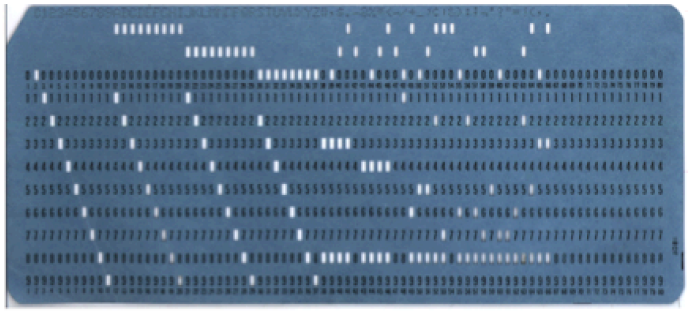
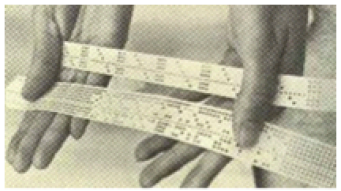
Event Analytics and the Visualization of Temporal Event Sequences

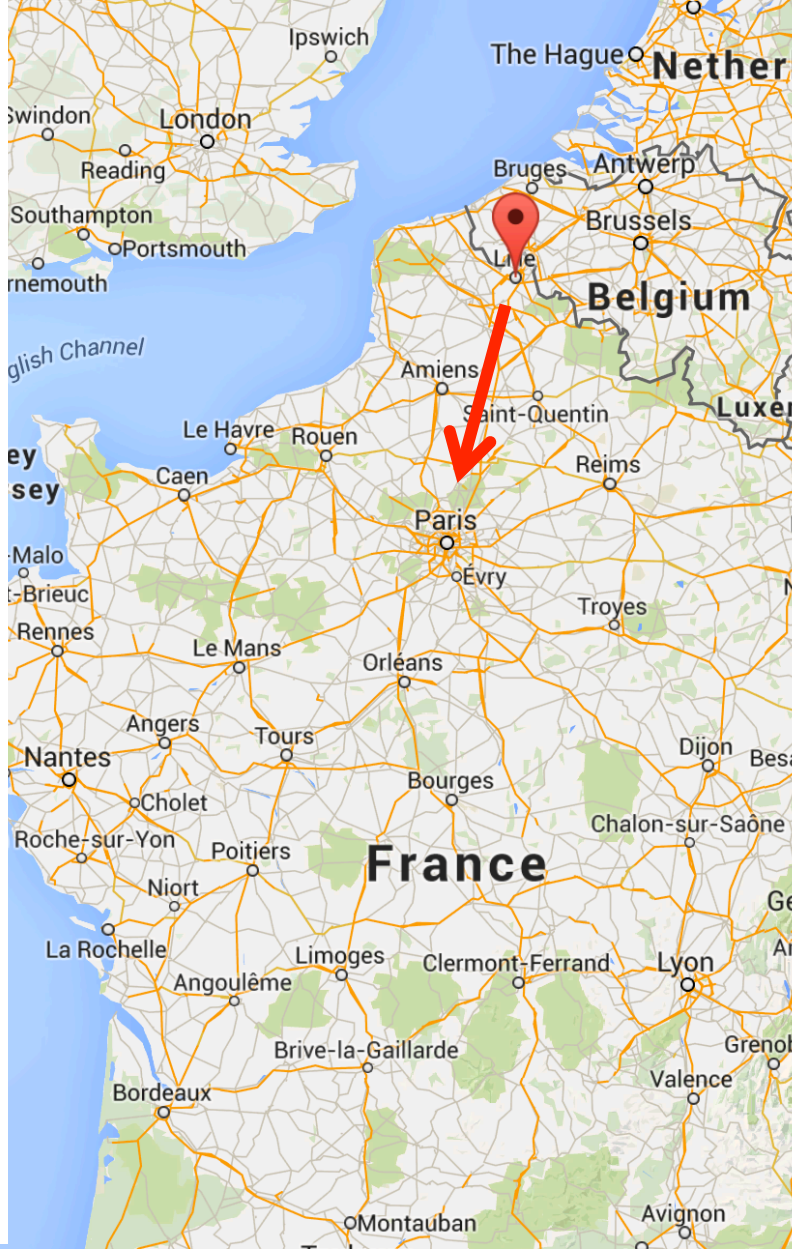
Catherine Plaisant

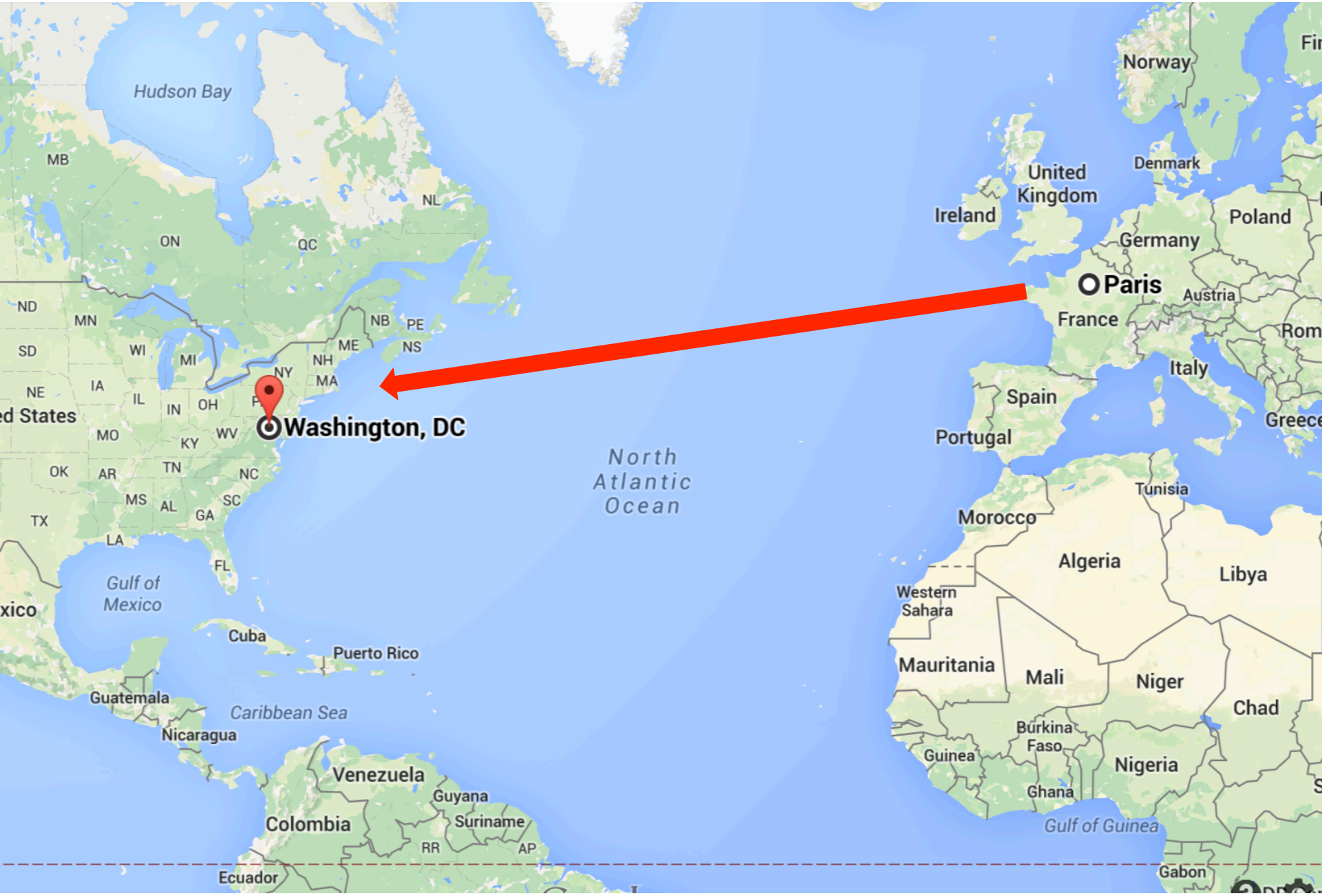
plaisant@c.umd.edu

Journées Vis~ June 8, 2017






Ipswich
The Hague
Nether
London
Reading
Bruges
Antwerp
Southampton
Portsmouth
Brussels
Belgium
Amiens
Saint-Quentin
Le Havre
Rouen
Caen
Reims
Paris
Évry
Troyes
Rennes
Le Mans
Orléans
Angers
Tours
Nantes
Dijon
Bourges
Cholet
Chalon-sur-Saône
Roche-sur-Yon
Poitiers
France
Niort
La Rochelle
Limoges
Clermont-Ferrand
Lyon
Angoulême
Brive-la-Gaillarde
Valence
Bordeaux
Montauban
Avignon
Toulouse
Nimes
Carcassonne
Montpellier
Marseille
Narbonne
Perpignan
Andorra
Gijón
A Coruña
Santander
Bilbao
Oviedo
San Sebastián
Lugo
Santiago de Compostela
Vitoria-Gasteiz
Ponferrada
León
Pamplona
Ourense
Burgos

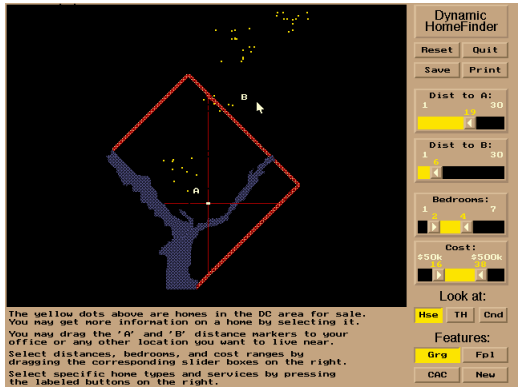

Hudson Bay
MB
NL
ON
QC
ND
MN
SD
WI
MI
NE
IA
IL
IN
OH
NY
PA
NB
PE
ME
NS
NH
MA
Washington, DC
MO
KY
WV
OK
AR
TN
NC
MS
AL
SC
GA
TX
LA
FL
Gulf of Mexico
Cuba
Puerto Rico
Guatemala
Caribbean Sea
Nicaragua
Venezuela
Guyana
Suriname
Colombia
RR
AP
Ecuador
North Atlantic Ocean
Norway
Denmark
United Kingdom
Ireland
Poland
Germany
Paris
France
Austria
Italy
Spain
Portugal
Tunisia
Morocco
Algeria
Libya
Western Sahara
Mauritania
Mali
Niger
Chad
Burkina Faso
Guinea
Nigeria
Ghana
Gulf of Guinea
Gabon

Love
12:35

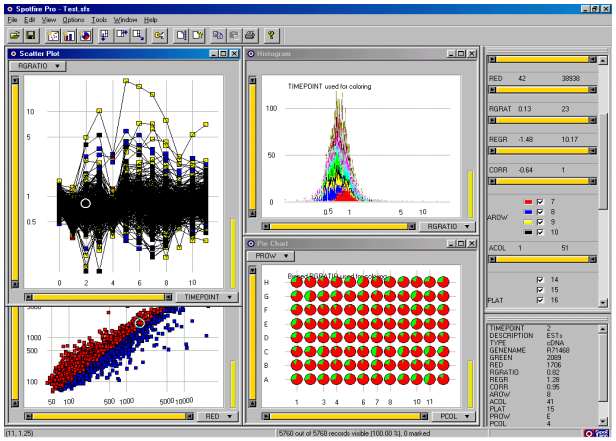
HCIL

Ben Shneiderman
Research collaborator for 30 years

Information visualization

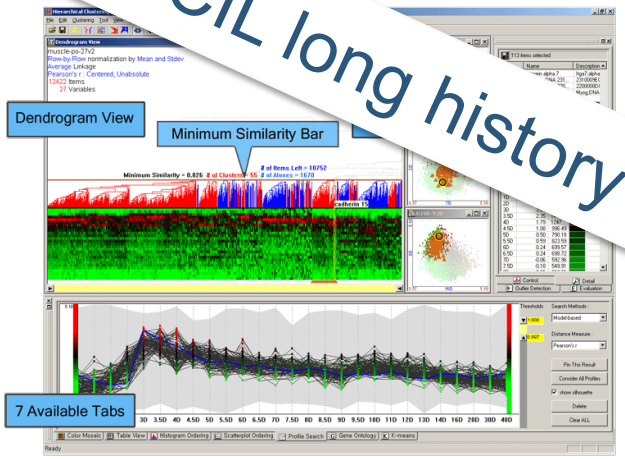


Spotfire

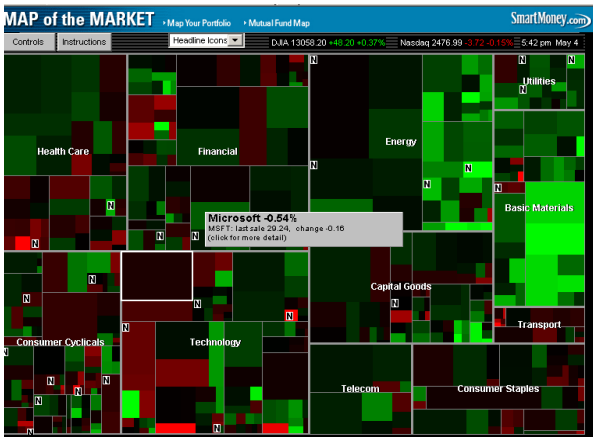


Home Finder and Filmfinder prototypes lead to SpotFire

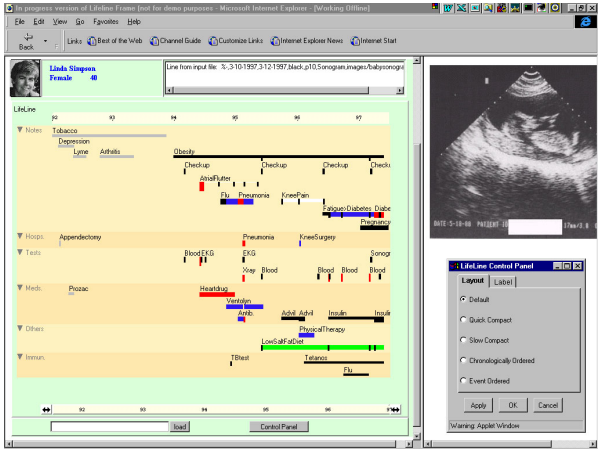
HCIL long history



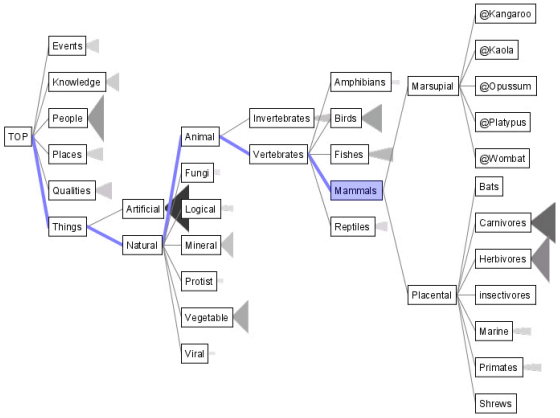
Hierarchical Clustering Explorer HCE



Treemap

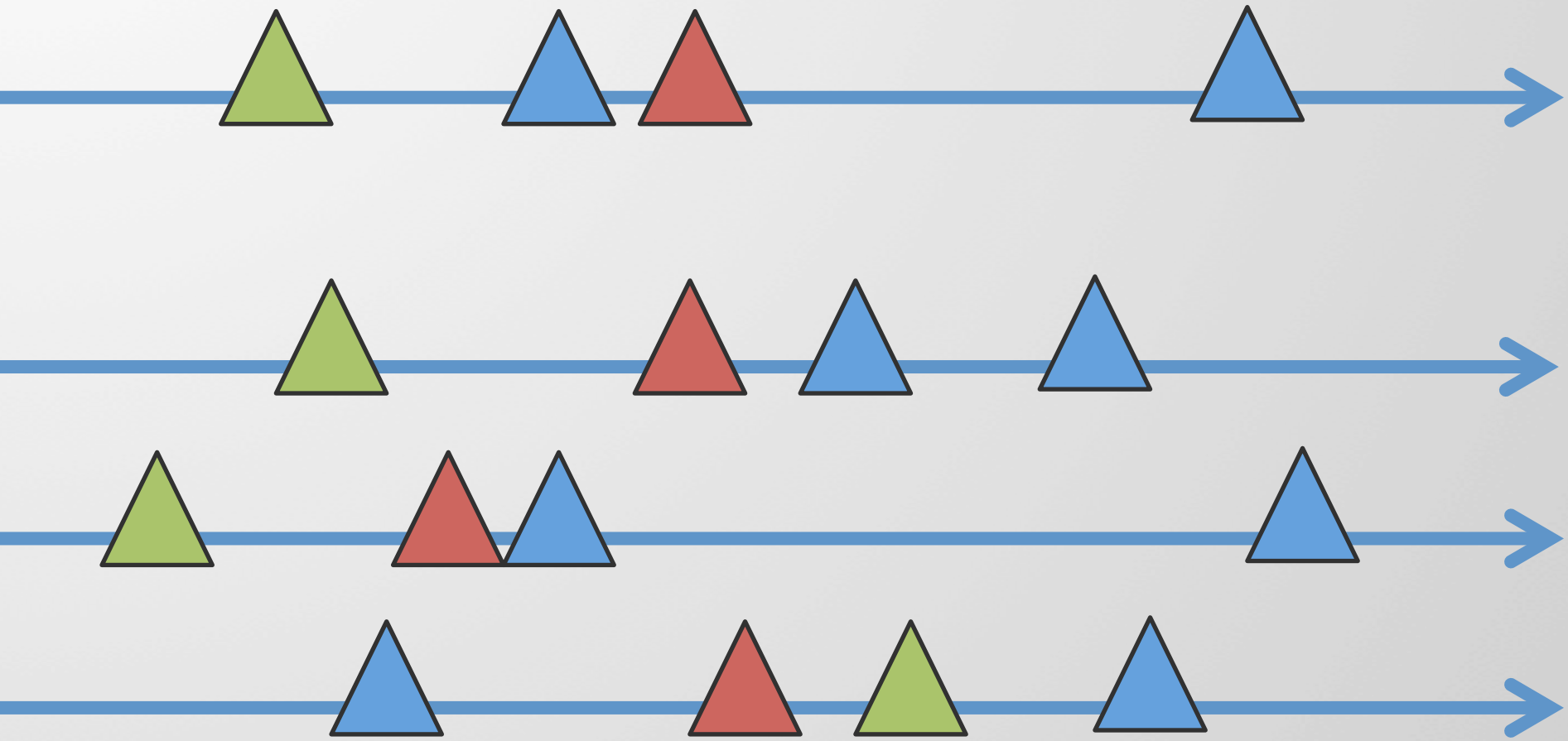


Lifelines



SpaceTree

Event Analytics

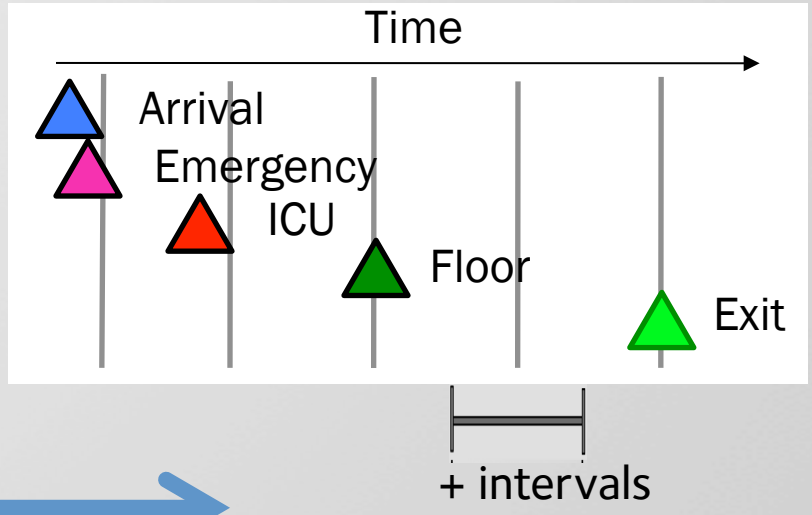


Numerical

Patient ID: 12345	
04/26/2010 10:00	31.03
04/26/2010 10:15	31.01
04/26/2010 10:30	31.02
04/26/2010 10:45	31.08
04/26/2010 11:00	31.16

Categorical

Patient ID: 45851737	
12/02/2008 14:26	Arrival
12/02/2008 14:36	Emergency
12/02/2008 22:44	ICU
12/05/2008 05:07	Floor
12/14/2008 06:19	Exit



e.g. High/Normal/Low

Prototype evolution

Tool	Event Types	Records	Display
LifeLines	Points, Intervals	One	Individual
LifeLines2	Points	Many	Individual, Summary
Similan	Points	Many	Individual
LifeFlow	Points	Many	Individual, Aggregate
EventFlow	Points, Intervals	Many	Individual, Aggregate

ORACLE Adobe Children's National MedStar Washington Hospital Center NIH National Institutes of Health

University of Maryland School of Pharmacy UF University of Florida Yale University School of Medicine UBC The University of British Columbia CDC Centers for Disease Control and Prevention

www.cs.umd.edu/hcil/*toolname*

Many application domains

Electronic Health Records: symptoms, treatment, lab test

Student records: course, paper, proposal, defense, etc.

