactive subgroup analysis. Participants reported that the major benefit of DPVis is to save time and effort for analyzing observational data. P2 mentioned that understanding disease progression patterns (*T1-2*) is "definitely doable without DPVis, but not easy, very difficult. Using DPVis makes it so much easier to do so." P9 mentioned that it helps her "quickly test an idea about whether my hypotheses are worth investigating." Participants described the utilities that DPVis uniquely provides. First, participants appreciated that the tool allows them "to see the entire cohort and to single out individuals at the same time" (P4). Second, it was regarded useful to be able to "build cohorts without scripts" (P1). The cohort building activity (*T5*) allows users to validate whether "what one has published or investigated earlier makes sense in other data, graphically" (P7). Third, participants were able to conduct sensitivity analysis repeatedly by refining subgroups and extracting patterns from data (*T5*): "You can conduct sensitivity analysis again and again. You can slightly change

what you look at. That's a real strength. [...] It lets users test the assumptions on particular subgroups. Even kinda test the robustness. Once you identify a pattern, then you could ask: it this pattern robust even after excluding a particular study [dataset]?" (P3).

DPVis helps users understand HMMs transparently. Despite the perceived usefulness of HMMs, participants had concerns about how to interpret them, but DPVis was helpful to mitigate the issues and to use the model in a more transparent manner. HMMs are generally regarded as useful to characterize disease progression patterns of individual patients (*T1*). P1 described the usefulness as: "HMMs are useful to understand the severity of diseases. It quantifies disease states and lets users find patterns." However, participants were aware of limitations with regards to how to interpret the outputs. HMM outputs alone were not perceived as intuitive because "there are always uncertainties in state assignment" (P7) and "states seem like black boxes" (P8). It is also challenging to interpret model outcomes without considering and keeping in mind how the data were collected: "Censoring is a big issue [in observational studies]. Everything we look at is very biased. Hard to get clarity on what data its trained on" (P9). What DPVis provides is to translate the HMM outputs into visual patterns. P4 said "The way it sets up helps to explain the model. It mitigates underlying criticism about the black-box nature of the model." P7 also commented "You can kind of see the quality of data: sampling and visit intervals. Thats quite useful." DPVis is agnostic to model configurations (e.g., number of states) or data quality (e.g., sampling frequency) and lets users ask difficult questions to answer without visualizations: "By looking at different models and data using DPVis, we can ask questions like, is there any genetic

polymorphism why children have different trajectories?" (P4).

Multiple views of DPVis help users summarize, search patterns, and build subgroups of disease progression trajectories. Among many views, Pathway Waterfall and Feature Matrix were considered the most useful. P2 also mentioned "[Pathway Waterfall] catches my eyes immediately. It shows how people progress across different states over time." Pathway Waterfall shows the trends without hiding individuals (*T2*), as participants noted "It shows the impression of the volume of subjects. It shows how people travel across states over time in different ways. Summary (aggregated visualization) hide those" (P3) and "Instead of quantiles and distributions, you can see individuals" (P5). Participants find Feature Matrix useful to characterize the states discovered by HMMs. P1 described the column of Feature Matrix as "semantics of states", which show what each state is (*T1*). Subject List View were helpful to check the details: "easy to understand, and there you can see huge variation among subjects" (P7). Dual Kernel Densities View were useful to determine characteristics of cohorts (*T3*): "it helps you determine the samples' rapid progressors, slow progressors, or non progressors" (P4). To build and refine subgroups (*T5*), State Sequence Finder was regarded helpful: "I really love State Sequence Finder. I think its very difficult to query state sequence patterns from longitudinal data. Using State Sequence Finder, I can say anybody who had GADA and then IAA, I want to see those people."

It takes time to learn how to use DPVis. Participants reported that it takes time and effort to learn how to use DPVis. The major difficulty was understanding and learning how to use the multiple views. P1 revealed, "each view makes sense, but at first I did not know where I need to look to understand transitions and semantics of states." P6 also commented, "Haven't seen this kind of visualization before, it is not easy to catch what it means at first [...] it's not something that easily pops up, it's more like you spend time and effort." The difficulties come from not having a computational background and lack of experience with visual analytics applications. P2 mentioned, "some views can be difficult to understand, depending on his/her background." P5 said, "Somebody who doesn't know HMMs need to learn a bit about HMMs first." P9 also added, "It might be misleading for those who are not familiar with the algorithm part." However, all participants agreed that they were able to understand how to use DPVis after the series of workshops. The key is to use DPVis multiple times repeatedly and to actively formulate and validate hypotheses on real data. P7 mentioned "It takes time and effort and many trials and errors. You have to be very eager to ask questions in order to understand them." P4 commented, "The series of workshops and discussions really help me to better understand the tool. [It shows] the way one builds the analysis using the cohorts from the literature. I immediately grasp the way the data displayed."

