PhUSE US Connect 2018

Paper DV01

Clinical Timelines Visualized

Spencer Childress, Rho, Inc., Chapel Hill, NC, USA
Jeremy Wildfire, Rho Inc., Chapel Hill, NC, USA
Ryan Bailey, Rho Inc., Chapel Hill, NC, USA
Britt Sikora, Rho Inc., Chapel Hill, NC, USA

ABSTRACT

The monitoring and analysis of longitudinal event data is a key aspect in all phases of clinical research. These data affect high-level decisions, such as those guiding adaptive study designs, as well as lower-level decisions like the continuation of a clinical trial participant. The Clinical Timelines library is a new open source tool designed to create interactive timeline visualizations from clinical trial data. The use of standard web technology to produce these graphics increases their accessibility and enables users to explore their data in ways that are not possible with traditional static data visualizations.

INTRODUCTION

Timelines are a fundamental tool for effective clinical project management, finding use at all levels of study execution from participant-level tracking, to site-level monitoring, to study-wide oversight. Unfortunately, the process of creating timelines is a tedious one that sees many timeline users resort to highly manual tools that lack rigorous data integration, like spreadsheets and project management systems. Even among tools that offer a more robust, programmatic approach like SAS and R, the output is typically static with limited ability for end users to explore the data in real time.

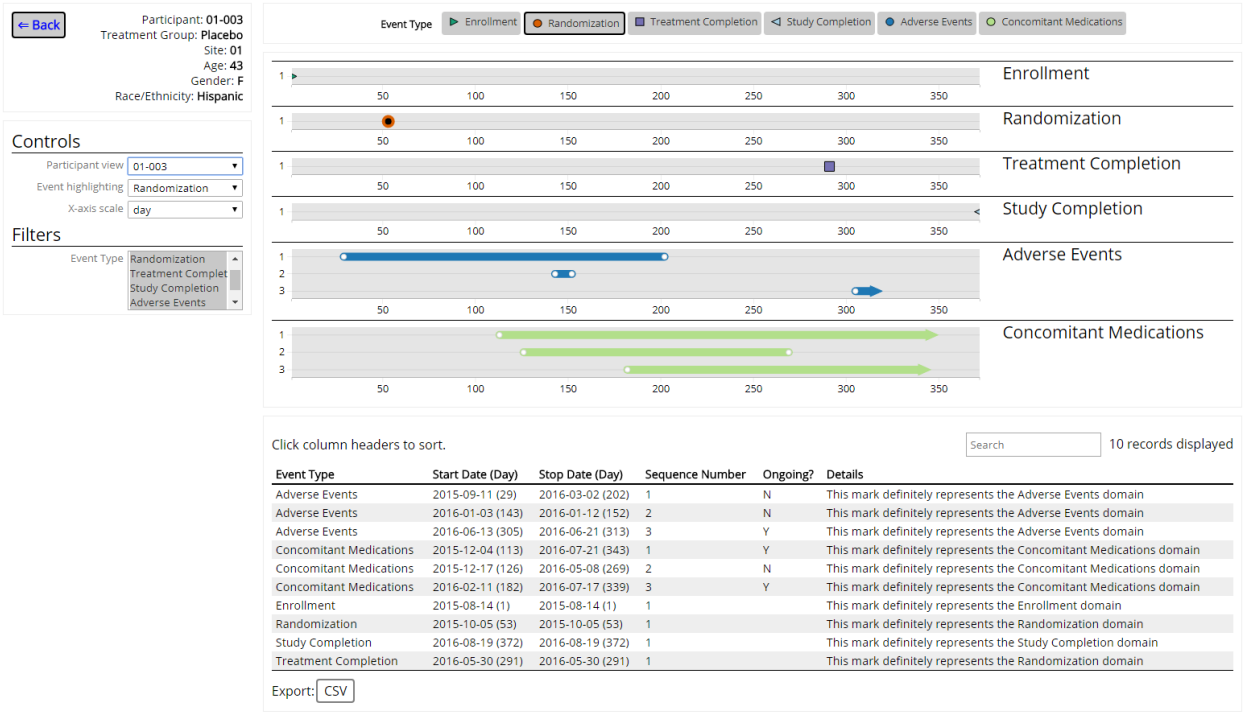
The Clinical Timelines library seeks to compactly and comprehensively represent longitudinal data from multiple data streams in a single, interactive interface (Figure 1). It renders a series of longitudinal charts faceted vertically by an identifier variable and displays events linearly from left to right. A single time scale represents multiple event types (e.g. study randomization, adverse events, etc.) and gives them context relative to each other, allowing for simple visual comparison. Patterns within an individual and across an entire population emerge in this macro view.