Web logs, usability logs, security etc.

Traffic incident logs: confirmed, unit arrived, lane closed etc.

A SINGLE RECORD

What is the situation?
What has been done?
What should we do now?

MULTIPLE RECORDS

RECORD

RECORD

RECORD

RECORD

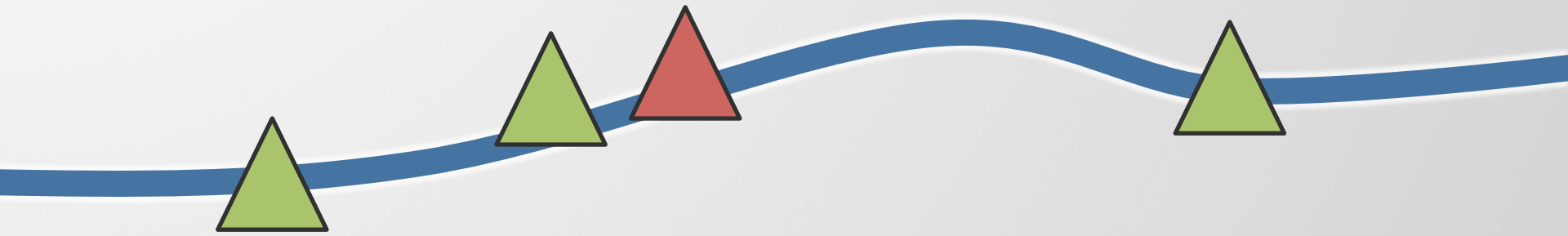
RECORD

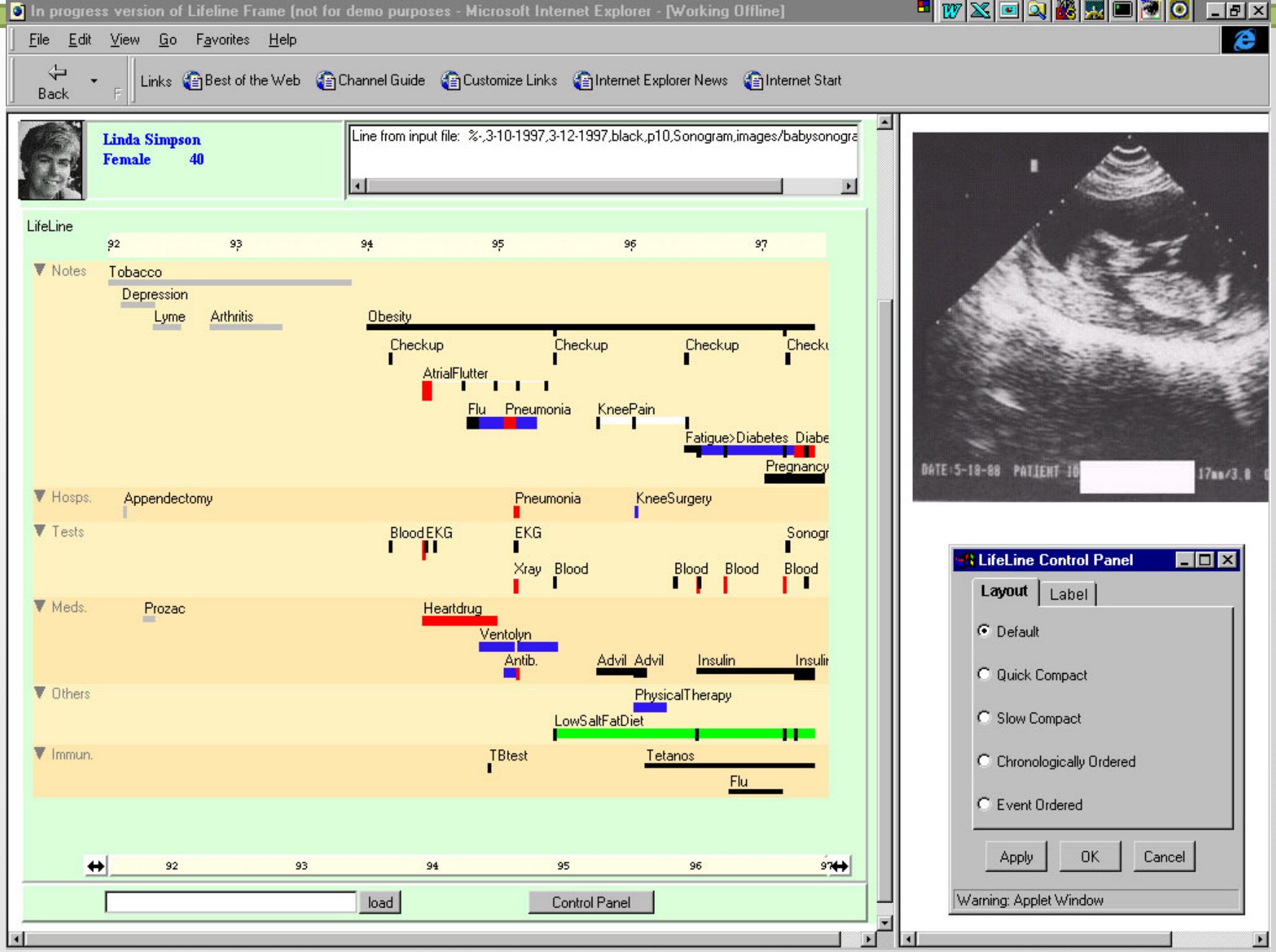
Are we following guidelines?

Do you have patients for my clinical trial?

Retrospective analysis:
How are opioids prescribed?
Patterns of readmissions?
Drug adverse reactions?
How can we improve care?

A SINGLE RECORD

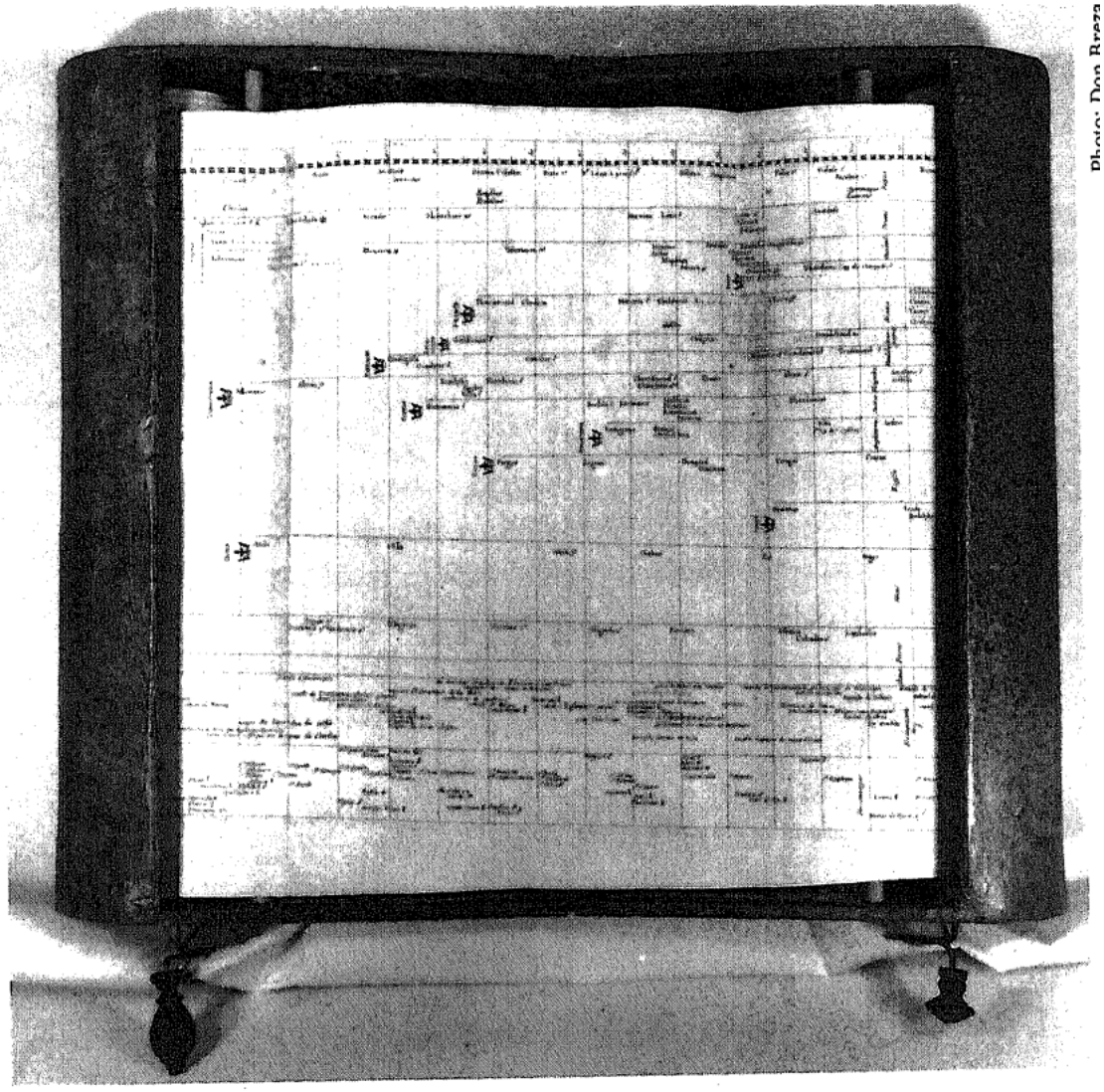




LifeLines – Single Patient

http://www.cs.umd.edu/hcil/lifelines

1753



Jacques Barbeu-Dubourg's *Machine chronologique*, opened to the years A.D. 360 to A.D. 510 on the *Carte chronographique*, showing the kingdoms of the Huns, Sweden, Scotland, England, Swabia, the Visigoths, the Vandals, Bourgogne, France, and Italy. Rare Books and Special Collections, Princeton University Library.

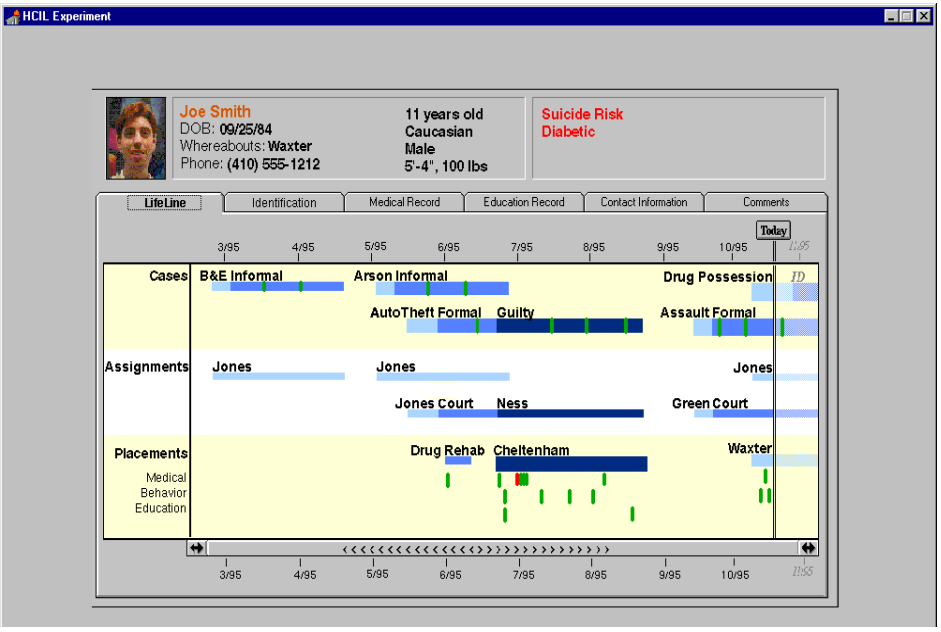
Jacques Barbeu Du Bourg - 1753

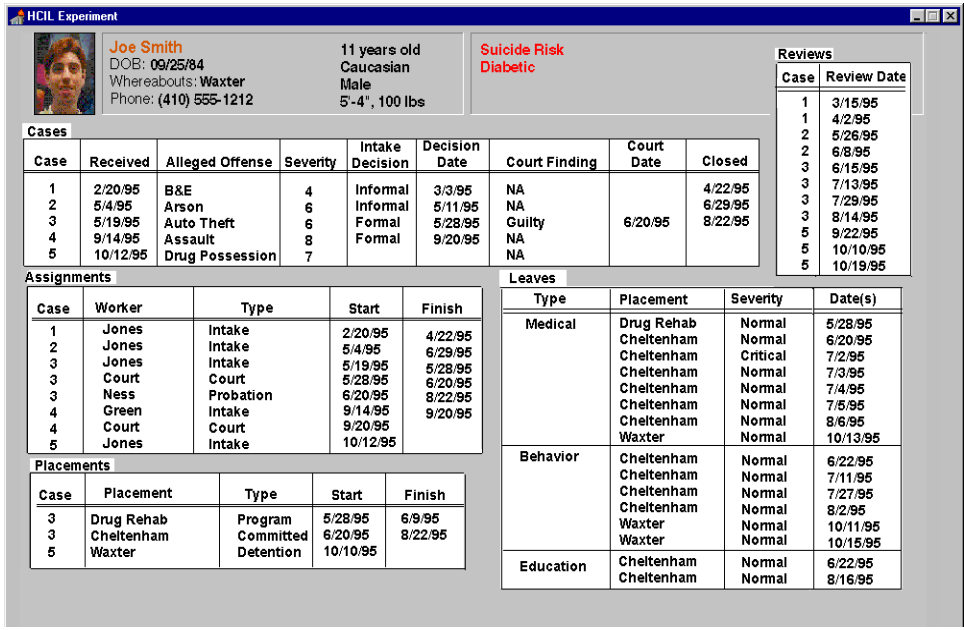
http://gallica.bnf.fr/ark:/12148/bpt6k1314025

Controlled experiment

- 36 participants used either LifeLines or Tabular display (static display only)
- Series of tests:

  first impression, 31 question quiz, subjective satisfaction questionnaire, recall test, spatial ability

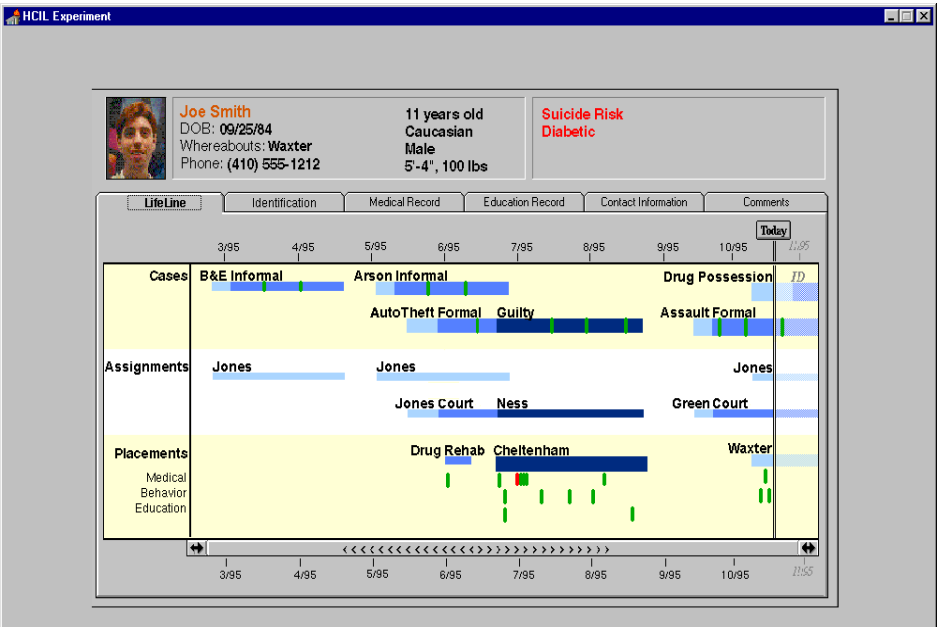


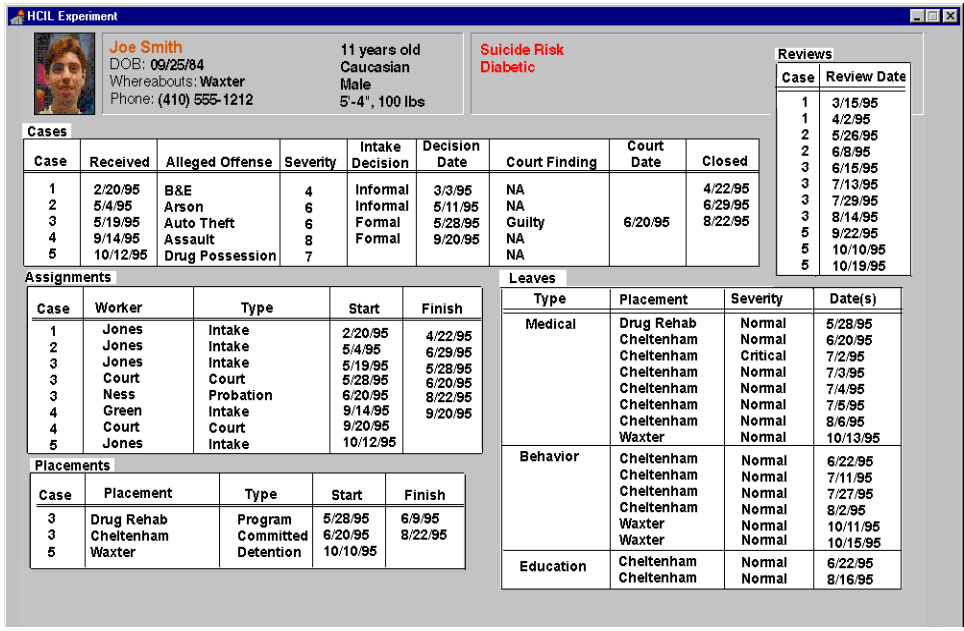


Alonso, D., Rose, A., Plaisant, C., Norman, K.
Viewing Personal History Records: A Comparison of Tabular Format and Graphical Presentation Using LifeLines
Behavior and Information Technology 17, 5, 1998, 249-262.

