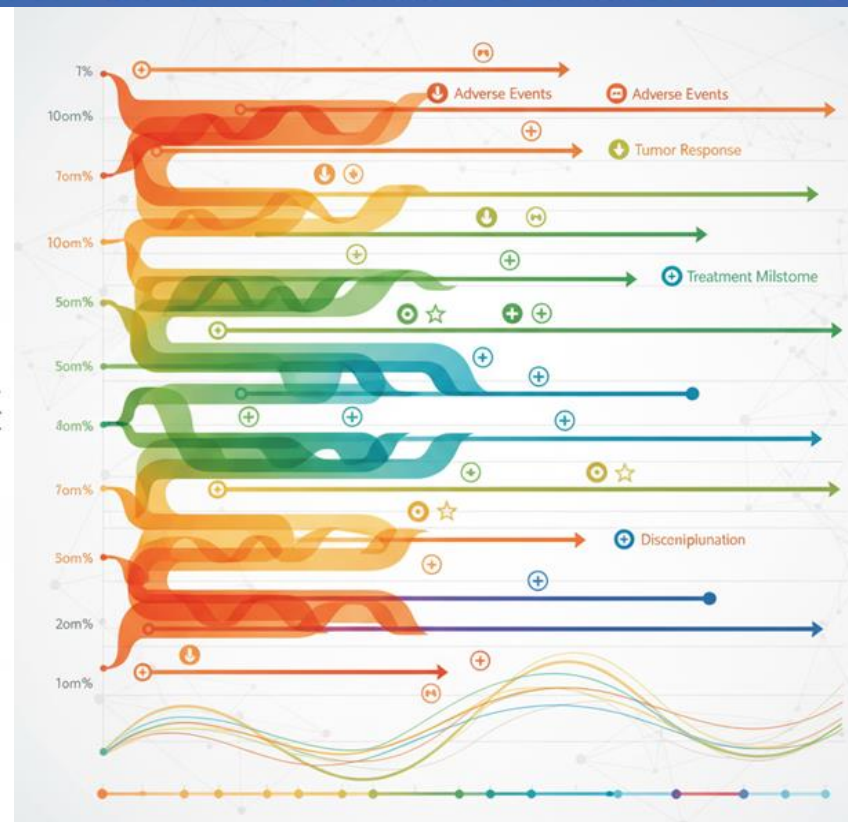
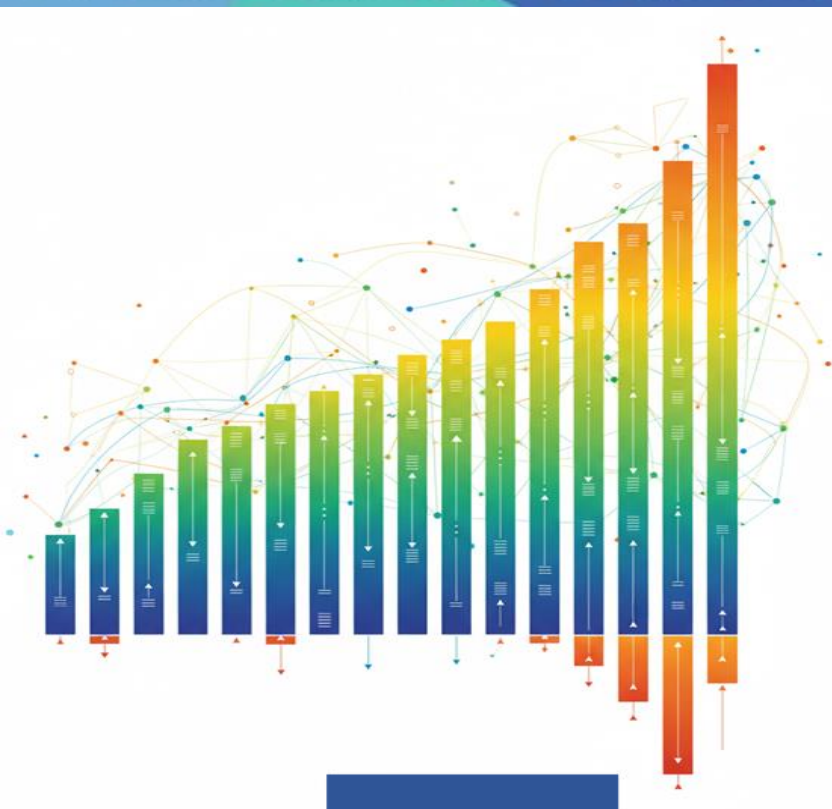




DV08: Visualizing the Invisible: Patient Journey with Waterfall, Spider and Swimmer Plots

Varsha Patil & Sunita Poojary, Mumbai, India

Agenda



Patient level visualization?