Users want to provide feedback to the HMM learning process. Participants shared ideas for future work. Some participants wanted to add more *automated analysis*. P1 wanted DPVis to "automatically identify and highlight outliers in Pathway Waterfall." Participants wanted to go beyond interpreting outcomes of trained HMM and to steer the model. P4 wanted an ability "to modify states and to establish new states." P9 also added "Allow me to specify my own state. I would like to train a model with states I define. I would like to see how the dataset fits my mental model. It's a human-supervised model. I would like to have more control over the training part." Participants also want to add statistical tests to confirm the patterns they find in the visualization: "Statistics will confirm the trends and help us build clinical trials. Tests like Chi-Square, T-test, [...] to test difference between subgroups in terms of progression rates and speeds" (P4). Some participants suggested the potential use of DPVis for *clinician-patient interaction*. P7 wanted to use Pathway Waterfall to explain a patient's progress: "Id like to have his/her waterfall with the background of others. My study subjects are at some point of the trajectory, so Id like to use this to show where the subject is now. Its more like a clinical tool. For patient education, I think it will be useful. For clinician education, it will also be useful." P4 said "The utility of this tool could be for physicians, families, and others, who are interested in the particular person in relation to others. I think it will be a fantastic educational tool for clinicians, scientists, and families."

6 DISCUSSION

In this section, we discuss some implications of DPVis, lessons learned from our study, and provide some thoughts for future work.

6.1 HMMs and DPVis for Disease Progression Patterns

Clinical researchers want to understand why (etiology) and how (pathogenesis) a certain disease develops from longitudinal observational study data. HMMs are great ways to summarize the course of disease development. The states discovered by HMMs are explainable by associated variables (*G1*) and naturally form trajectories that share the same state transition patterns (*G2*). However, as our participants revealed in Section 5.3, HMMs alone are not interpretable nor useful for clinicians to understand and explore the disease progression patterns effectively. DPVis allows users to understand the state characteristics and disease progression patterns transparently through multiple, coordinated visualizations. Furthermore, the interactive features of DPVis, such as State Sequence Finder, allow users to flexibly define subgroups and to test their relationships with various health-related outcomes (*G3*). As Section 5.2 describes, users can gain clinically meaningful insights by driving their analysis using DPVis. The three clinical research goals (*G1-3*), the five analytic tasks (*T1-5*), and the design of DPVis can provide useful guidance for future visual analytics applications for clinical research. Our study can provide insights for other domains, where research questions involve summarizing and understanding temporal event sequence patterns.

There are recommendations and limitations when one uses HMMs and DPVis. As mentioned in Section 5, it is recommended to train an HMM with less than 20 states in order to maximize the interpretability. Users can use any number of variables for running an HMM. In some cases, users use latent variables, which are reduced from original variables using dimensionality reduction algorithms, for inferring states. This can also reduce the interpretability of the model. Pathway Waterfall can suffer from overplotting when the number of subjects and the number of visits increase. It had no issues of overplotting for one of the largest observational study data from T1DI described in Section 5.2. In case of overplotting, one can adjust the canvas size and/or hide dots in Pathway Waterfall temporarily. Feature Matrix could be overwhelming to users due to numbers written over bubbles, so it is recommended to only show the numbers when requested.

6.2 Supporting Subgroup Building for Clinical VA

Our study demonstrates the usefulness of subgroup building activities (*T5*) in clinical visual analytics. As Section 5.3 reports, users want to test hypotheses from prior literature and to conduct sensitivity analyses repeatedly by stratifying subjects into relevant

subgroups. Thus, it is very important for a visual analytics system to support creating and updating subgroups. While designing DPVis, we discovered that every view can serve as a tool to create and refine subgroups of subjects based on users' constraints. Since every view has its own perspective, users can express their constraints in different ways. By combining multiple constraints, users can refine cohorts that meet their analysis goals. For instance, a user can set a filter based on time of state transition using State Sequence Finder and another filter based on heights of subjects at a state using Feature Distribution. The flexible methods of DPVis for constructing filters was useful for our experts to formulate and test their hypotheses. The features blends well with users' analytic workflow, where they define and express cohorts and test their hypotheses by comparing the patterns of those different cohorts.