Strong benefits

- LifeLines twice as fast for many tasks
  (e.g. time interval comparison, or task across categories)
- Better recall
  4.3 vs 2.8 correct out of 6 questions
- More accurate “1st impression”





MULTIPLE RECORDS

RECORD

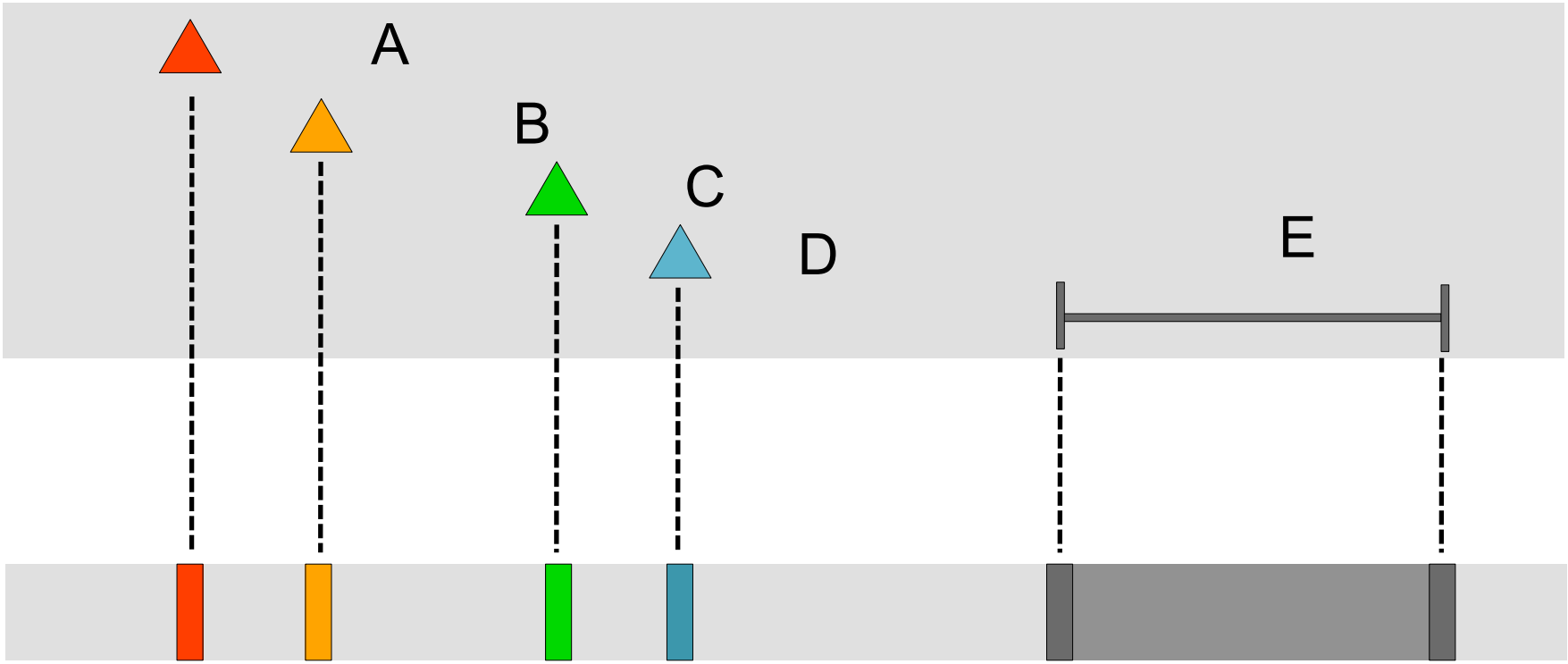
RECORD

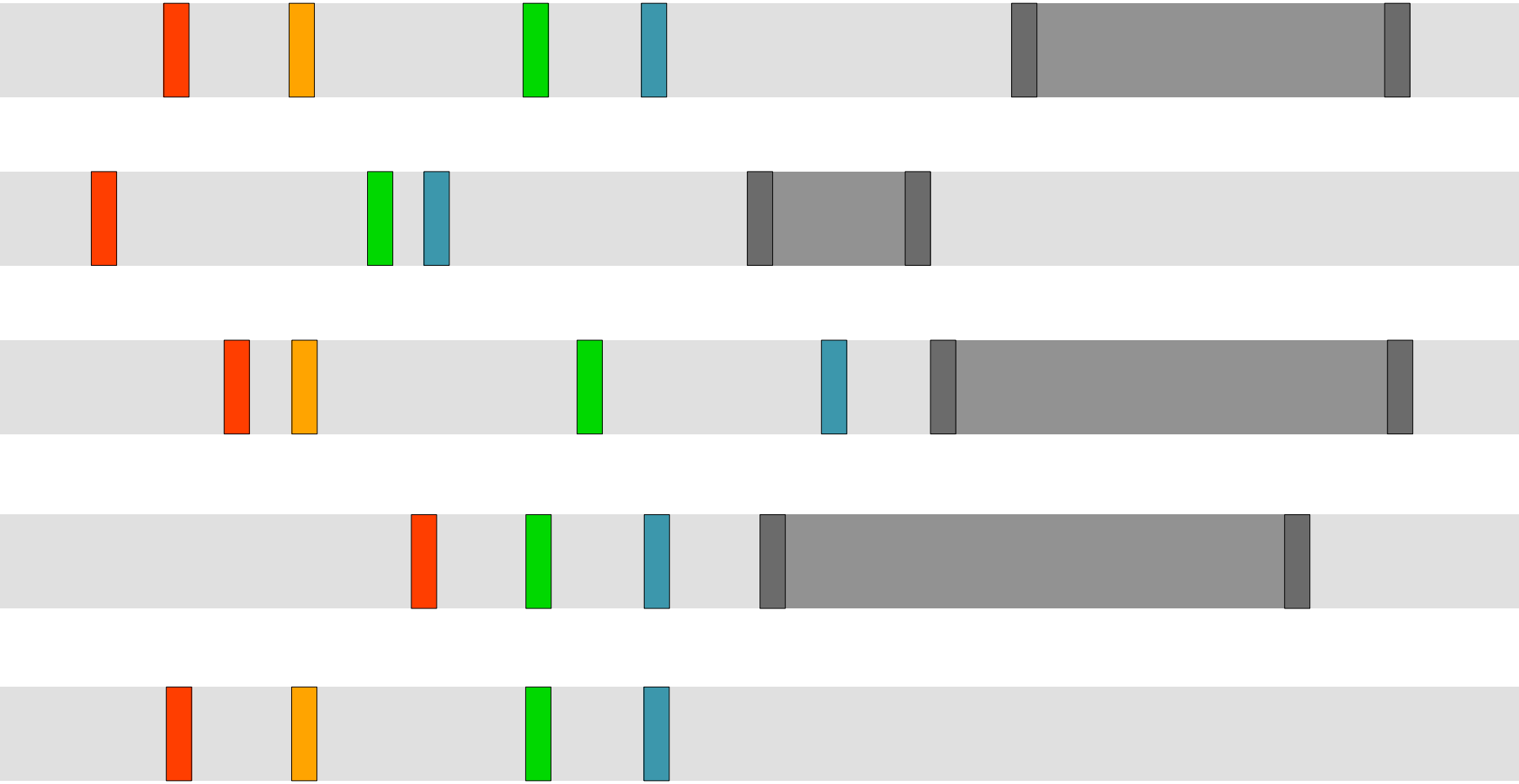
RECORD

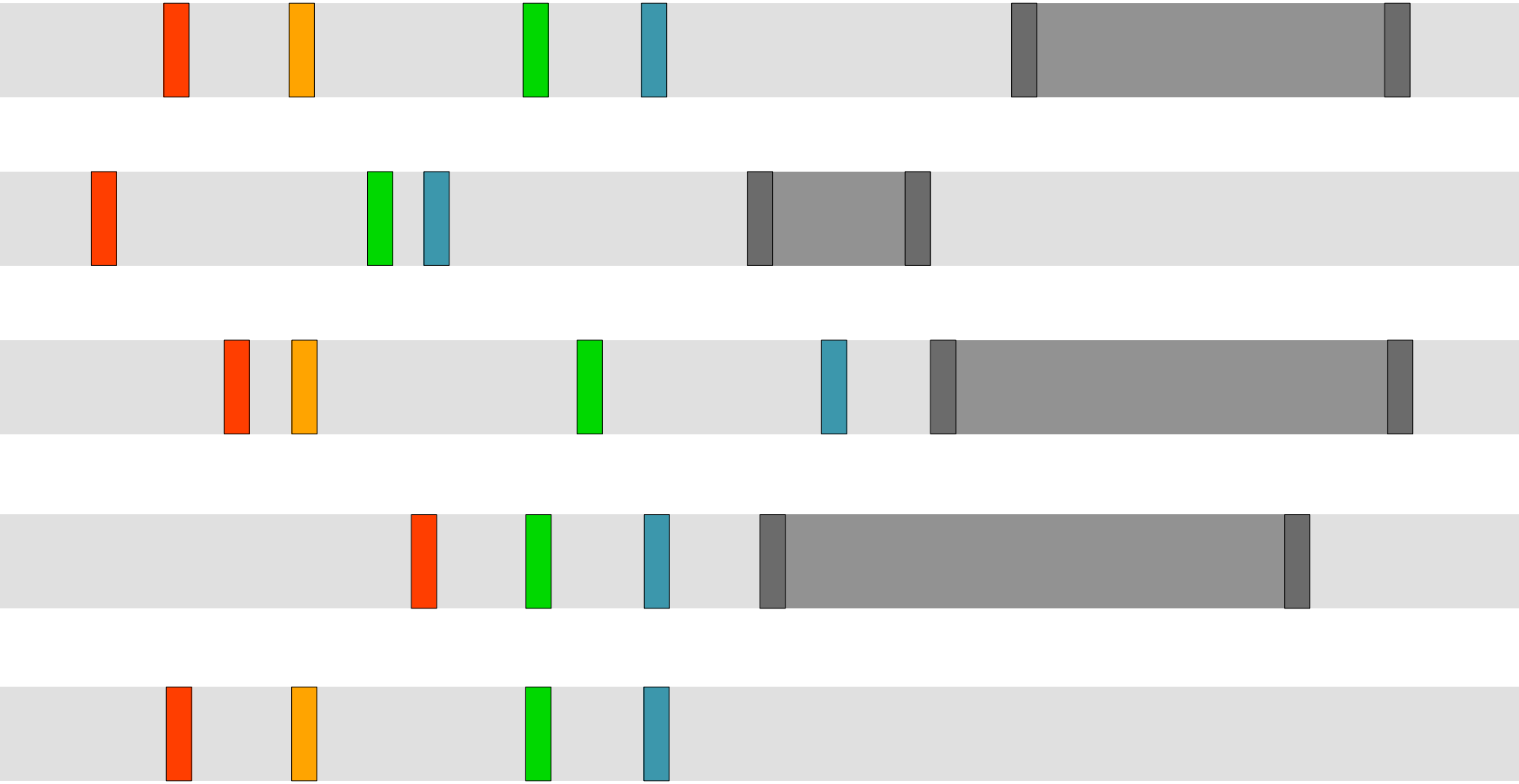
RECORD

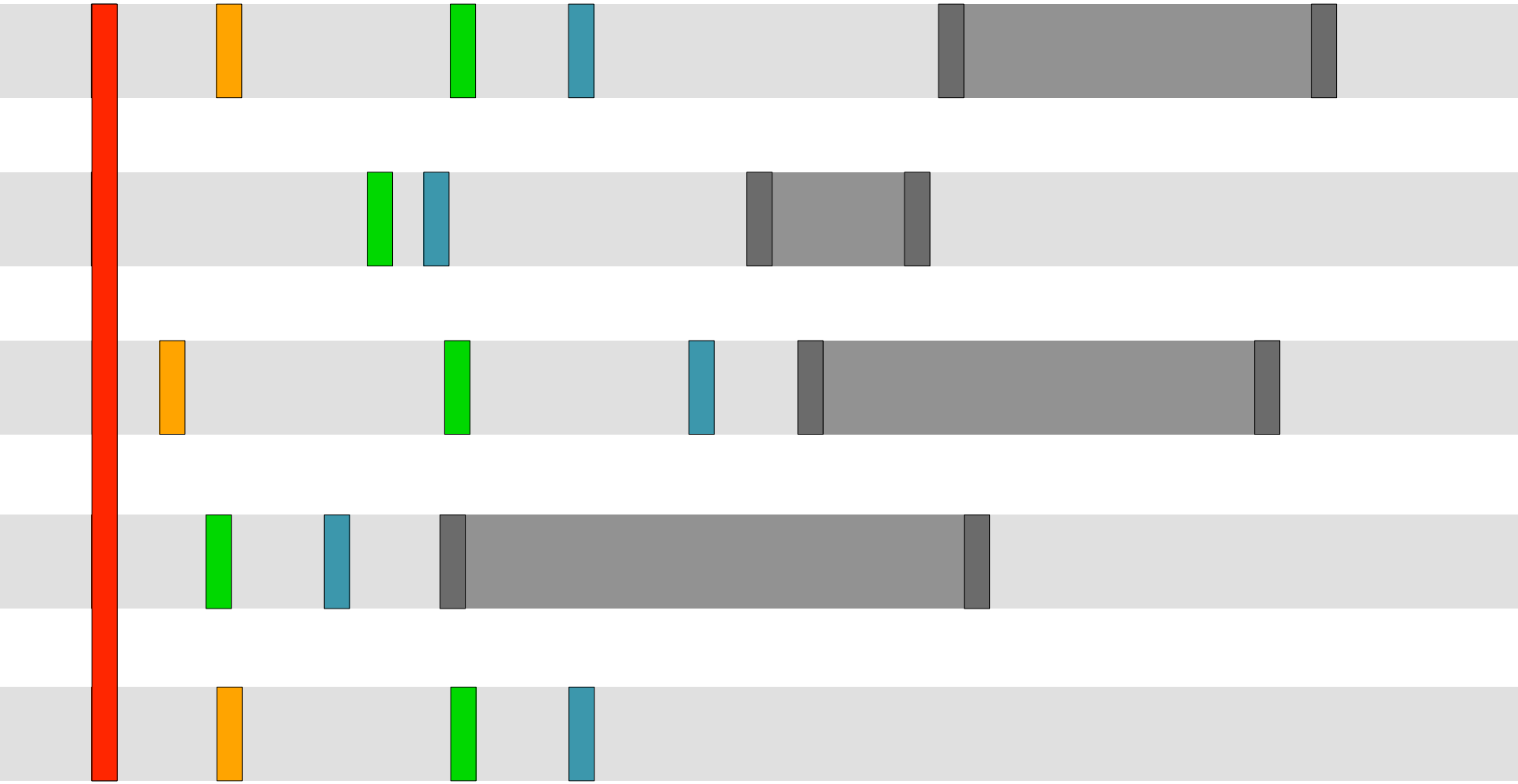
RECORD

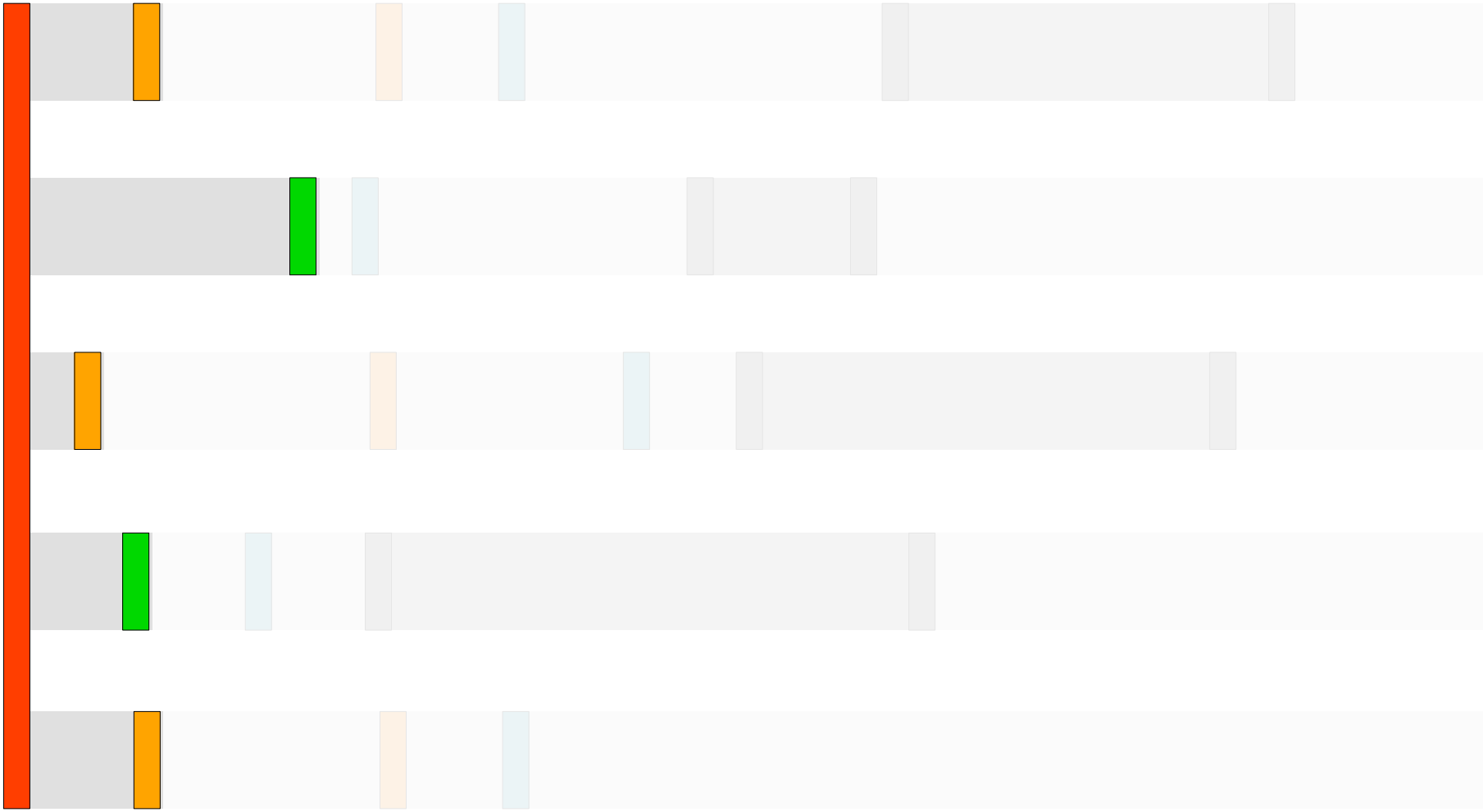
Constructing the EventFlow Overview

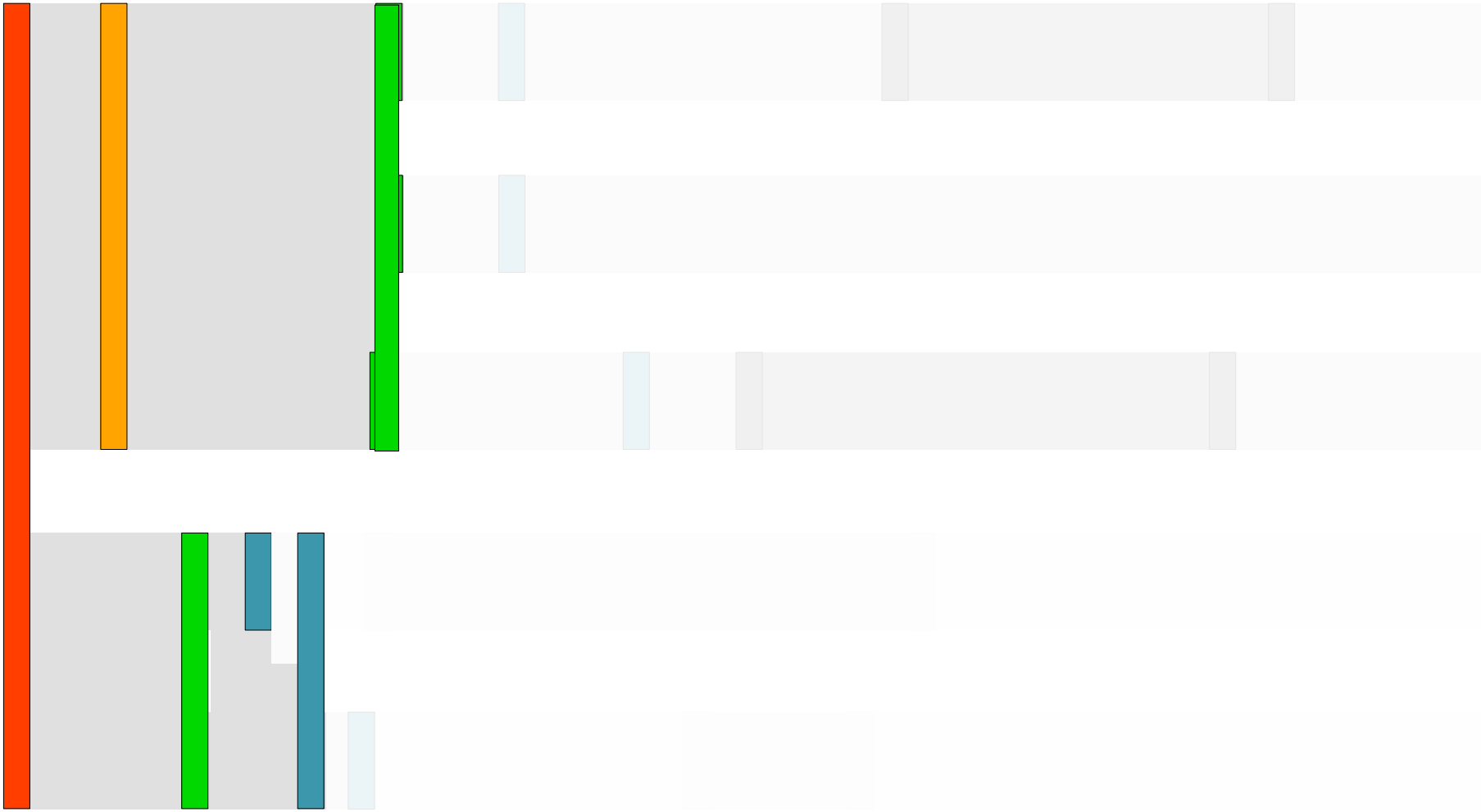


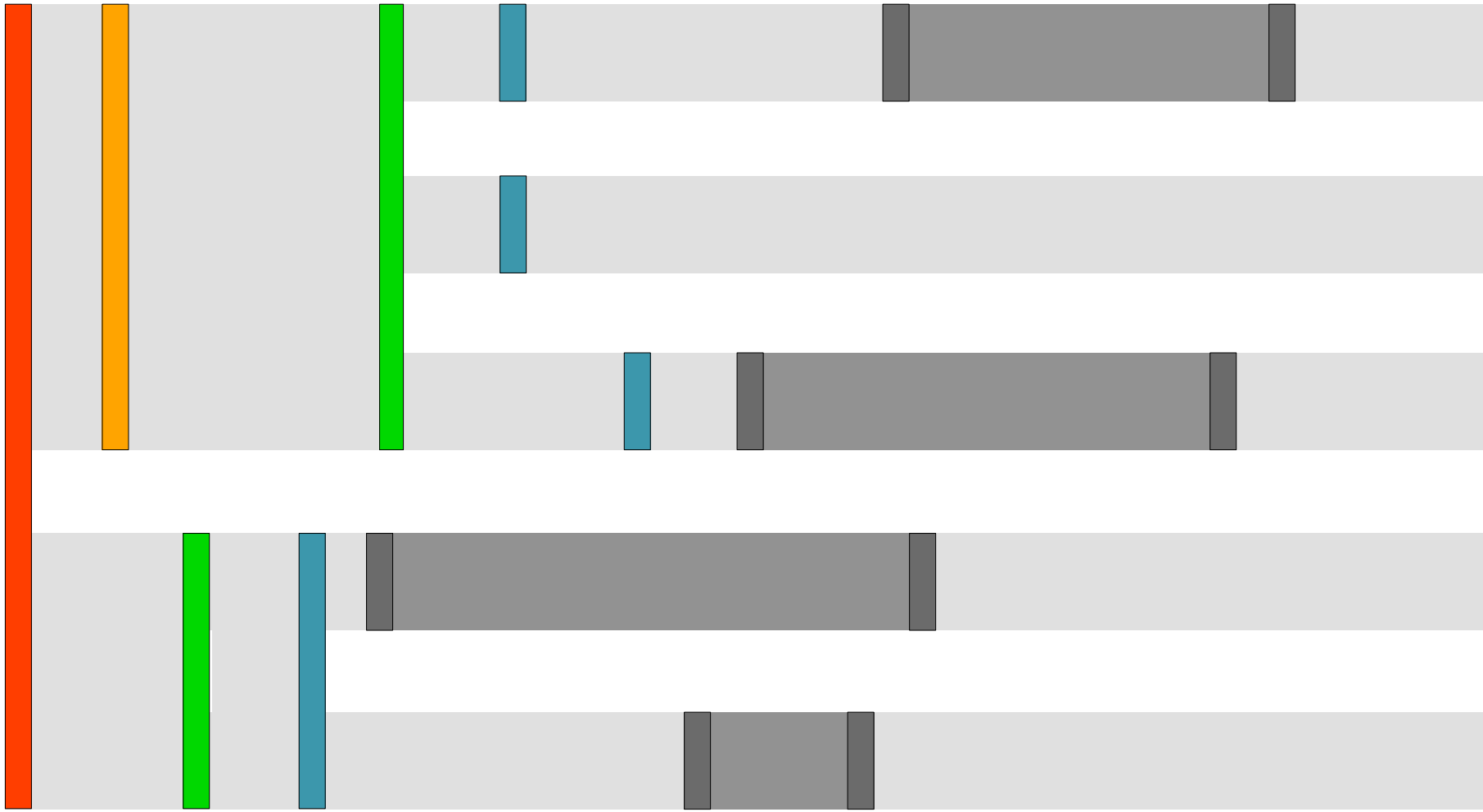


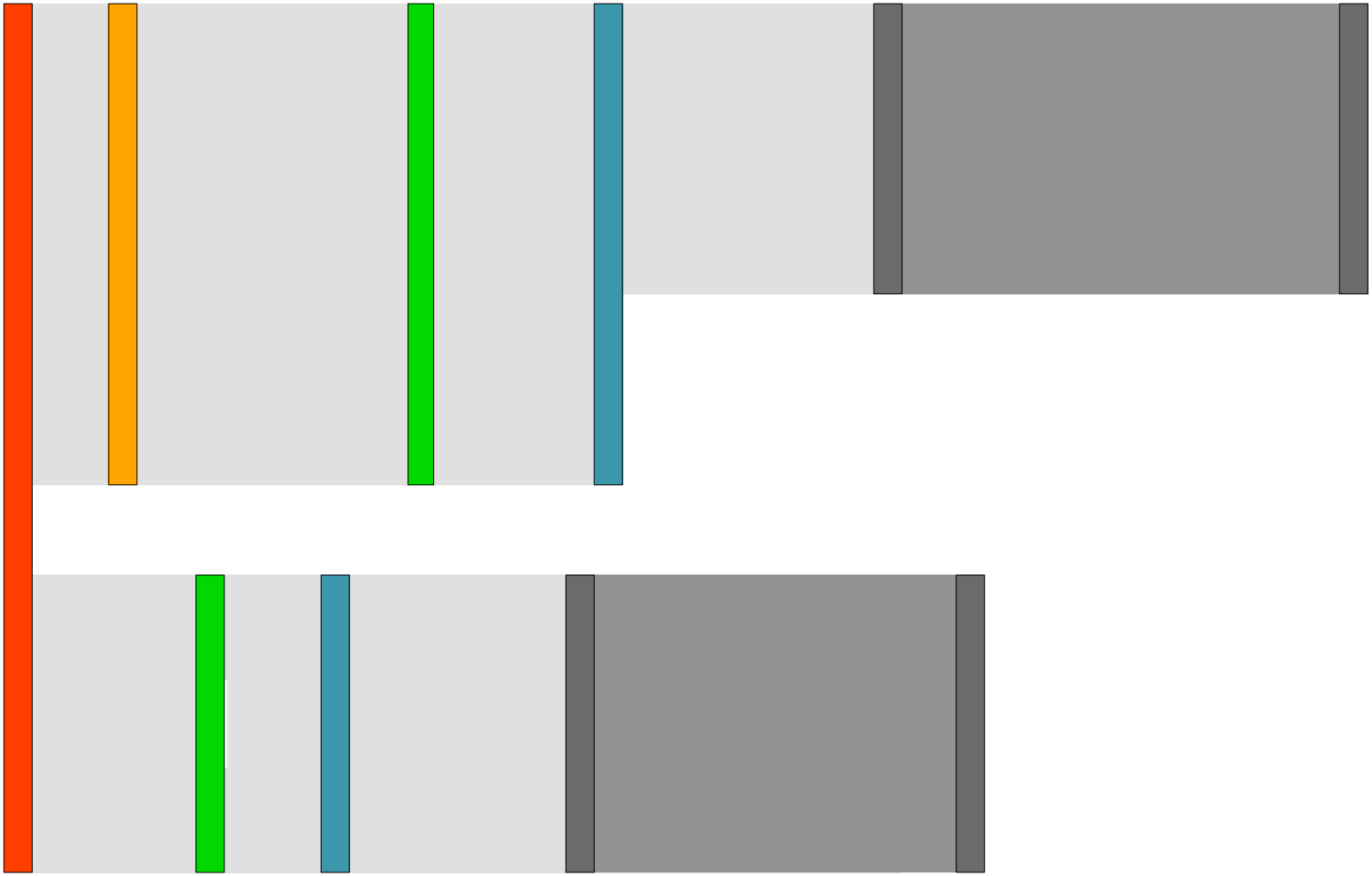


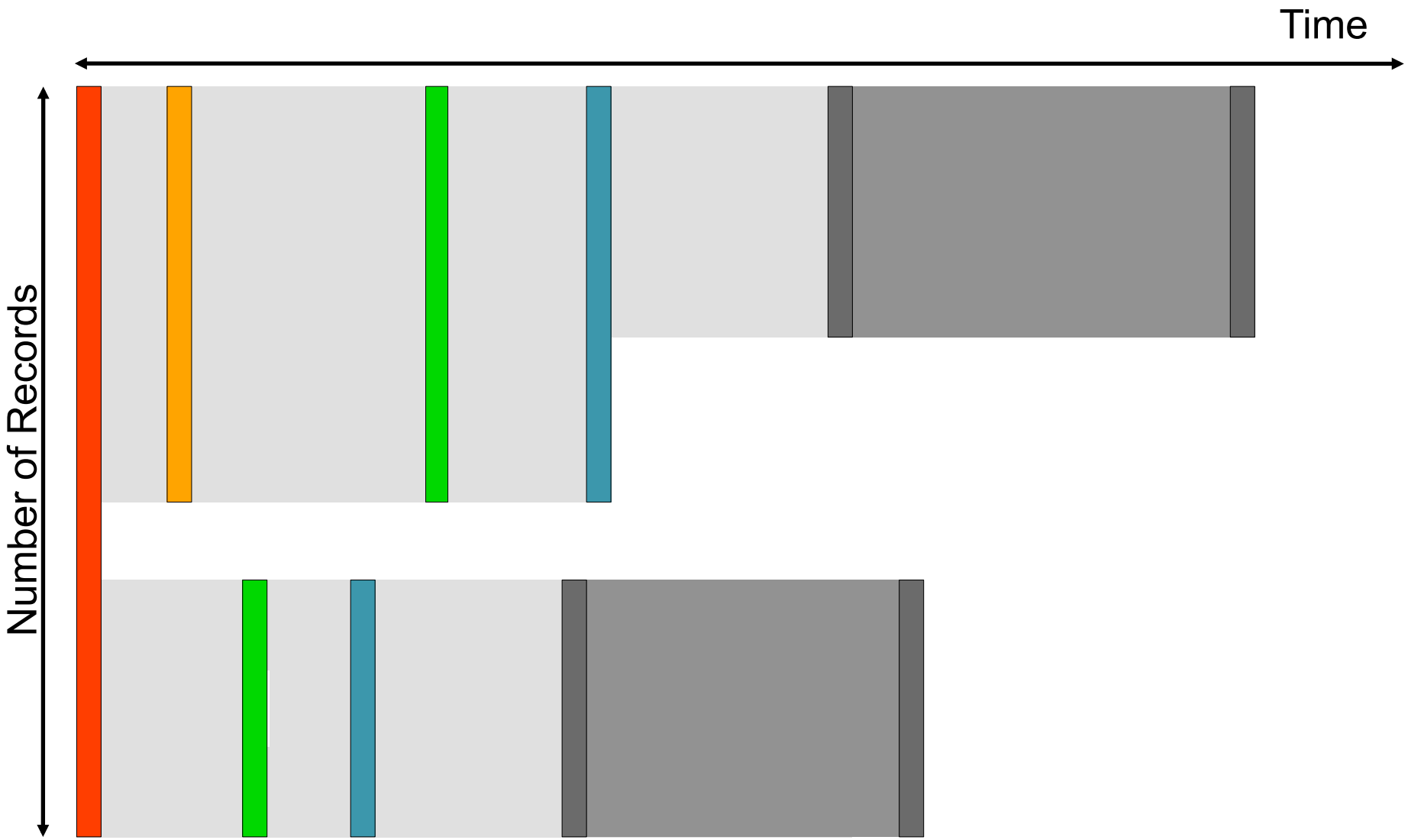