Figure 1 – Clinical Timelines

Data marks such as circles and triangles represent single day events or timepoints while lines represent multiday events or time intervals. The library supports both date and reference day time scales to provide both a holistic and a within-individual chronological context. For example, to give adverse events and concomitant medications context relative to study participation, data analysts measure the difference in days between the occurrence of the event and a baseline timepoint such as randomization.

This comprehensive view of all individuals in a population lets users visually tease out patterns and easily detect unexpected phenomena. It mitigates the comparison of expected events that should occur across all individuals as well as the identification of events that occur within a subset of the population. From here users can drilldown to a view of the individual (Figure 2). A profile of sorts, the individual timeline combines individual characteristics, a timeline faceted by event type, and a listing of the data associated with those events.

Figure 2 – Individual Timeline



The Clinical Timelines library is fully customizable and can be adapted to many common clinical trial data types. Possible configurations include:

- **Study population timelines,** faceted by participant, which include enrollment duration and visit patterns. These details facilitate granular risk-based monitoring tasks (assessment of treatment compliance, identification of safety signals, adherence to the schedule of events, etc.).
- **Study operational timelines,** grouped by clinical site, which allow principal investigators and study coordinators to track site-level events such as those leading to site activation and those related to study conduct (participant enrollment, early termination, protocol deviations, etc.).
- **Program-level timelines,** stratified by study, which allow sponsors to track study milestones, FDA submissions, etc. across an entire program.

Finally, because the library is open-source, users are free to tailor the tool as needed to their data and purpose. The code, technical documentation, and an interactive example built with simulated study data are available on GitHub.

INTERACTIVITY

The Clinical Timelines differs most widely from traditional static timelines because it introduces interactivity to the user. A user interface along the left side of the timelines enables broad control over the data, the time range, and the display of the timelines (Figure 4). Explanatory labels accompany the controls, which fall into a few categories: data filters that subset the underlying data and redraw the timelines; time controls that set the time scale and alter the time range; and y-axis controls that change the sorting and grouping of individuals. The *Participant view* dropdown contains a list of individuals and toggles between the population view and the individual profile view. And finally the *Event highlighting* dropdown contains a list of event types which when selected highlights that event type, giving it greater visual attraction.

Figure 3 – Interactive Legend

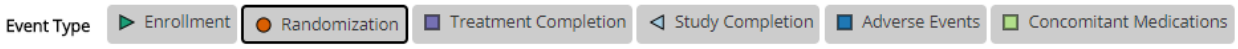
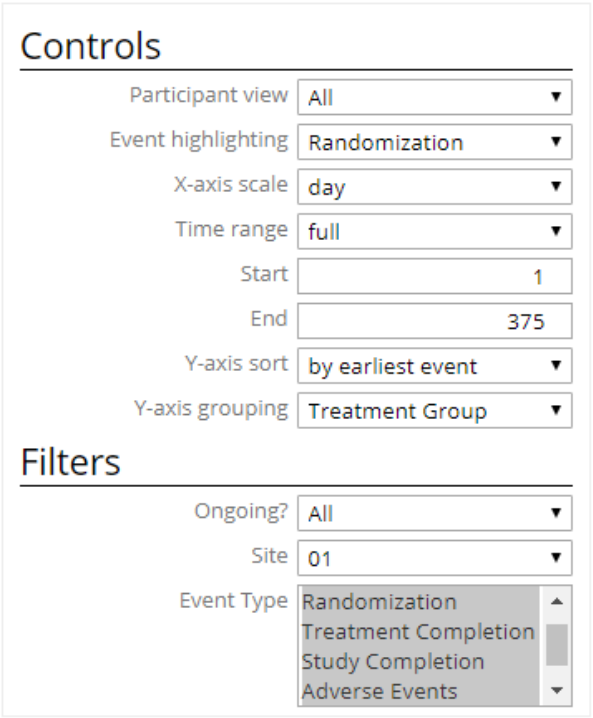


Figure 4 – Controls

The legend at the top controls the visible set of event types (Figure 3) Each legend item is a button that toggles the display of that event. Additionally, the outlined legend item indicates which event type is currently highlighted. Each item in the legend corresponds to an option in the Event Type multi-select control in the controls interface (Figure 4).