Supporting subgroup building activity can be adapted into future clinical visual analytics applications. Clinical researchers often test hypotheses by comparing different subgroups. The subgroups are frequently referred to with special names given by clinicians or previous literature, and they are the main character of stories they discuss in order to gain clinically meaningful insights. For instance, clinical researchers speculate whether the "Multiple-AAB First" group, discovered in Section 5.2, can be replicated in other observational studies. Future researchers can investigate methods to replicate subgroups from prior literature in an automated or human-in-the-loop approach, as requested by participants in Section 5.3. In future work, we aim to investigate intuitive ways to build cohorts, which include drawing expected patterns over time-series visualizations like AxiSketcher [34].

6.3 Design Study Reflection: Working in a Triad

An important key to success of a design study is to effectively facilitate communication between experts in their domains, mainly health professionals, statisticians, and visualization experts. By combining the pieces from the triad of experts, we can build an effective visual analytics application. Though it is essential for the experts to carefully evaluate each other's inputs, it is very challenging due to the knowledge gaps between domains. What facilitates effective communication between experts in different fields can be the visualization itself [35]. Throughout meetings, experts discussed over DPVis, which resulted in several changes to try out in the next phase. Therefore, what makes a visual analytics tool interpretable and engaging for users is how effectively the representations facilitate communication and new insights among the team. DPVis was used to translate a fundamental question from health professionals about HMMs, "what is a state and how does it differ from clinical stages?" As Section 5 revealed, DPVis allowed our team to understand how multiple states may fit a clinical stage and to hypothesize more granular states that can describe within and outside clinical stages. An expert also reported that Pathway Waterfall allowed him to make sense of the overall scale and dimension of patients' visits over time. In the meetings, we gained numerous insights into the data aided by the various visualizations presented in this paper. In the end, the design process takes long but mutual learning across domains as Hall et al. [24] nicely depict.

6.4 Learning Curve, Uncertainties, and Trust

We find that clinicians often encounter problems while understanding layered, unknown uncertainties in the visual analytics pipeline. First, users need to understand how multiple views are set up. Second, they need to understand what the underlying model conveys in terms of its inputs and outputs. Third, they also need to know how each original study was designed and how the subjects were sampled and measured. Depending on the background of users, they might fall into any of the three problems, lose trust in the visualizations, and/or make false interpretation of the results, as Sacha et al. provides [60]. We overcome the issues through a series of workshops. Participants reported that the key to overcoming such problems was to repeatedly formulate and validate hypotheses with small examples from real data. Participants often expect insights to pop out without any actions. It is important for them to keep in mind that the clinicians need to drive the analysis by actively formulating and validating hypotheses. In addition, designers need to become semi-experts in clinical and analytic domains by actively engaging with the end users for a long period of time. We believe that clinicians start engaging with the tool when they are convinced by useful findings, insights, and hypotheses proven through the tool. Thus, designers of the system need to spend time in understanding how the model works and interpreting the results in clinical context. It is often very difficult to acquire additional skills in machine learning and clinical domains. We believe that tight collaboration through a study group will be useful to learn the perspectives of clinicians. The design study protocol clearly specifies the role of liaison, which we need to pursue in order to properly implement the visual analytics approach for medical experts [64].

7 CONCLUSION

In this work, we conducted a design study with clinical researchers, including physician scientists, statisticians, and visualization experts, to develop a visual analytics application for exploring disease progression pathways and their interaction with various measures. As a result, we developed DPVis, which seamlessly integrates HMMs and provides views and interactive features that facilitate users to formulate and test hypotheses by iteratively constructing multiple subgroups based on state transitions and distributions over multiple measures and then analyzing and comparing them. The usage scenario and user experiences sections demonstrate the usefulness of the application to gain a summary of disease progression trajectories transparently, to construct subgroups in a flexible manner, and to characterize the patterns using relevant health outcomes. Our collaborators have been continuously researching disease progression patterns for type 1 diabetes, Huntington's disease, and Parkinson's disease using DPVis. In future work, we plan to extend DPVis so that users can visually supervise the training process of HMMs. We also plan to apply DPVis into clinical studies of other diseases (e.g., Alzheimer's Disease).

REFERENCES

[1] P. Angelelli, S. Oeltze, J. Haász, C. Turkay, E. Hodneland, A. Lundervold, A. J. Lundervold, B. Preim, and H. Hauser. Interactive visual analysis of heterogeneous cohort-study data. *IEEE computer graphics and applications*, 34(5):70–82, 2014.