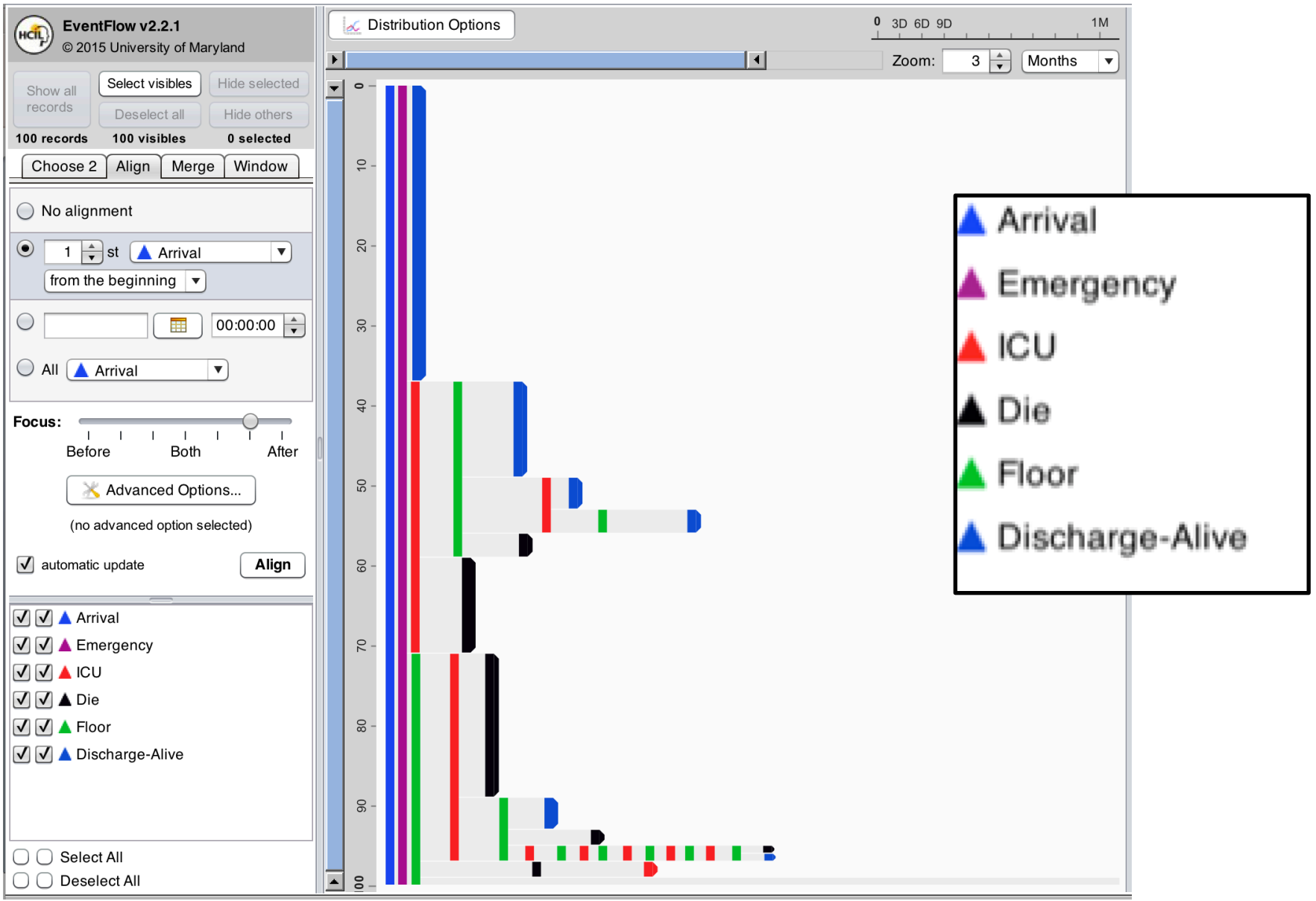


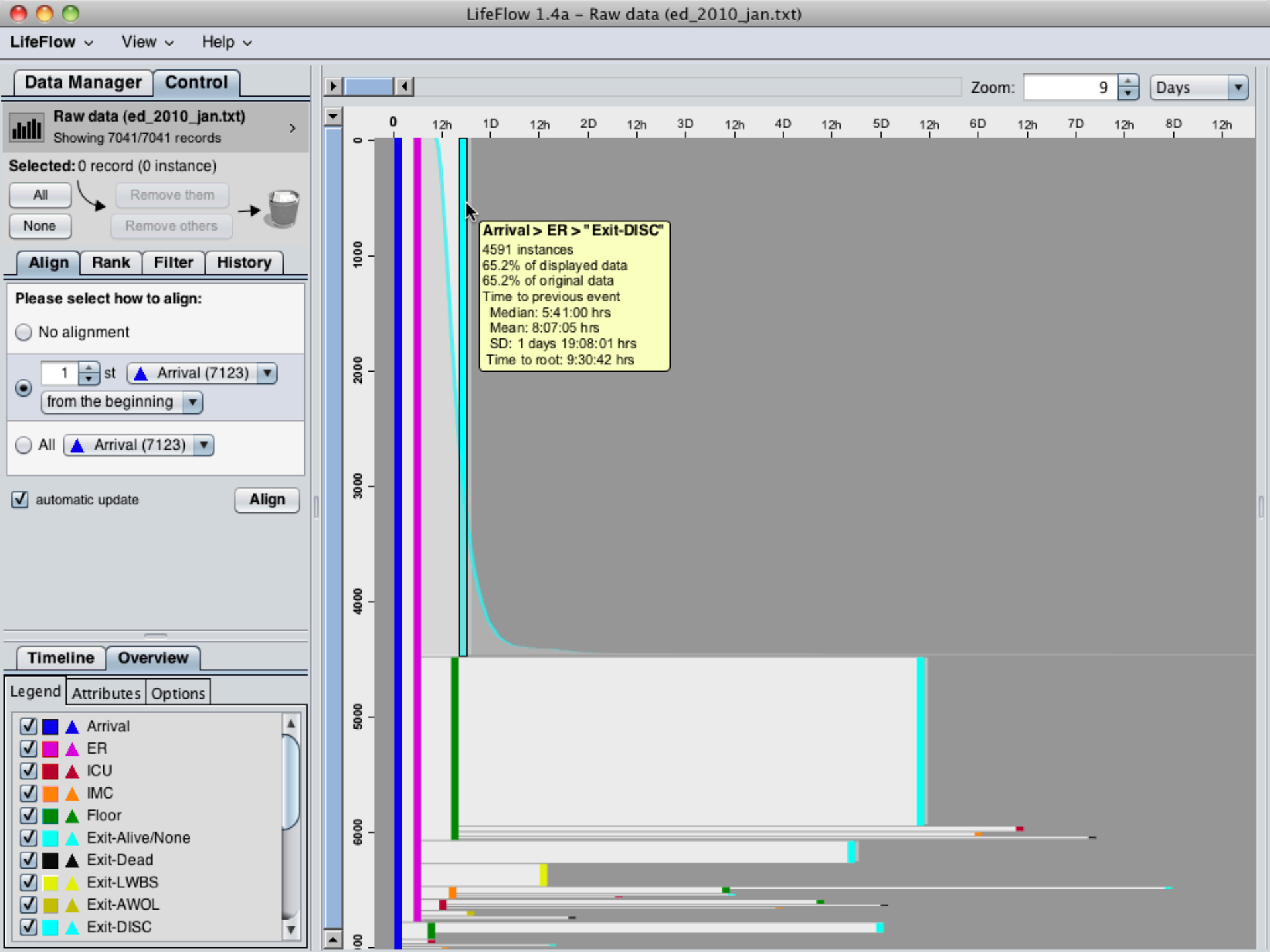





Time
Number of Records

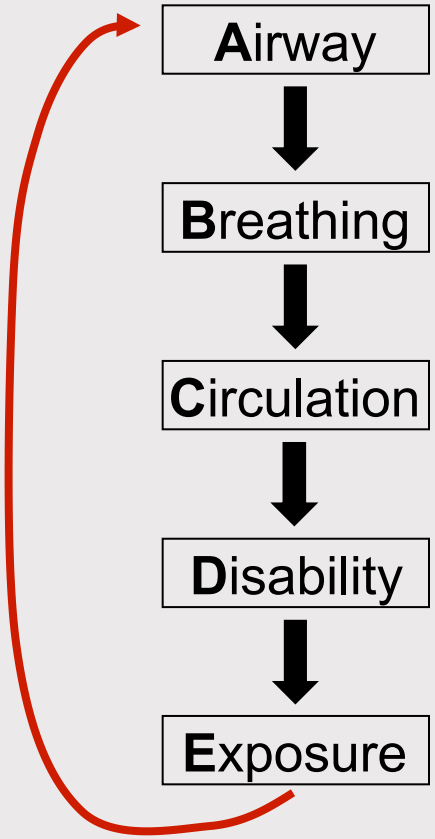
EventFlow ~ e.g. Transfers within Hospital




LifeFlow 1.4a – Raw data (ed_2010_jan.txt)
LifeFlow
View
Help
Data Manager
Control
Raw data (ed_2010_jan.txt)
Showing 7041/7041 records
Selected: 0 record (0 instance)
All
None
Remove them
Remove others
Align
Rank
Filter
History
Please select how to align:
No alignment
1
st
Arrival (7123)
from the beginning
All
Arrival (7123)
automatic update
Align
Timeline
Overview
Legend
Attributes
Options
Arrival
ER
ICU
IMC
Floor
Exit-Alive/None
Exit-Dead
Exit-LWBS
Exit-AWOL
Exit-DISC
Zoom:
9
Days
0
12h
1D
2D
3D
4D
5D
6D