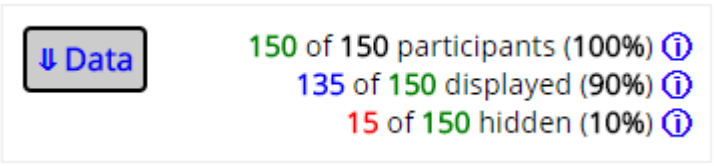
One of the drawbacks of data visualizations and data summaries in general is the disconnect with the underlying data. Reproducibility is a feature that every data analysis should include to rectify that disconnect. Standalone graphics however are difficult to reproduce and validate because the raw data rarely accompany them.

With graphics generated in a webpage, the browser stores the data and passing them to the user is trivial. In the timelines view the Clinical Timelines exposes the data via a download button in the population details container (Figure 5). Thus the user can answer his or her own questions about the chart by exporting and inspecting the data.

WEB TECHNOLOGY

Web technology makes interactivity possible and accessible both to the developer and to the user. A webpage may look like a singular document but every component of it is an element or group of elements. Each element has display properties, child and/or parent elements, and most importantly, event listeners. Event listeners wait for something to happen to the associated element, such as a mouse click or keyboard input, and trigger another event. This functionality makes web-based data visualizations active documents in contrast with passive graphics stored in image files.

Figure 5 – Data Export



The Clinical Timelines relies on two libraries to incorporate interactivity: D3 and Webcharts. D3 is a low-level data visualization library written in JavaScript that leverages HTML, SVG, and CSS to render graphics. Its central functionality relies on the binding of data to elements on the webpage, which defines a direct relationship between the data and the chart, including the dimensions of the plotting area and the coordinates of the data marks. D3 is incredibly powerful but to produce even a simple chart requires a significant amount of code because the user must define every aspect of the chart, from the canvas to the axes to the marks.

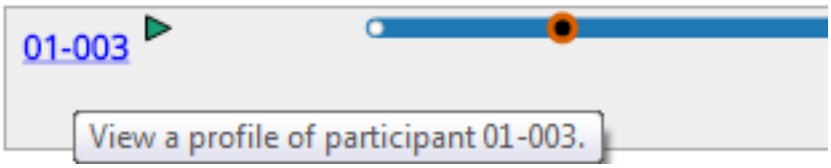
Built with D3, Webcharts simplifies the chart-making process by internalizing the steps commonly taken with D3 to generate a chart. Instead, it accepts settings objects that describe the axes, the types of marks to plot, and the color scheme. The Webcharts application programming interface (API) also exposes methods to generate controls, HTML tables, and small multiples, a series of identical charts faceted by a categorical variable in the data.

For instance, the settings for the timelines chart define a time or linear x-axis (date or reference day) and an ordinal y-axis (participant, site, study, etc.). Data marks include circles for timepoints and lines for time intervals. Finally the entire set of event types maps to a set of colors for chromatic differentiation.

Webcharts also couples data displays with controls such as data filters and settings toggles. This relationship between controls and data display provides the backbone of its interactivity. Every data mark in the timelines represents and links to one or more data points, and every data point is accessible by the filters.

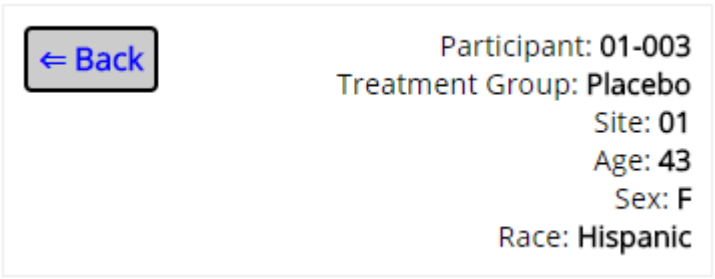
DUAL VIEW

Figure 6 – Navigation to Profile View



The Clinical Timelines opens to a macro view of all individuals, each contained within a horizontal band. The horizontal bands visually contain each individual's events, which are laid out from left to right chronologically. When events share a time point or time interval, the later event is offset vertically to avoid overlap. From this view, users can navigate to a summary of the individual either via the *Participant view* dropdown (Figure 3) or by clicking a y-axis label (Figure 6).