[2] R. Arias-Hernandez, L. T. Kaastra, T. M. Green, and B. Fisher. Pair analytics: Capturing reasoning processes in collaborative visual analytics. In *2011 44th Hawaii international conference on system sciences*, pp. 1–10. IEEE, 2011.

[3] J. Ayres, J. Flannick, J. Gehrke, and T. Yiu. Sequential PAttern Mining Using a Bitmap Representation. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '02, pp. 429–435. ACM, New York, NY, USA, 2002.

[4] M. Battaglia, M. S. Anderson, J. H. Buckner, S. M. Geyer, P. A. Gottlieb, T. W. H. Kay, . Lernmark, S. Muller, A. Pugliese, B. O. Roep, C. J. Greenbaum, and M. Peakman. Understanding and preventing type 1 diabetes through the unique working model of TrialNet. *Diabetologia*, 60(11):2139–2147, Nov. 2017.

[5] N. Beerenwinkel, J. Rahnenfhrer, M. Dumer, D. Hoffmann, R. Kaiser, J. Selbig, T. Lengauer, D. Gusfield, P. Bourne, S. Istrail, P. Pevzner, and M. Waterman. Learning multiple evolutionary pathways from cross-sectional data. *RECOMB 2004 : proceedings of the Eighth Annual International Conference on Research in Computational Molecular Biology, ACM*, pp. 36–44, 03 2004.

[6] J. Bernard, D. Sessler, J. Kohlhammer, and R. A. Ruddle. Using Dashboard Networks to Visualize Multiple Patient Histories: A Design Study on Post-Operative Prostate Cancer. *IEEE Transactions on Visualization and Computer Graphics*, 25(3):1615–1628, Mar. 2019.

[7] B. C. Cappers and J. J. van Wijk. Exploring Multivariate Event Sequences Using Rules, Aggregations, and Selections. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):532–541, Jan. 2018.

[8] Y. Chen, A. Puri, L. Yuan, and H. Qu. StageMap: Extracting and Summarizing Progression Stages in Event Sequences. In *2018 IEEE International Conference on Big Data (Big Data)*, pp. 975–981, Dec. 2018.

[9] Y. Chen, P. Xu, and L. Ren. Sequence Synopsis: Optimize Visual Summary of Temporal Event Data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):45–55, Jan. 2018.

[10] M. Cho, B. Kim, H. Bae, and J. Seo. Stroscope: Multi-Scale Visualization of Irregularly Measured Time-Series Data. *IEEE Transactions on Visualization and Computer Graphics*, 20(5):808–821, May 2014.

[11] S. F. Cook and R. R. Bies. Disease progression modeling: Key concepts and recent developments. *Current Pharmacology Reports*, 2(5):221–230, 2016. doi: 10.1007/s40495-016-0066-x

[12] R. DESPER, F. JIANG, O. Kallioniemi, H. MOCH, C. H. PAPADIM-ITRIOU, and A. Schaffer. Inferring tree models for oncogenesis from comparative genome hybridization data. *Journal of computational biology : a journal of computational molecular cell biology*, 6:37–51, 02 1999. doi: 10.1089/cmb.1999.6.37

[13] D. Dingen, M. van't Veer, P. Houthuizen, E. H. J. Mestrom, E. H. Korsten, A. R. Bouwman, and J. van Wijk. RegressionExplorer: Interactive Exploration of Logistic Regression Models with Subgroup Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):246–255, 2019.

[14] A. Eklund. beeswarm: The Bee Swarm Plot, an Alternative to Stripchart version 0.2.3 from CRAN, 2017.

[15] M. Glueck, A. Gvozdik, F. Chevalier, A. Khan, M. Brudno, and D. Wigdor. PhenoStacks: Cross-Sectional Cohort Phenotype Comparison Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):191–200, Jan. 2017.

[16] M. Glueck, P. Hamilton, F. Chevalier, S. Breslav, A. Khan, D. Wigdor, and M. Brudno. PhenoBlocks: Phenotype Comparison Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):101–110, 2016.

[17] M. Glueck, M. P. Naeini, F. Doshi-Velez, F. Chevalier, A. Khan, D. Wigdor, and M. Brudno. PhenoLines: Phenotype Comparison Visualizations for Disease Subtyping via Topic Models. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):371–381, 2018.

[18] D. Gotz and H. Stavropoulos. DecisionFlow: Visual Analytics for High-Dimensional Temporal Event Sequence Data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1783–1792, Dec. 2014.

[19] D. Gotz, J. Zhang, W. Wang, J. Shrestha, and D. Borland. Visual analysis of high-dimensional event sequence data via dynamic hierarchical aggregation. *IEEE transactions on visualization and computer graphics*, 26(1):440–450, 2019.