7D
8D
Arrival > ER > "Exit-DISC"
4591 instances
65.2% of displayed data
65.2% of original data
Time to previous event
Median: 5:41:00 hrs
Mean: 8:07:05 hrs
SD: 1 days 19:08:01 hrs
Time to root: 9:30:42 hrs

Children's Hospital in DC: Trauma Bay

Primary Survey (ABCDE)

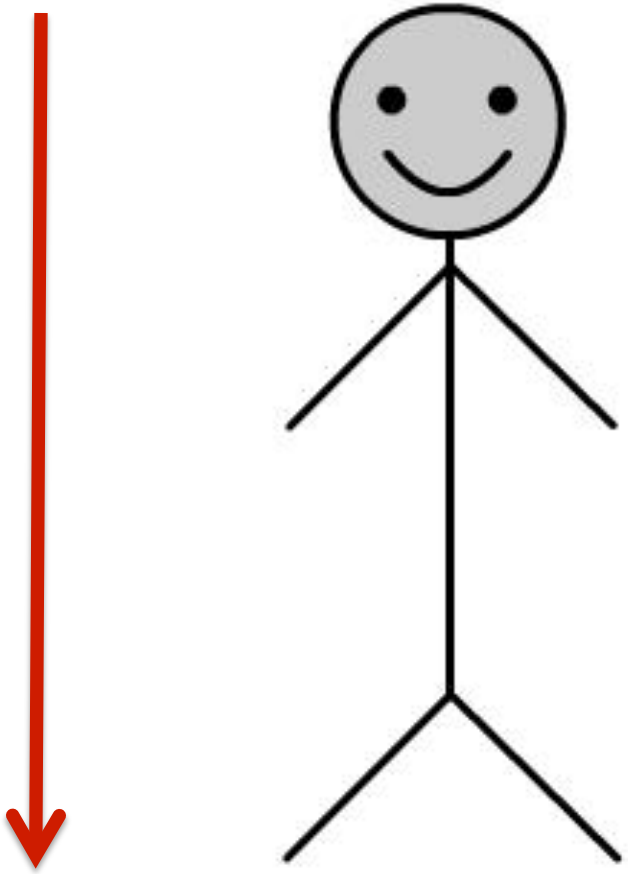


Identify and manage life-threatening conditions in a sequential manner

Children's National Medical Center

Secondary Survey

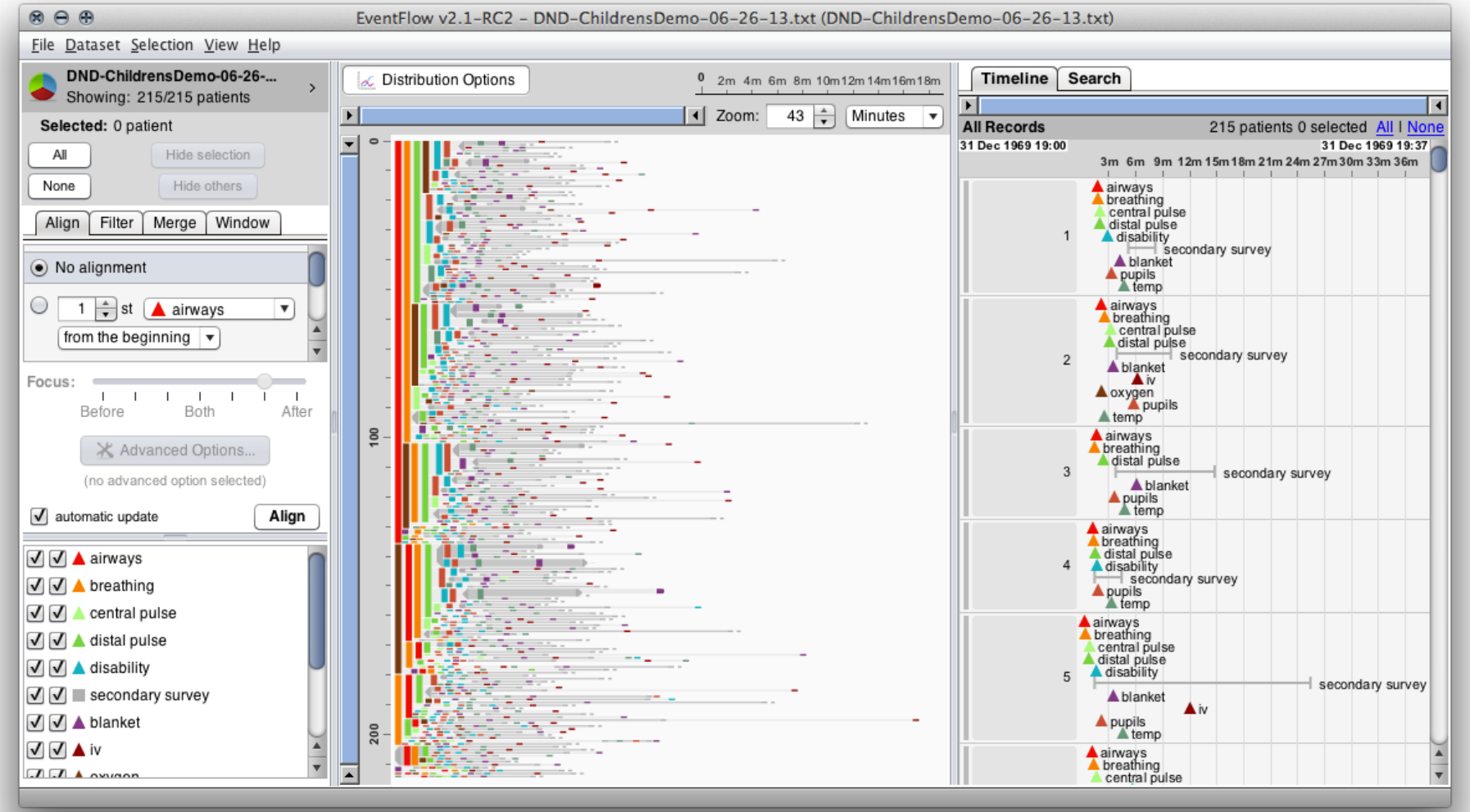
Begins once Primary Survey is complete and patient is stable



Head to Toe Examinati on

Children's National Medical Center

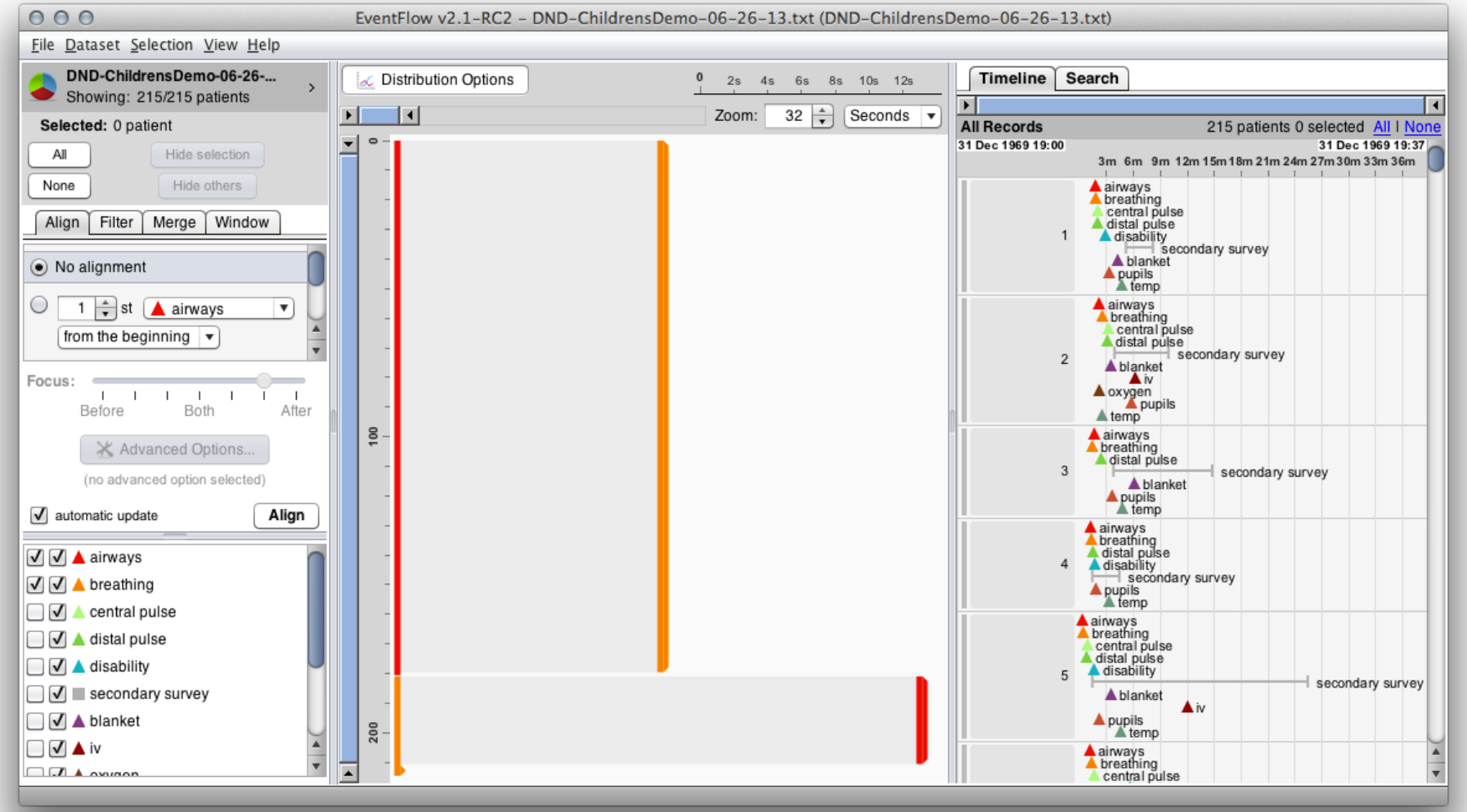
At first… Confetti.
Need strategy to get answers