Figure 7 – Individual Characteristics



This view includes a breakdown of the individual's characteristics in the top left that provides relevant demographic and clinical context to the events experienced by the individual (Figure 7). Faceted by event type, the timelines plot each event on its own horizontal band (Figure 2). Finally, below the timeline the associated data lists out in a sortable, searchable, exportable, paginated data listing (Figure 8).

Figure 8 – Individual Data Listing

Click column headers to sort. Search 10 records displayed

Event Type	Start Date (Day)	Stop Date (Day)	Sequence Number	Ongoing?	Details
Adverse Events	2015-09-11 (29)	2016-03-02 (202)	1	N	This mark definitely represents the Adverse Events domain
Adverse Events	2016-01-03 (143)	2016-01-12 (152)	2	N	This mark definitely represents the Adverse Events domain
Adverse Events	2016-06-13 (305)	2016-06-21 (313)	3	Y	This mark definitely represents the Adverse Events domain
Concomitant Medications	2015-12-04 (113)	2016-07-21 (343)	1	Y	This mark definitely represents the Concomitant Medications domain
Concomitant Medications	2015-12-17 (126)	2016-05-08 (269)	2	N	This mark definitely represents the Concomitant Medications domain
Concomitant Medications	2016-02-11 (182)	2016-07-17 (339)	3	Y	This mark definitely represents the Concomitant Medications domain
Enrollment	2015-08-14 (1)	2015-08-14 (1)	1		This mark definitely represents the Enrollment domain
Treatment Completion	2016-06-01 (293)	2016-06-01 (293)	1		This mark definitely represents the Randomization domain
Study Completion	2016-08-19 (372)	2016-08-19 (372)	1		This mark definitely represents the Study Completion domain
Randomization	2015-10-05 (53)	2015-10-05 (53)	1		This mark definitely represents the Randomization domain

Export: CSV

In order to return to the population timelines, simply click the *Back* button in the individual characteristics container (Figure 7) or select *All* in the *Participant view* dropdown (Figure 4).

DATA STRUCTURE AND DATA STREAMS

The data structure the Clinical Timelines expects is straightforward and closely adheres to the Analysis Data Model (ADaM) Data Structure for Adverse Event Analysis defined by the Clinical Data Interchange Standards Consortium (CDISC). The data should contain one record per event where a variable such as DOMAIN in the Study Data Tabulation Model (SDTM) data standard distinguishes different event types (Figure 9). Each record should also include variables that define a timepoint or time interval, and can represent either a date or reference day scale. A sequence variable differentiates events from the same domain experienced by the same individual.

Figure 9 – Data Structure

USUBJID	DOMAIN	STDT	STDY	ENDT	ENDY	SEQ
01-001	Adverse Events	7/16/2015	57	8/14/2015	86	1
01-001	Concomitant Medications	7/16/2015	57	8/14/2015	86	1
01-001	Enrollment	5/21/2015	1	5/21/2015	1	1
01-001	Randomization	7/17/2015	58	7/17/2015	58	1
01-001	Study Completion	5/9/2016	355	5/9/2016	355	1
01-002	Adverse Events	7/2/2015	34	7/16/2015	48	1
01-002	Adverse Events	11/18/2015	173	5/6/2016	343	2

To create a singular, comprehensive view of a clinical trial operations area, the data facilitator must combine a variety of data sources and locations. With clinical

data, this task is more straightforward because the clinical data streams typically live in one place, such as in an electronic data capture (EDC) database or on a network drive. Furthermore, analysis datasets often already incorporate subsets of those data streams and hopefully adhere to a data standard such as SDTM or ADaM; this simplifies the identification of needed observations and variables. The combination of central location and data standards eases the creation of a clinical trial timelines dataset.

With site and study-level data, however, the data facilitator must be more resourceful. Non-EDC data streams rarely subscribe to a data standard, generally live in different locations or databases, and are not of the same format. To create a singular view of these data requires in-depth exploration of the source data and conversion to a standard format. To minimize difficulties with data configuration, the format the Clinical Timelines library requires is simple and well-defined. The data structure is one record per ID per event and requires a short list of variables: ID, event type, start date/day, and end date/day.

Site-level data of interest may include: site activation timelines, clinical monitoring visits, participant screening and enrollment, and protocol deviations. Study-level data of interest may include: protocol development, site activations, first participant first visit (FPFV), last participant last visit (LPLV), delivery timelines, and FDA submission. In the past and into the present, these data streams are stored in disparate locations and in a variety of formats such as text files, spreadsheets, and databases. With the proliferation of clinical trial management systems (CTMS) that can integrate data from multiple sources, these data should become more centralized and accessible.