[20] S. Guo, F. Du, S. Malik, E. Koh, S. Kim, L. Liu, D. Kim, H. Zha, and N. Cao. Visualizing Uncertainty and Alternatives in Event Sequence Predictions. In *ACM SIGCHI Conference on Human Factors in Computing System*, May 2019.

[21] S. Guo, Z. Jin, D. Gotz, F. Du, H. Zha, and N. Cao. Visual Progression Analysis of Event Sequence Data. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):417–426, Jan. 2019.

[22] S. Guo, K. Xu, R. Zhao, D. Gotz, H. Zha, and N. Cao. EventThread: Visual Summarization and Stage Analysis of Event Sequence Data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):56–65, Jan. 2018.

[23] C. H. Jackson, L. D. Sharples, S. G. Thompson, S. Duffy, and E. Couto. Multistate markov models for disease progression with classication error. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52:193 – 209, 07 2003. doi: 10.1111/1467-9884.00351

[24] K. W. Hall, A. J. Bradley, U. Hinrichs, S. Huron, J. Wood, C. Collins, and S. Carpendale. Design by Immersion: A Transdisciplinary Approach to Problem-Driven Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2019.

[25] D. Holten and J. J. v. Wijk. Force-Directed Edge Bundling for Graph Visualization. *Computer Graphics Forum*, 2009. doi: 10.1111/j.1467-8659 .2009.01450.x

[26] J. Hu. AI Offers Hope for Earlier Screening for Type 1 Diabetes, June 2019.

[27] H. Hubert, M. Feinleib, P. McNamara, and W. Castelli. Obesity as an independent risk factor for cardiovascular disease: A 26-year follow-up of participants in the framingham heart study. *Circulation*, 67:968–977, 1983.

[28] R. A. Insel, J. L. Dunne, M. A. Atkinson, J. L. Chiang, D. Dabelea, P. A. Gottlieb, C. J. Greenbaum, K. C. Herold, J. P. Krischer, . Lernmark, R. E. Ratner, M. J. Rewers, D. A. Schatz, J. S. Skyler, J. M. Sosenko, and A.-G. Ziegler. Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association. *Diabetes Care*, 38(10):1964–1974, Oct. 2015. doi: 10.2337/dc15-1419

[29] B. Jonsdottir, C. Larsson, M. Lundgren, A. Ramelius, I. Jnsson, H. E. Larsson, and DiPiS study Group. Childhood thyroid autoimmunity and relation to islet autoantibodies in children at risk for type 1 diabetes in the diabetes prediction in skne (DiPiS) study. *Autoimmunity*, 51(5):228–237, Aug. 2018.

[30] P. Klemm, S. Oeltze-Jafra, K. Lawonn, K. Hegenscheid, H. Völzke, and B. Preim. Interactive visual analysis of image-centric cohort study data. *IEEE transactions on visualization and computer graphics*, 20(12):1673–1682, 2014.

[31] J. Krause, A. Perer, and H. Stavropoulos. Supporting Iterative Cohort Construction with Visual Temporal Queries. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):91–100, Jan. 2016.

[32] B. C. Kwon, M. Choi, J. T. Kim, E. Choi, Y. B. Kim, S. Kwon, J. Sun, and J. Choo. RetainVis: Visual Analytics with Interpretable and Interactive Recurrent Neural Networks on Electronic Medical Records. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):299–309, Jan. 2019.

[33] B. C. Kwon, B. Eysenbach, J. Verma, K. Ng, C. D. Filippi, W. F. Stewart, and A. Perer. Clustervision: Visual supervision of unsupervised clustering. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):142–151, Jan 2018.

[34] B. C. Kwon, H. Kim, E. Wall, J. Choo, H. Park, and A. Endert. Axisketcher: Interactive nonlinear axis mapping of visualizations through user drawings. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):221–230, Jan 2017.

[35] B. C. Kwon, S.-H. Kim, S. Lee, J. Choo, J. Huh, and J. S. Yi. Visohc: Designing visual analytics for online health communities. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):71–80, 2016.

[36] B. C. Kwon, J. Verma, and A. Perer. Peekquence: Visual analytics for event sequence data. *ACM SIGKDD Workshop on Interactive Data Exploration and Analytics*, 2016.

[37] P. Law, Z. Liu, S. Malik, and R. C. Basole. MAQUI: Interweaving Queries and Pattern Mining for Recursive Event Sequence Exploration. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):396–406, Jan. 2019.