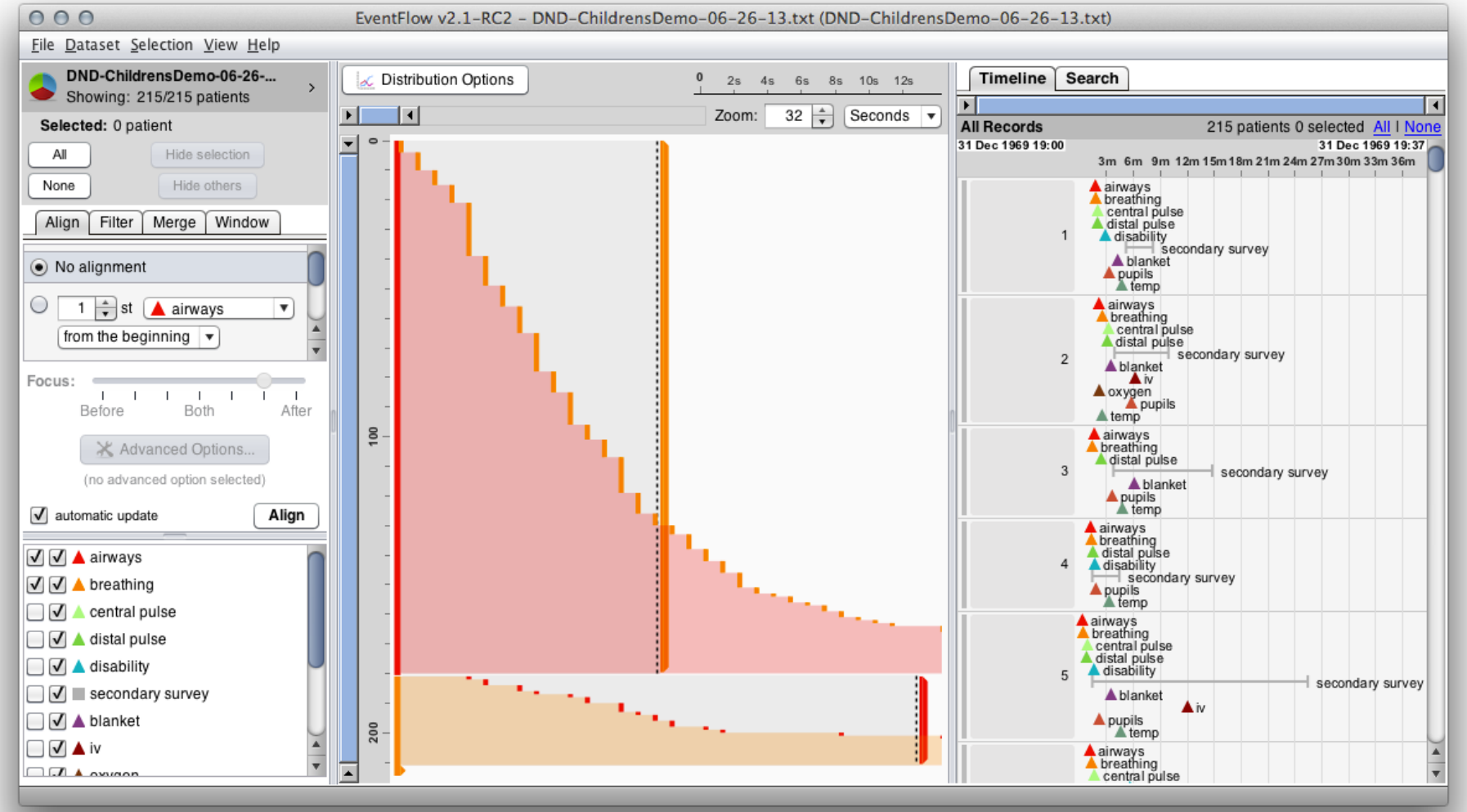
skip

Focusing on the first two events

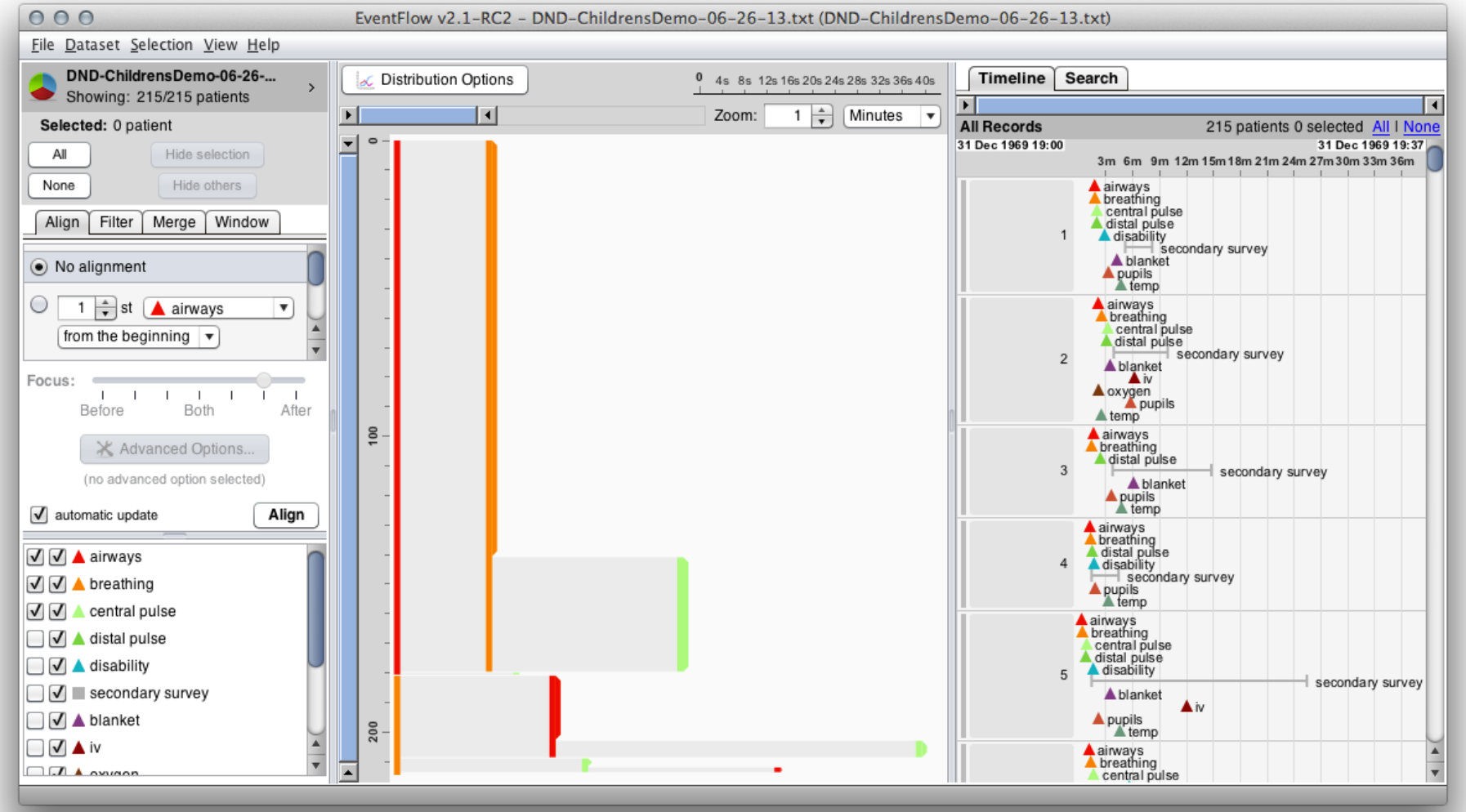
84% of patients are checked in the correct order.
The most common deviation is that the breathing is checked before the airways (14% of patients)
Reversed group takes longer on average than the correct sequence



Distributions

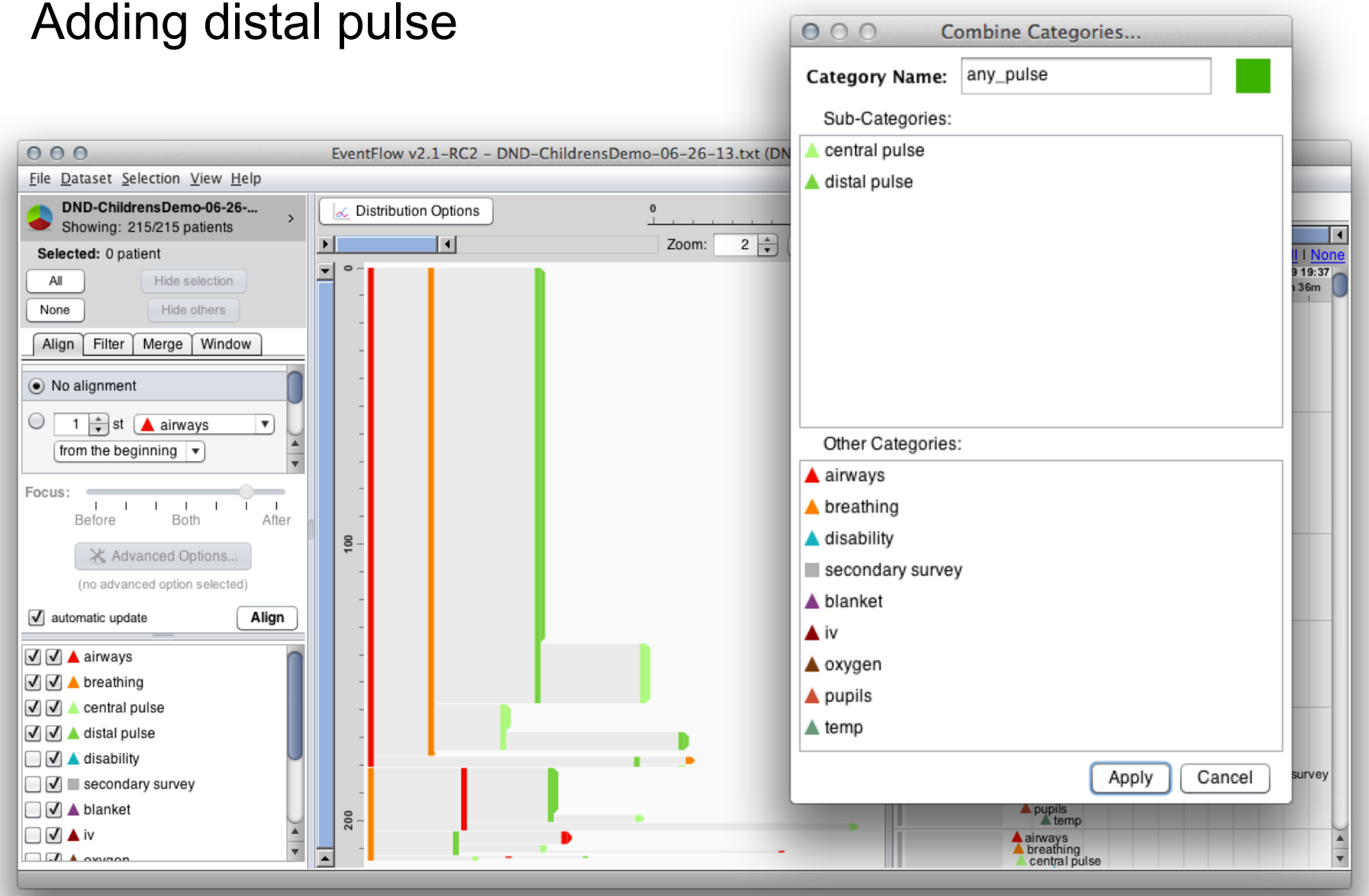


Adding third event type (central pulse)

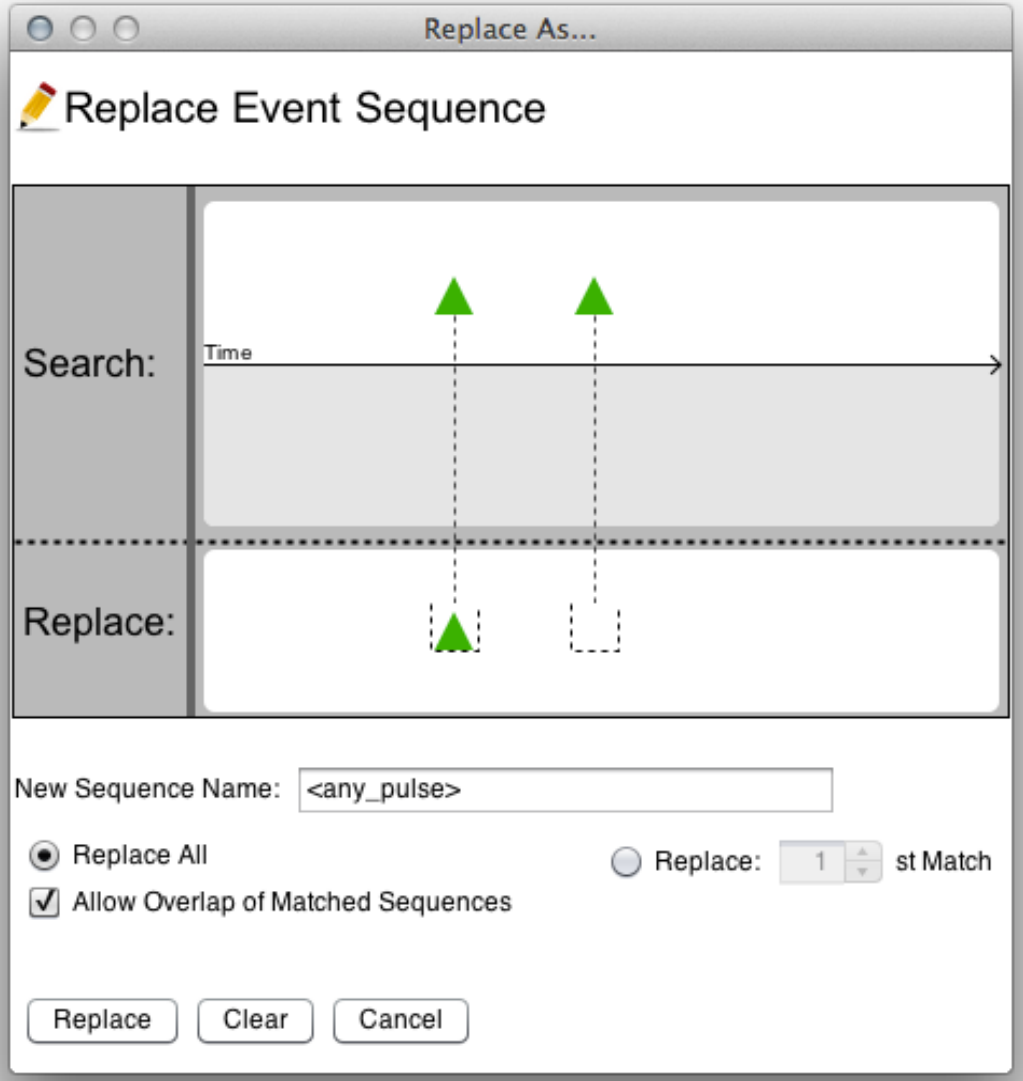


Adding distal pulse

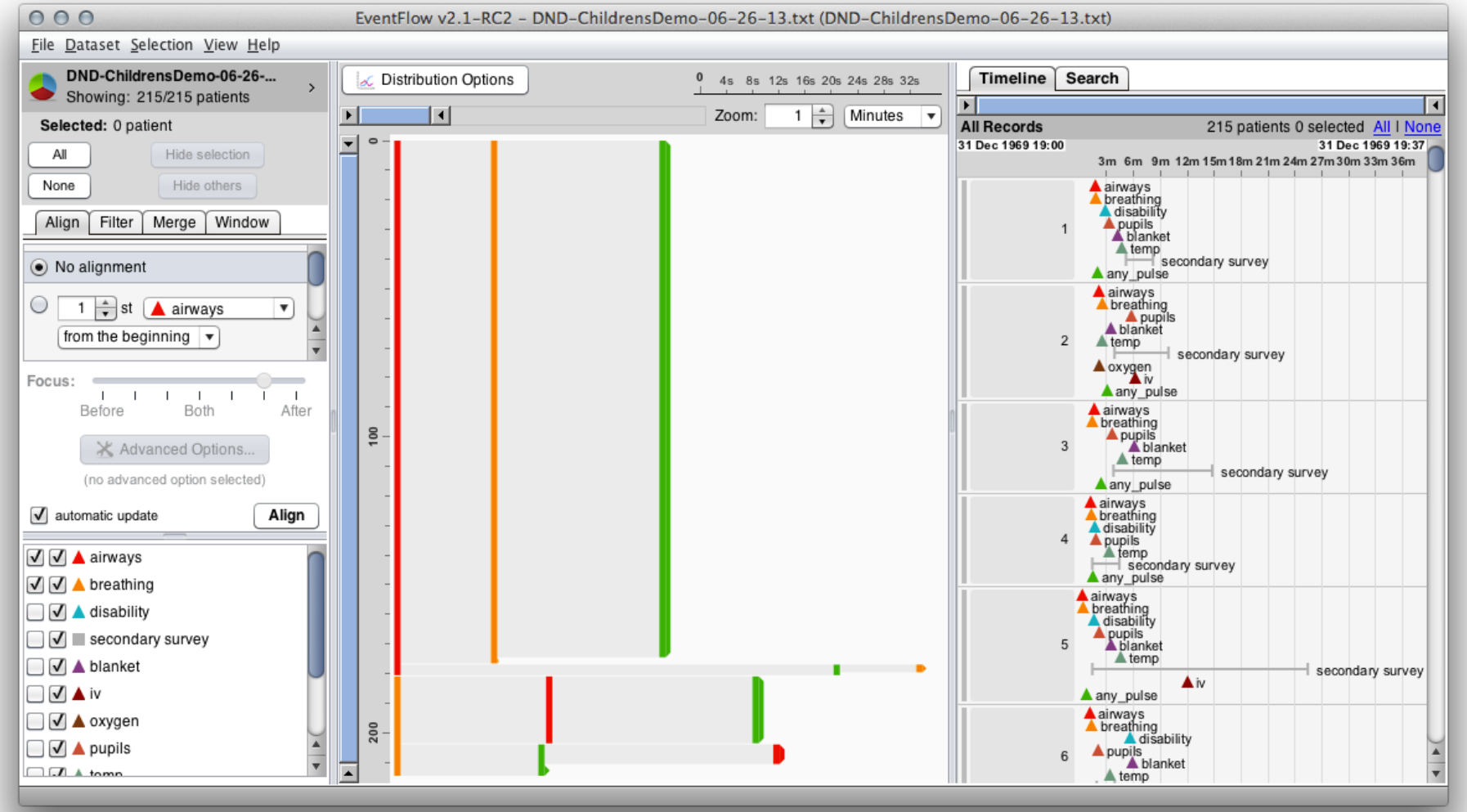
Combine the 2 pulse types



Graphical search & replace to remove duplicates

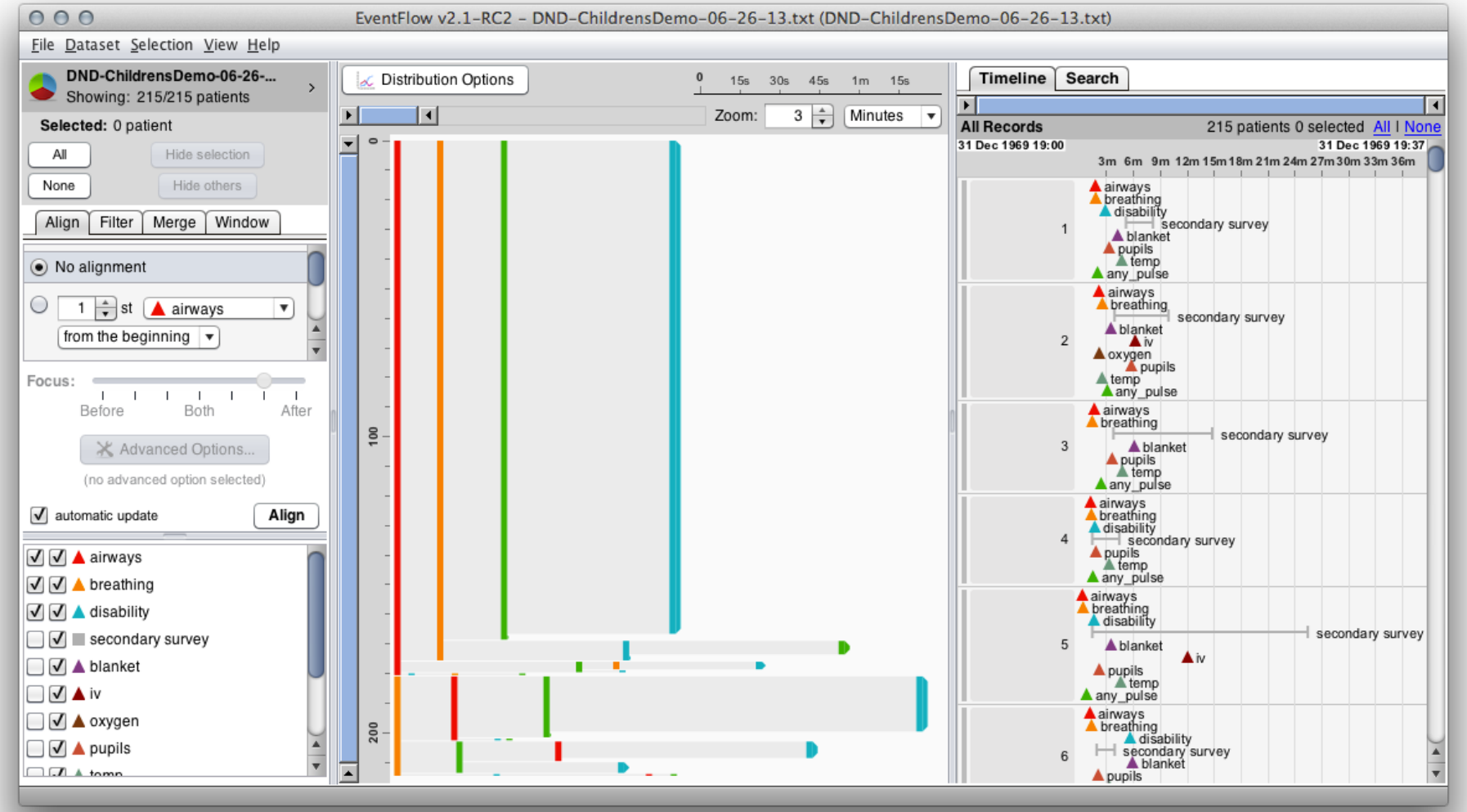


81% of the patients are treated in the correct order.
The largest deviation is still the airway and breathing being out of order,
but there are also instances where the circulation is checked too early.