COMPARISON: STATIC VS INTERACTIVE

Many of the interactive data visualization libraries Rho develops are inspired and informed by common static visualizations produced regularly with SAS and R. The Clinical Timelines originated from a plot of transplant follow-up with days since transplant on the x-axis and participants on the y-axis. The follow-up period is represented by a horizontal line, and a series of triangles, circles, arrows, stars, asterisks, and diamonds adorn it. These marks indicate clinical events such as unscheduled visits, the occurrence of medical signs of interest, and the timing of the primary outcome (Figure 10).

Affectionately referred to as a “lucky charms plot,” the graphic suffers from visual overload given the large variety of symbols and a lack of vertical space. The webpage solves the vertical issue because vertical space is infinite (Figure 11). To ameliorate the visual overload, we offset overlapping marks and give each ID more spacing to accommodate the offset. Circles, squares, diamonds, and directional triangles outlined in black represent timepoints. Borders around the symbols give each event type visual distinction that makes them easier to identify.

Figure 10 – Lucky Charms (static)

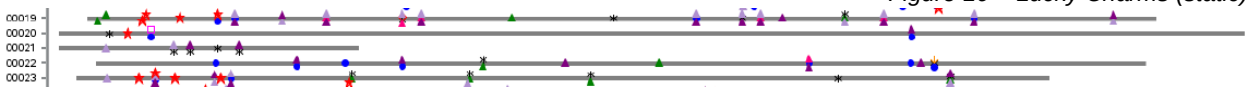
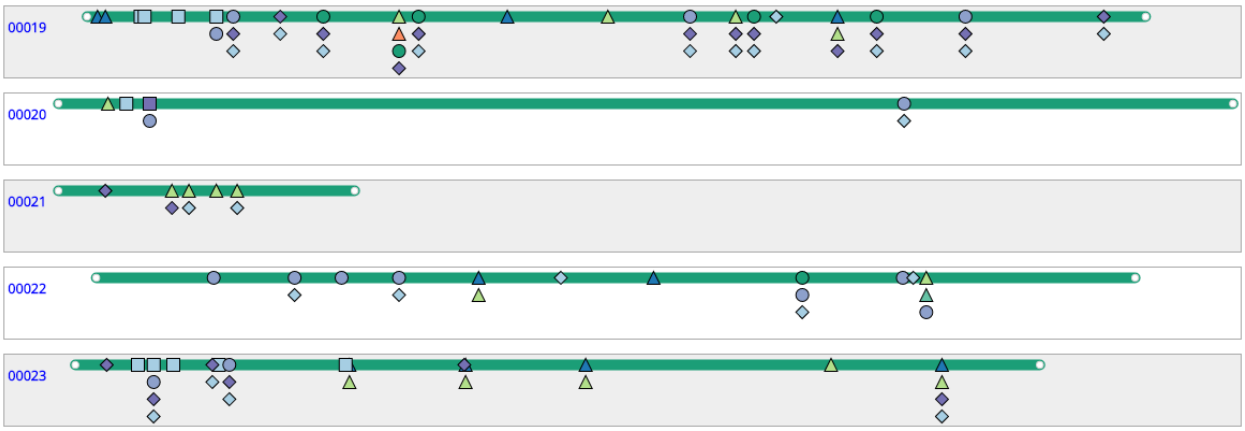


Figure 11 – Lucky Charms (interactive)



The interactive timelines chart provides ample space for each individual. Additionally, the marks have well-defined shapes and borders to give them distinction, and those that overlap are offset. Finally, striping clarifies the transition between one individual and the next.

CONCLUSION

The Clinical Timelines library provides robust control over the configuration and use of a timeline graphic. It expects data that match industry standards and so works with minimal configuration of the settings and input data. The user has substantial control over the timelines via time range controls, data filters, group stratification, and a toggle that switches to a profile view of the individual. Finally the library is open-source and available on GitHub. The data are out there; make them work for you.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Author Name:	Spencer Childress
Company:	Rho, Inc.
Address:	6330 Quadrangle Drive
City / Postcode:	Chapel Hill, NC 27517
Work Phone:	919 408 8000
Fax:	919 408 0999
Email:	spencer_childress@rhoworld.com
Web:	http://www.rhoworld.com/