[38] D. Levy, R. J. Garrison, D. Savage, W. Kannel, and W. Castelli. Prognostic implications of echocardiographically determined left ventricular mass in the framingham heart study. *New England Journal of Medicine*, 322:1561–1566, 1990.

[39] Y.-Y. Liu, S. Li, F. Li, L. Song, and J. M. Rehg. Efficient learning of continuous-time hidden markov models for disease progression. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, eds.,

Advances in Neural Information Processing Systems 28, pp. 3600–3608. Curran Associates, Inc., 2015.

[40] Z. Liu, B. Kerr, M. Dontcheva, J. Grover, M. Hoffman, and A. Wilson. CoreFlow: Extracting and Visualizing Branching Patterns from Event Sequences. *Computer Graphics Forum*, 36(3):527–538, 2017.

[41] Z. Liu, Y. Wang, M. Dontcheva, M. Hoffman, S. Walker, and A. Wilson. Patterns and Sequences: Interactive Exploration of Clickstreams to Understand Common Visitor Paths. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):321–330, Jan. 2017.

[42] M. H. Loorak, C. Perin, N. Kamal, M. Hill, and S. Carpendale. TimeSpan: Using Visualization to Explore Temporal Multi-dimensional Data of Stroke Patients. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):409–418, Jan. 2016.

[43] M. Lorezi, F. M., G. Frisoni, D. Alexander, S. Ourselin, and A. D. N. Initiative. Proabilistic disease progression modeling to characterize diagnostic uncertainty: Application to staging and prediction in alzheimer's disease. *NeuroImage*, in press, 2017.

[44] G. E. Marai, C. Ma, A. T. Burks, F. Pellolio, G. Canahuate, D. M. Vock, A. S. R. Mohamed, and C. D. Fuller. Precision Risk Analysis of Cancer Therapy with Interactive Nomograms and Survival Plots. *IEEE Transactions on Visualization and Computer Graphics*, 25(4):1732–1745, 2019.

[45] K. Marek *et al.* The parkinson progression marker iniative (ppmi). *Progress in Neurobiology*, 95:629–635, 2011.

[46] A. Mathisen and K. Grnbk. Clear Visual Separation of Temporal Event Sequences. In *2017 IEEE Visualization in Data Science (VDS)*, pp. 7–14, Oct. 2017.

[47] M. Monroe, R. Lan, H. Lee, C. Plaisant, and B. Shneiderman. Temporal Event Sequence Simplification. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2227–2236, Dec. 2013.

[48] D. Mould. Models for disease progression: new approaches and uses. *Clinical Pharmacology & Therapeutics*, 92(1):125–131, 2012.

[49] M. Muller. Curiosity, Creativity, and Surprise as Analytic Tools: Grounded Theory Method. In J. S. Olson and W. A. Kellogg, eds., *Ways of Knowing in HCI*, pp. 25–48. Springer, New York, NY, 2014.

[50] S. Nejentsev, M. Sjroos, T. Soukka, M. Knip, O. Simell, T. Lvgren, and J. Ilonen. Population-based genetic screening for the estimation of Type 1 diabetes mellitus risk in Finland: selective genotyping of markers in the HLA-DQB1, HLA-DQA1 and HLA-DRB1 loci. *Diabetic Medicine: A Journal of the British Diabetic Association*, 16(12):985–992, Dec. 1999.

[51] P. H. Nguyen, R. Henkin, S. Chen, N. Andrienko, G. Andrienko, O. Thonnard, and C. Turkay. Vasabi: Hierarchical user profiles for interactive visual user behaviour analytics. *IEEE transactions on visualization and computer graphics*, 26(1):77–86, 2019.

[52] J. Pei, J. Han, M. Asl, H. Pinto, Q. Chen, U. Dayal, and M. Hsu. PrefixSpan: Mining Sequential Patterns Efficiently by Prefix Projected Pattern Growth. In *Proceedings 17th International Conference on Data Engineering*, pp. 215–226, 2001.

[53] A. Perer and F. Wang. Frequence: Interactive Mining and Visualization of Temporal Frequent Event Sequences. In *Proceedings of the 19th International Conference on Intelligent User Interfaces*, IUI '14, pp. 153–162. ACM, New York, NY, USA, 2014.

[54] A. Perer, F. Wang, and J. Hu. Mining and exploring care pathways from electronic medical records with visual analytics. *Journal of Biomedical Informatics*, 56:369–378, Aug. 2015.