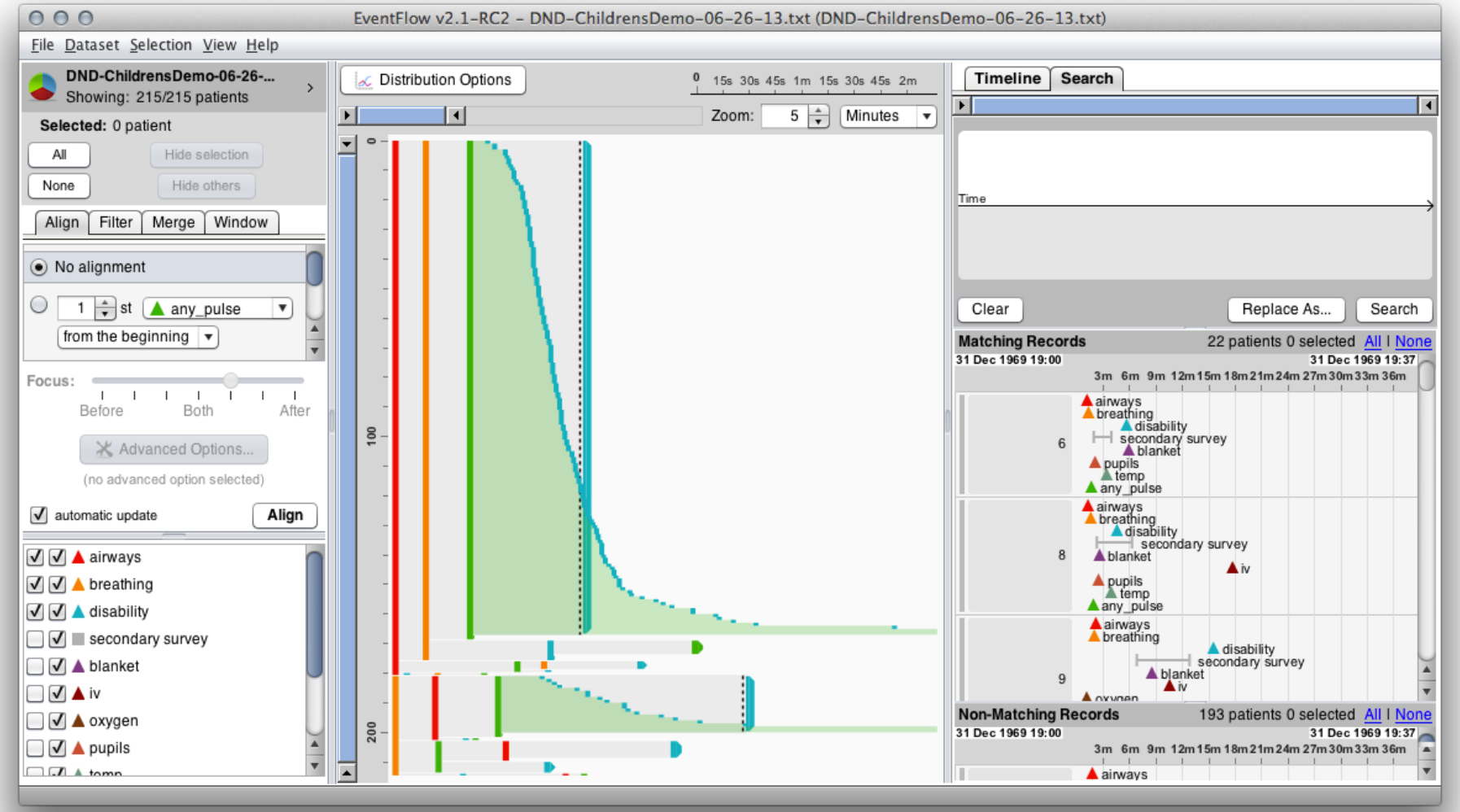


Add disability check

78% of patients treated in the correct order…

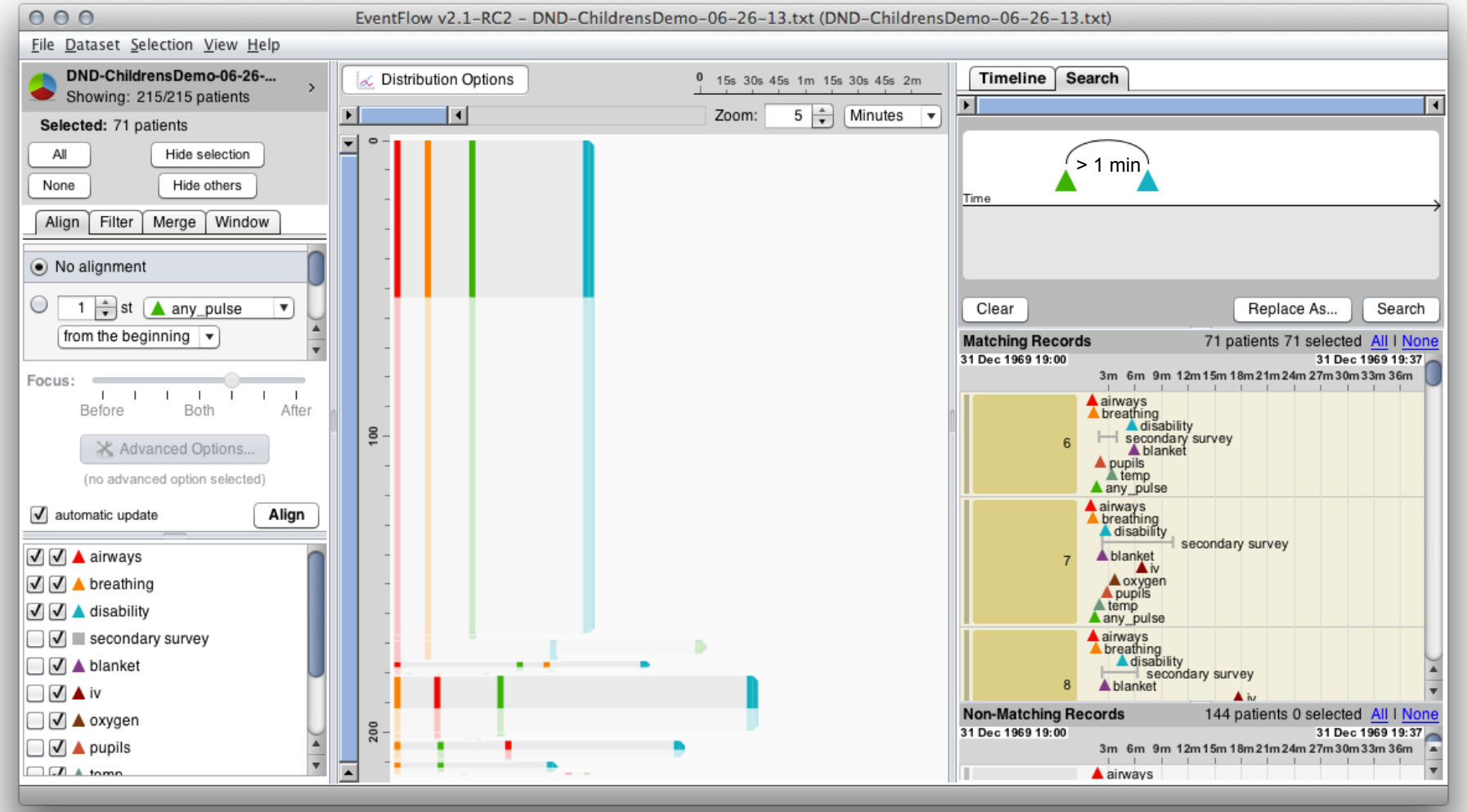


The average time between steps continues to be longer for the reversed airway and breathing check groups. The time between circulation and disability checks is about 1 minute for the correct sequence and 2 minutes and 15 seconds for the reversed order group.



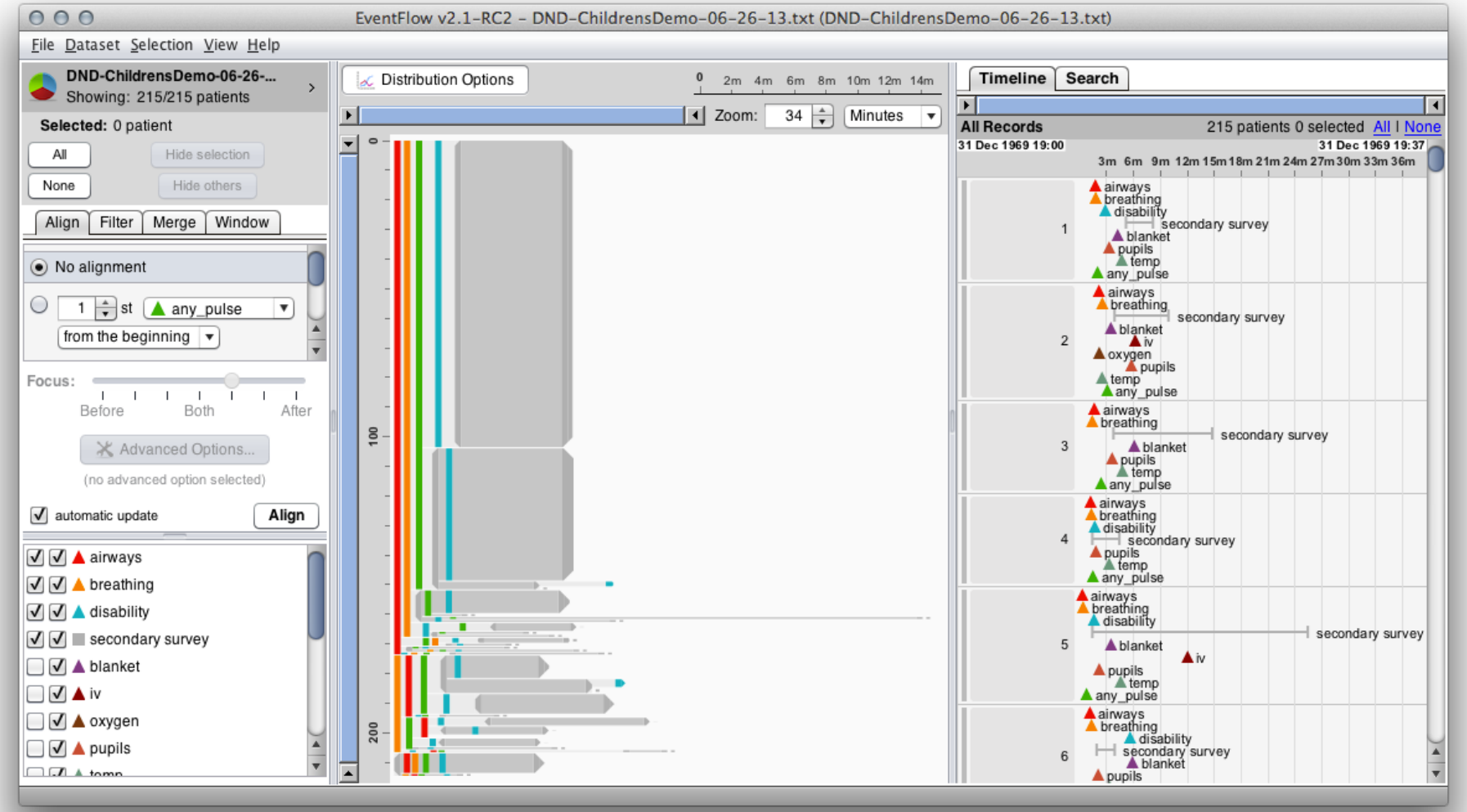
Searching with a time constraint.

See where check is taking longer than one minute



Add secondary survey

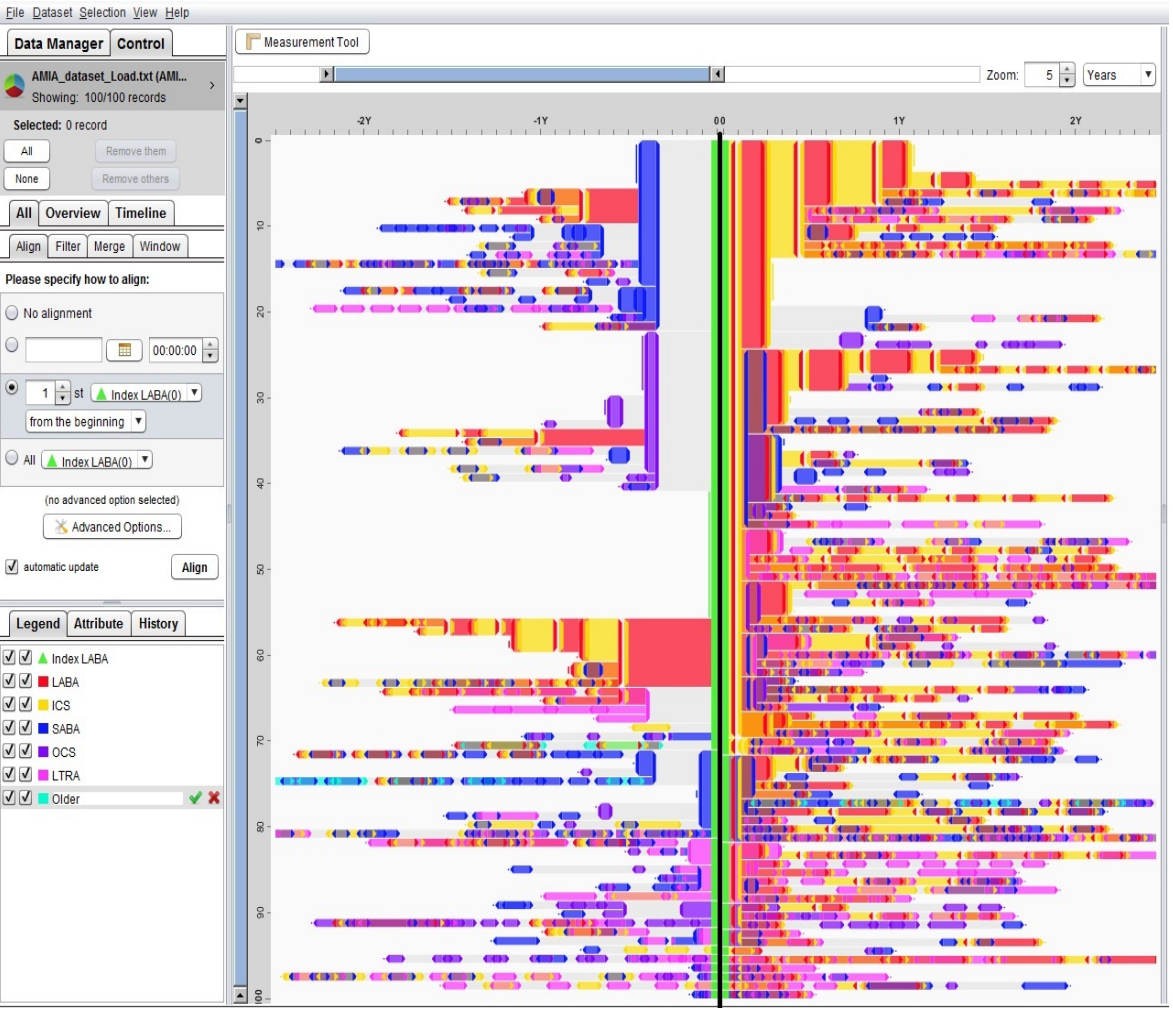
Correct procedure drops to only 48% of patients,