[55] C. Plaisant, R. Mushlin, A. Snyder, J. Li, D. Heller, and B. Shneiderman. LifeLines: using visualization to enhance navigation and analysis of patient records. *Proceedings of the AMIA Symposium*, pp. 76–80, 1998.

[56] P. J. Polack, and, and and. TimeStitch: Interactive multi-focus cohort discovery and comparison. In *2015 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 209–210, Oct. 2015.

[57] P. J. Polack Jr., S.-T. Chen, M. Kahng, K. D. Barbaro, R. Basole, M. Sharmin, and D. H. Chau. Chronodes: Interactive Multifocus Exploration of Event Sequences. *ACM Transactions on Interactive Intelligent System*, 8(1):2:1–2:21, Feb. 2018.

[58] V. Raveneau, J. Blanchard, and Y. Pri. Progressive sequential pattern mining: steerable visual exploration of patterns with PPMT. In *Visualization in Data Science (VDS at IEEE VIS 2018)*. Berlin, Germany, Oct. 2018.

[59] M. Rewers, T. L. Bugawan, J. M. Norris, A. Blair, B. Beaty, M. Hoffman, R. S. McDuffie, R. F. Hamman, G. Klingensmith, G. S. Eisenbarth, and H. A. Erlich. Newborn screening for HLA markers associated with IDDM: diabetes autoimmunity study in the young (DAISY). *Diabetologia*, 39(7):807–812, July 1996.

[60] D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. A. Keim. The Role Of Uncertainty, Awareness, And Trust In Visual Analytics. *Visualization and Computer Graphics, IEEE Transactions on*, 22(1):240–249, 2016.

[61] P. Schulam and S. Saria. A Framework for Individualizing Predictions of Disease Trajectories by Exploiting Multi-Resolution Structure. In *NIPS*, 2015.

[62] B. Shneiderman and C. Plaisant. Strategies for Evaluating Information Visualization Tools: Multi-dimensional In-depth Long-term Case Studies. In *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization*, pp. 1–7. ACM, 2006.

[63] B. Shneiderman and C. Plaisant. Interactive Visual Event Analytics: Opportunities and Challenges. *Computer*, 52(1):27–35, Jan. 2019.

[64] S. Simon, S. Mittelstädt, D. A. Keim, and M. Sedlmair. Bridging the gap of domain and visualization experts with a liaison. In *Eurographics Conference on Visualization 2015*, pp. 127–131, 2015.

[65] C. D. Stolper, A. Perer, and D. Gotz. Progressive Visual Analytics: User-Driven Visual Exploration of In-Progress Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1653–1662, Dec. 2014.

[66] R. Sukkar, E. Katz, Y. Zhang, D. Raunig, and B. T. Wyman. Disease progression modeling using hidden markov models. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2845–2848, 2012. doi: 10.1109/EMBC.2012.6346556

[67] Z. Sun, Z. Ghosh, Y. Li, Y. Cheng, A. Mohan, C. Sampaio, and J. Hu. A probabilistic disease progression modeling approach and its application to integrated huntington's disease observational data. *JAMIA OPEN*, 2019.

[68] C. Trn, M. LandinOlsson, J. stman, B. Scherstn, H. Arnqvist, G. Blohm, E. Bjrk, J. Bolinder, J. Eriksson, B. Littorin, L. Nystrm, G. Sundkvist, and . Lernmark. Glutamic acid decarboxylase antibodies (GADA) is the most important factor for prediction of insulin therapy within 3 years in young adult diabetic patients not classified as Type 1 diabetes on clinical grounds. *Diabetes/Metabolism Research and Reviews*, 16(6):442–447, 2000.

[69] B. Vogelstein, E. R. Fearon, S. R. Hamilton, S. E. Kern, A. C. Preisinger, M. Leppert, A. M. Smits, and J. L. Bos. Genetic alterations during colorectal-tumor development. *New England Journal of Medicine*, 319(9):525–532, 1988. doi: 10.1056/NEJM198809013190901

[70] J. Wang and J. Han. BIDE: efficient mining of frequent closed sequences. In *Proceedings. 20th International Conference on Data Engineering*, pp. 79–90, Apr. 2004.

[71] T. D. Wang, C. Plaisant, B. Shneiderman, N. Spring, D. Roseman, G. Marchand, V. Mukherjee, and M. Smith. Temporal Summaries: Supporting Temporal Categorical Searching, Aggregation and Comparison. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1049–1056, Nov. 2009.