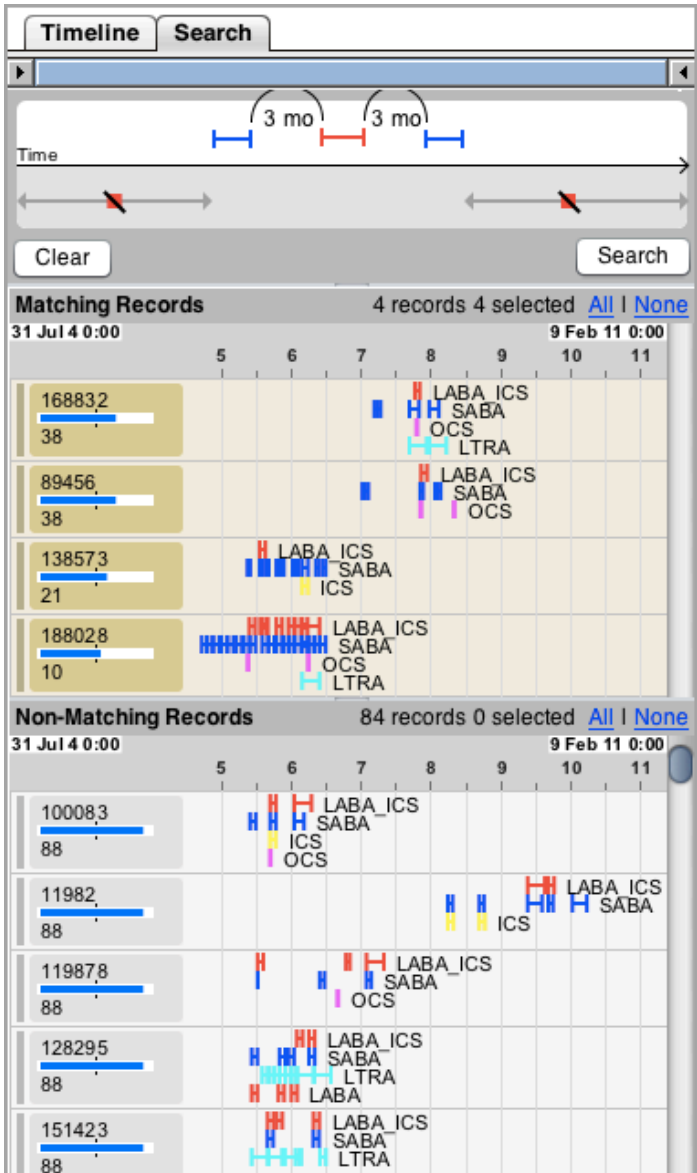


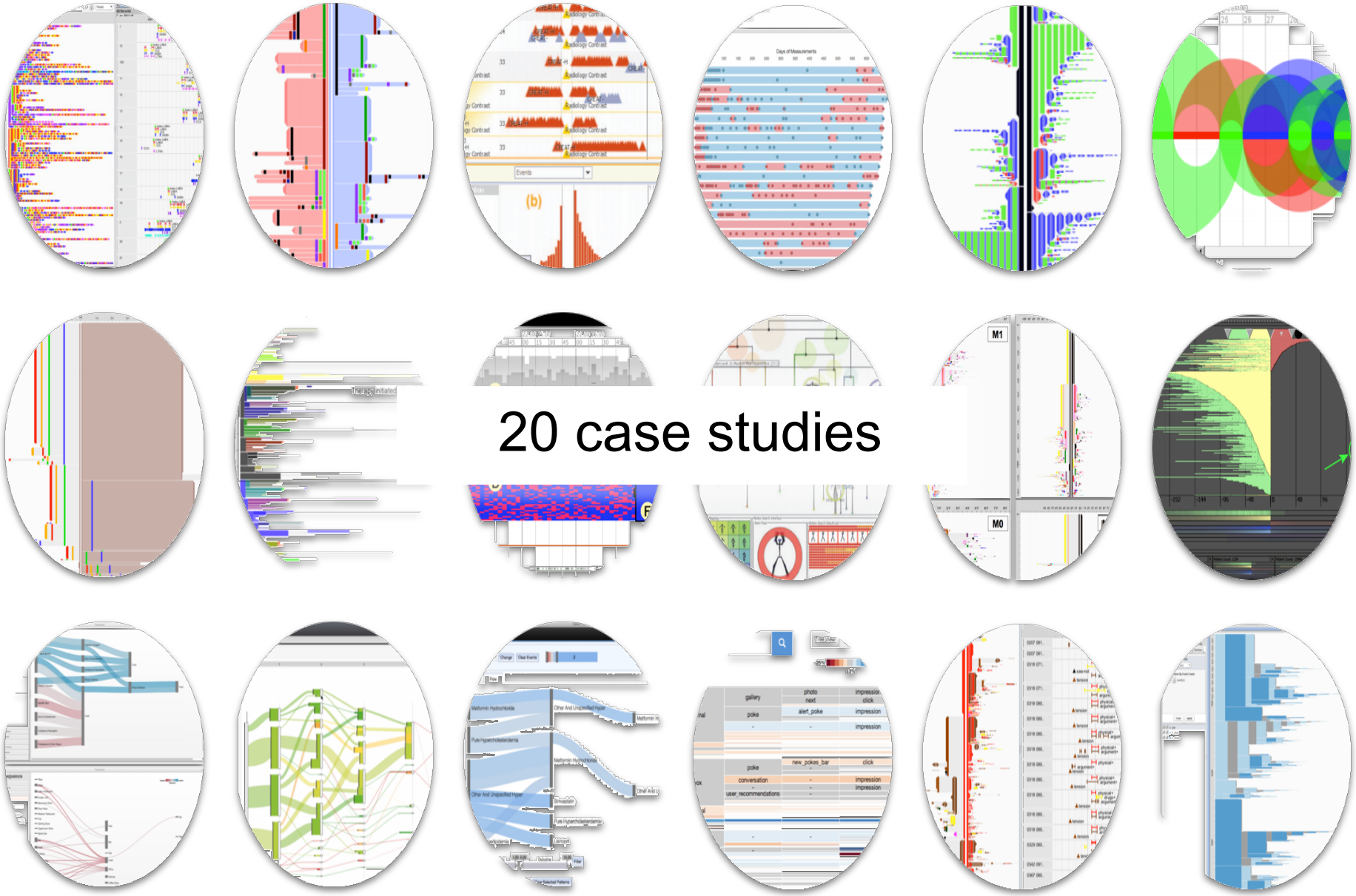
Other example

in collaboration with Army PharmacoVigilance Center

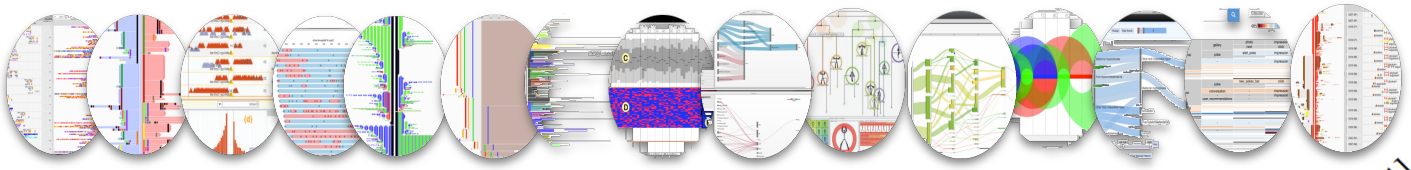
Analysis of prescription patterns of asthma medication






20 case studies

20 case studies



15 strategies

	ASTHMA [38]	BASKETBALL [40]	CONTRAST [59]	DEVIATION [14]	DRUG [39]	EHR COHORTS [19]	EPILEPSY [39]	FOURSQUARE [45]	HEART [44]	HYPERLIPIDEMIA [46]	HYPERTENSION [56]	LIVER [33]	LOG [39]	MEMuRY [32]	MOTION [12]	PROSTATE CANCER [43]	PROSTATE CANCER 2 [11]	TWITTER [62]	VIOLENCE [47]	WORKFLOW [39]
Extraction	●	●	●	●	●	●	●	●	●	●	●	●	●	●				●	●	●
S1: Goal-Driven Record Extracting	●	●		●	●		●	●	●	●	●	●	●			●	●	●		●
S2: Goal-Driven Event Category Extracting								●	●	●	●		●				●			
S3: Identifying Features Linked to Outcome																				
S4: Aligning	●	●	●	●	●		●	●	●	●		●	●		●	●	●	●	●	●
S5: Temporal Windowing					●	●	●	●	●	●	●	●	●	●	●	●	●	●		●
S6: Selecting Milestone Events								●	●	●			●							
S7: Random Sampling of Records																				
Folding	●				●															
S8: Temporal Folding		●	●					●	●					●					●	
Pattern Simplification	●			●	●		●	●	●	●	●	●	●		●	●	●	●	●	●
S9: Grouping Event Categories				●		●		●	●	●	●		●	●						●
S10: Coalescing Repeating Point Events into One	●				●							●	●							
S11: Coalescing Repeating Interval Events into 1	●			●	●					●										
S12: Converting Hidden Complex Events into 1																				
S13: Bucketing by Time Period																				
Iteration	●				●															
S14: Analyzing Small Subset then Larger One					●							●			●			●		
S15: Partitioning																				

Du et al. (2017) Coping with Volume and Variety in Temporal Event Sequences: Strategies for Sharpening Analytic Focus
IEEE Transactions Visualization and Computer Graphics

Event Analytics

Eventflow to:

- **review** the data from individual records
- **search** for temporal patterns of interest
- **summarize** all the event sequences
- perform **data transformations**
- **select cohorts** of interest for further studies

see also lots of other projects in event analytics
e.g. workshop at IEEE Vis 2016

Merci

Thank you to U. of Maryland colleagues

Ben Shneiderman

Fan Du, Sana Malik

Megan Monroe, Krist Wongsuphasawat, David Wang

+ all case study partners

plaisant@cs.umd.edu

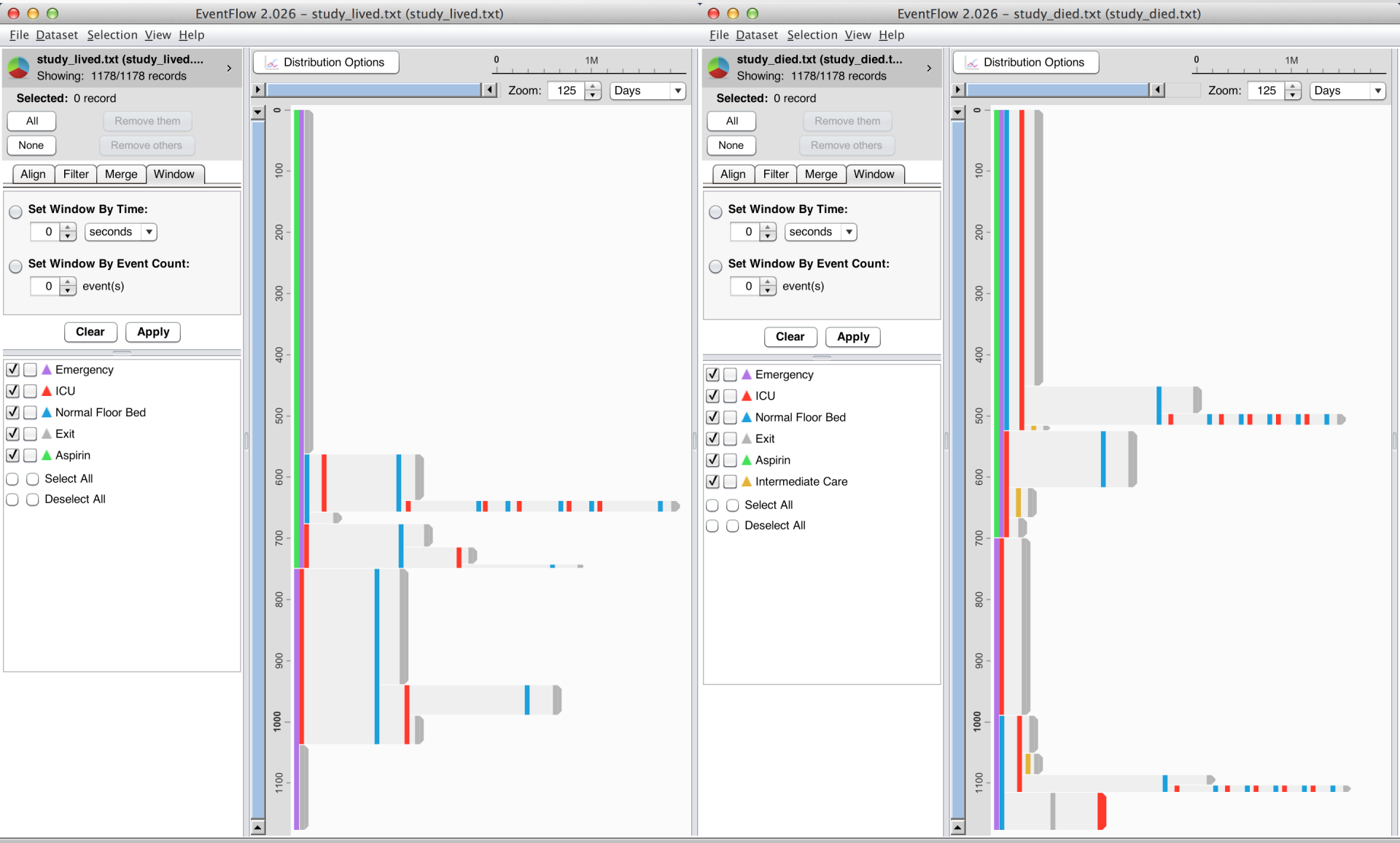
hcil.umd.edu/eventflow

Contact me for software access

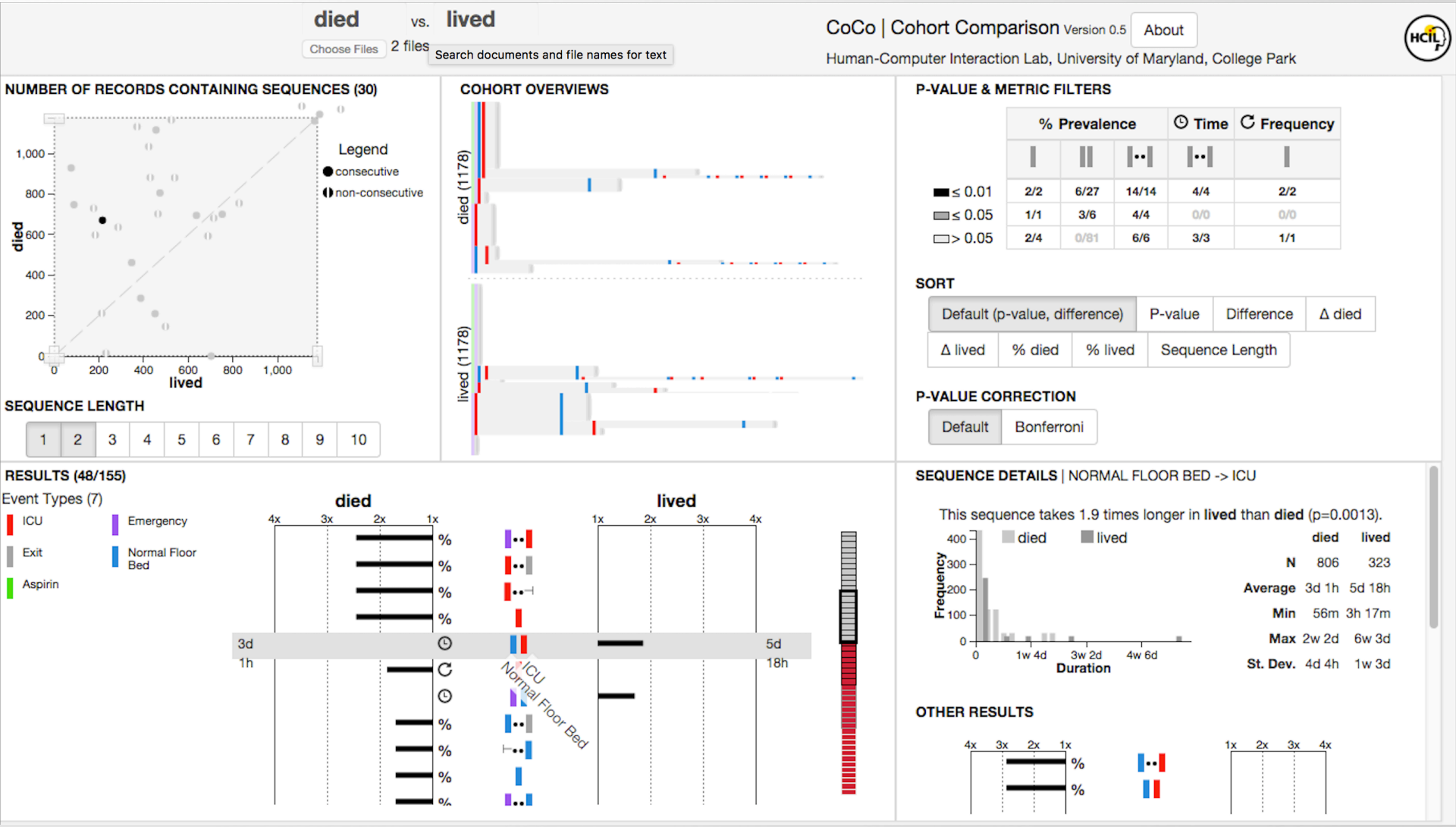
FYI: full day course June 30th

www.aviz.fr/DayCourse2017/EventAnalytics

Common question:
Comparison of 2 groups of records.
Two Eventflow side by side (hard to find differences)

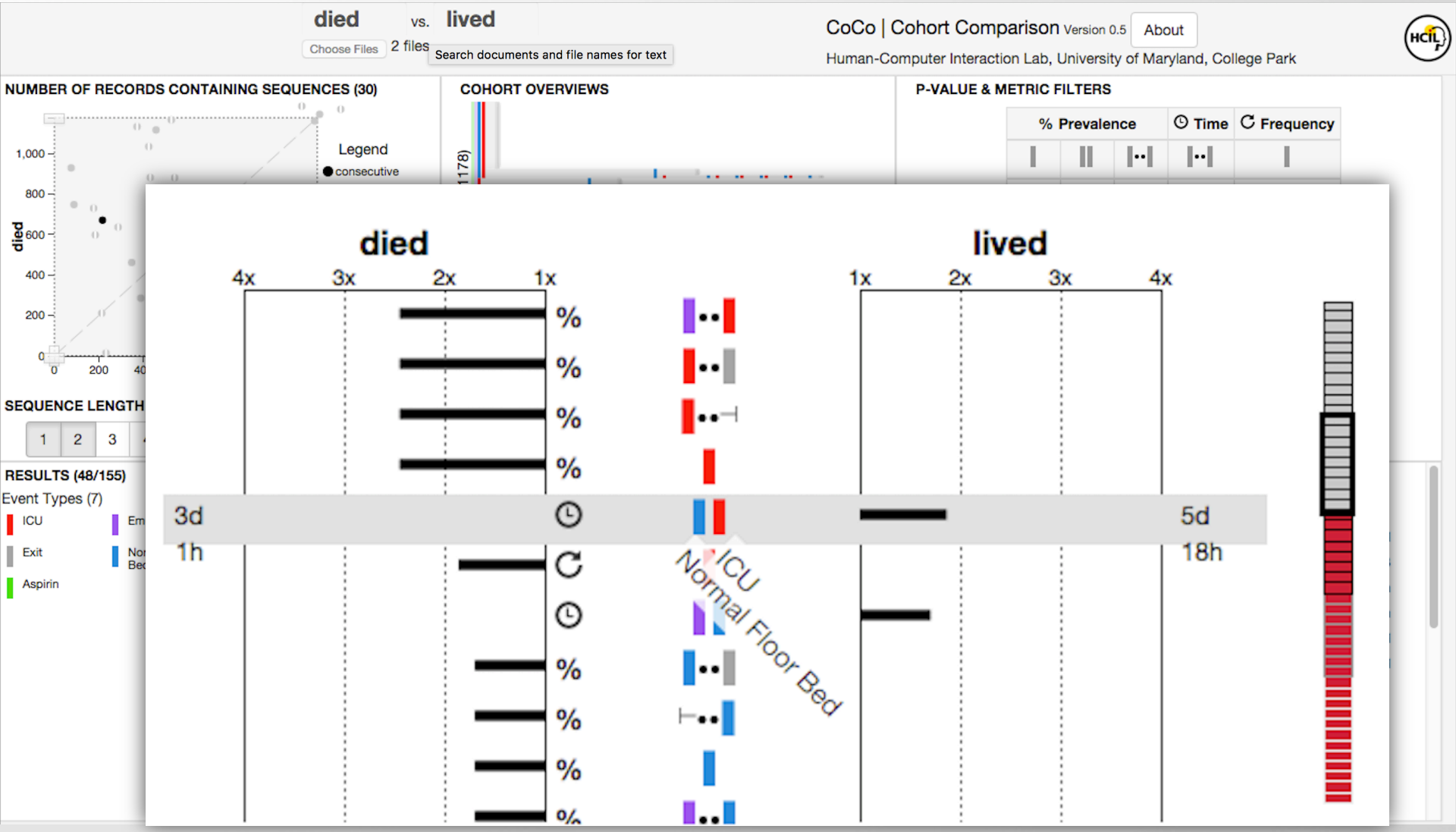


Cohort Comparison: Coco



www.cs.umd.edu/hcil/coco

Cohort Comparison: Coco



www.cs.umd.edu/hcil/coco