[72] X. Wang, D. Sontag, and F. Wang. Unsupervised learning of disease progression models. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, pp. 85–94. ACM, New York, NY, USA, 2014. doi: 10.1145/2623330.2623754

[73] E. Wion, M. Brantley, J. Stevens, S. Gallinger, H. Peng, M. Glass, and W. Hagopian. Population-wide infant screening for HLA-based type 1 diabetes risk via dried blood spots from the public health infrastructure. *Annals of the New York Academy of Sciences*, 1005:400–403, Nov. 2003.

[74] K. Wongsuphasawat and D. Gotz. Exploring Flow, Factors, and Outcomes of Temporal Event Sequences with the Outflow Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2659–2668, Dec. 2012.

[75] K. Wongsuphasawat, J. A. Guerra Gmez, C. Plaisant, T. D. Wang, M. Taieb-Maimon, and B. Shneiderman. LifeFlow: Visualizing an Overview of Event Sequences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 1747–1756. ACM, New York, NY, USA, 2011.

[76] E. Zgraggen, S. M. Drucker, D. Fisher, and R. DeLine. (s, Qu)Eries: Visual Regular Expressions for Querying and Exploring Event Sequences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pp. 2683–2692. ACM, New York, NY, USA, 2015.

[77] Y. Zhang, K. Chanana, and C. Dunne. IDMVis: Temporal Event Sequence Visualization for Type 1 Diabetes Treatment Decision Support. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):512–522, Jan. 2019.

[78] A.-G. Ziegler, M. Pflueger, C. Winkler, P. Achenbach, B. Akolkar, J. P. Krischer, and E. Bonifacio. Accelerated progression from islet autoimmunity to diabetes is causing the escalating incidence of type 1 diabetes in young children. *Journal of Autoimmunity*, 37(1):3–7, Aug. 2011. doi: 10.1016/j.jaut.2011.02.004

Bum Chul Kwon is Research Staff Member at IBM Research. His research area includes visual analytics, data visualization, human-computer interaction, healthcare, and machine learning. Prior to joining IBM Research, he worked as postdoctoral researcher at University of Konstanz, Germany. He received his M.S. and Ph.D. from Purdue University in 2010 and 2013, respectively. He received his B.S. in Systems Engineering from University of Virginia in 2008.

Vibha Anand is a Research Staff Member at IBM Research. Her research focuses on applying machine learning techniques to prior observational studies, clinical trials, EMR data and real-time biosensors to derive actionable insights. Before IBM Research, she was an Assistant Professor and Professional Staff at the Cleveland Clinic and at Indiana University, School of Medicine in Dept. of Pediatrics. She holds a BE in Electrical Engineering, MS in Computer Science, PhD in Informatics and is a Fellow of AMIA.

Kristen A. Severson is a post-doctoral researcher at IBM Research. Her research focuses on the development and application of machine learning techniques to derive insight from real-world health data. Prior to joining IBM, she received a PhD from the Massachusetts Institute of Technology. She also holds a BS in chemical engineering from Carnegie Mellon University.

Soumya Ghosh is a Research Staff Member at IBM Research. His research focuses on the design of flexible statistical models and efficient inference algorithms for reasoning about noisy, high-dimensional data. His recent work has examined approaches for combining the complementary strengths of Bayesian methods, graphical models, and deep neural networks for modeling progression of neuro-degenerative diseases. He holds a PhD in computer science from Brown University.

Zhaonan Sun is a Research Staff Member at IBM Research. Her research focuses on developing statistical and machine learning methods in the healthcare domain for generating insights from real-world data. Prior to joining IBM, she received PhD in statistics from Purdue University.

Markus Lundgren is a M.D, Ph.D. specializing in pediatric endocrinology and a researcher at the Department of clinical sciences Malmö at Lund University, Lund, Sweden. His research focuses on pediatric type 1 diabetes Prediction and Prevention including the effects of early disease detection. He received his Ph.D. and M.D. from Lund University in 2017 and 2003, respectively.

Brigitte I. Frohnert is a M.D., Ph.D. who specializes in a pediatric endocrinology in University of Colorado Denver, Aurora, Colorado. Her research interest is in understanding the environmental factors which contribute to the development of islet autoimmunity and progression to type 1 diabetes, with the ultimate goal of preventing type 1 diabetes in the future. She received her M.D and Ph.D from University of Minnesota in 2002 and 2000, respectively.

Kenney Ng is a Principal Research Staff Member and manager of the Health Analytics Research Group at IBM Research. His research focus is on the development and application of AI techniques to analyze, model and derive actionable insights from real world health data. He received B.S., M.S., and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology. He is a member of IEEE and AMIA.