Why patient level visualizations matter?



Patient-level vs aggregate visualizations



Synthetic data for implementation of visuals



Waterfall plots



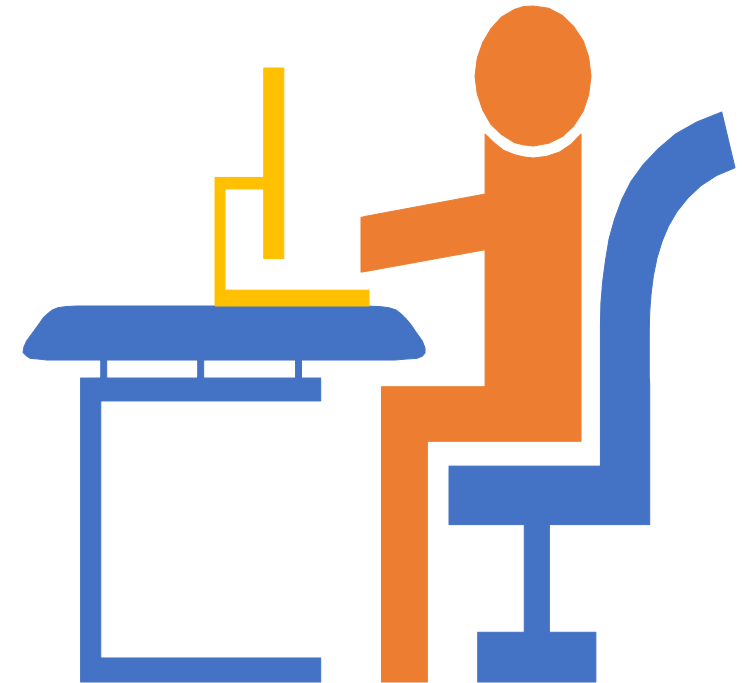
Spider plots



Swimmer plots



Full picture & summary





Patient-level visualization?



Patient-level visualization is a way of showing **data for each individual patient separately**, rather than summarizing everyone into averages or percentages. It focuses on **personal treatment journeys**, making it easier to see differences, patterns, and outliers that might be hidden in overall statistics.

Example

Imagine a cancer trial:

A summary chart might say:

“Average treatment duration = 12 months.”

A patient-level visualization (like a swimmer plot) would show:

Patient A: 8 months, response started at month 3, ongoing at cutoff.

Patient B: 20 months, response ended at month 15.

Patient C: 5 months, discontinued early..

Why Patient-Level Visualizations Matter?



Focus – Aggregate Statistics

Conclusion focused on Aggregate statistics – Masking Individual Variability



Listings

Historically, **listings** were used to review patient-level data, but they lacked visual clarity and interpretability.



Shift

Shift from listings and summary plots to **patient-level graphs** reflects a growing emphasis on **individual patient's trajectories**

In TA like **oncology and neurology**, patient responses are **highly heterogeneous**, making individual-level insights essential.

Patient-level visualizations offer a more intuitive and granular view of clinical data.



Advantages

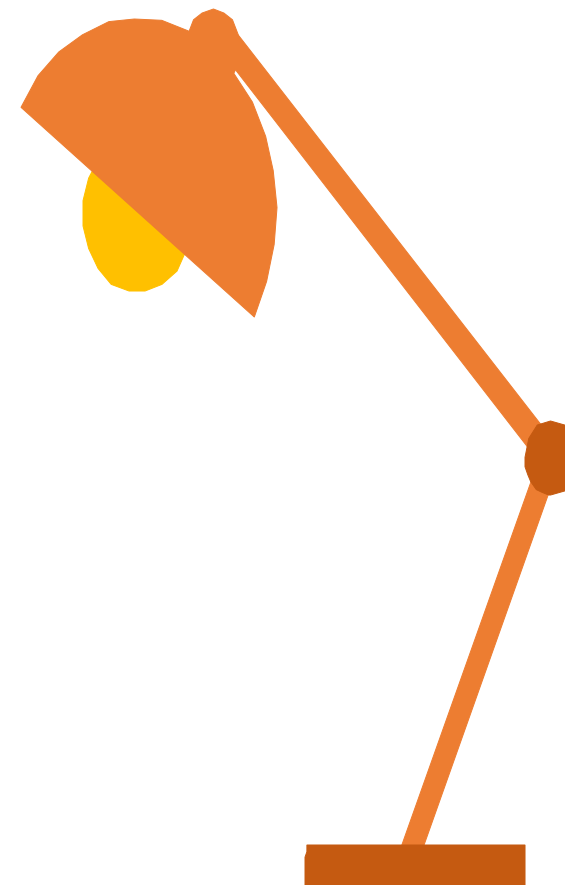
Patient-level plots (like waterfall, spider and swimmer plots) help:

- Detect outliers and subgroup patterns

- Visualize treatment dynamics over time

- Support personalized medicine and adaptive trial designs

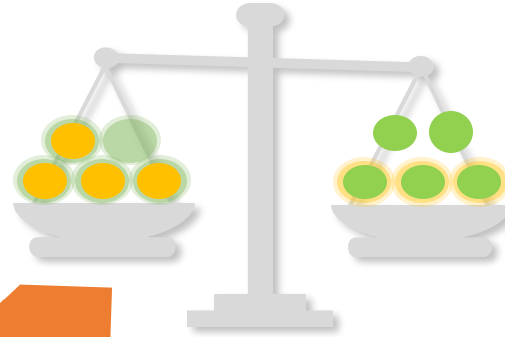
- Enhance regulatory submissions and clinical dashboards



MARS⁷

www.marsgroupinc.com
USA | INDIA | EUROPE

Patient-Level Vs Aggregate Visualizations



AGGREGATE PLOT

Mean Change

Median Survival

Summary Treatment
Duration

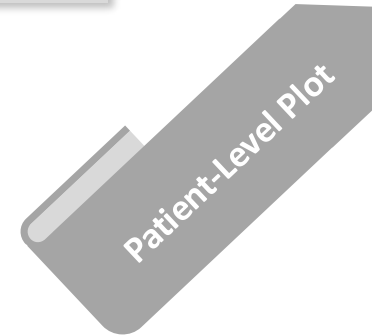


MISSING INSIGHTS

Individual response variability

Trajectory of each patient

Treatment milestones per
patient

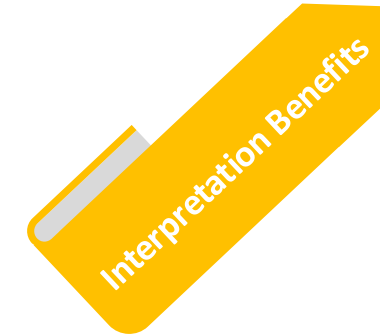


PATIENT-LEVEL PLOT

Waterfall Plot

Spider Plot

Swimmer Plot



INTERPRETATION BENEFITS

Identify responders vs non-responders

Visualize progression/regression over
time

Understand timelines and key events

Waterfall Plots



CONCEPTUAL OVERVIEW

WHY IT'S USED?

WHAT IT DISPLAY?

The “Ruler of Response”

- Bars showing %change in tumor size from baseline
- Each bar = one patient; height = tumor shrinkage or growth
- Sorted bars resemble a cascading waterfall

Heterogeneity (How responses differ between individuals)

- Assess treatment efficacy at individual level
- Identify responders vs non-responders
- Visualize variability in patient outcomes

Interpretation

X-axis: Each patient

Y-axis: % change in tumor size from baseline

- Bars below zero: Tumor shrinkage (Response)
- Bars above zero: Tumor growth (Progression)
- Optional color coding for response categories

Implementation with Data

Synthetic Data For Waterfall & Spider Plot

Patient	Visit	VisitDate	TumorSize	Percent Change	Treatment StartDate	Treatment EndDate	Treatment Durationweeks	Treatment Status	Treatment	OverallResponse
P001	0	21-01-2025	72.45	0	21-01-2025	03-06-2025	19	Ongoing	Trt B	CR
P001	4	18-02-2025	62.93	-13.15	21-01-2025	03-06-2025	19	Ongoing	Trt B	CR
P001	8	18-03-2025	52	-28.23	21-01-2025	03-06-2025	19	Ongoing	Trt B	CR
P001	12	15-04-2025	35	-44.38	21-01-2025	03-06-2025	19	Ongoing	Trt B	PR
P001	16	13-05-2025	25	-51.92	21-01-2025	03-06-2025	19	Ongoing	Trt B	PR
P001	20	10-06-2025	10	-71.43	21-01-2025	03-06-2025	19	Ongoing	Trt B	PR
P001	24	08-07-2025	0	-100	21-01-2025	03-06-2025	19	Ongoing	Trt B	CR
P006	0	23-01-2025	68.13	0	23-01-2025	11-12-2025	46	Completed	Trt A	SD
P006	4	20-02-2025	68.13	45.935	23-01-2025	11-12-2025	46	Completed	Trt A	PD
P006	8	20-03-2025	52	-23.74	23-01-2025	11-12-2025	46	Completed	Trt A	SD
P006	12	17-04-2025	45	-3.449	23-01-2025	11-12-2025	46	Completed	Trt A	SD
P006	16	15-05-2025	67	45.549	23-01-2025	11-12-2025	46	Completed	Trt A	PD
P006	20	12-06-2025	67	45.549	23-01-2025	11-12-2025	46	Completed	Trt A	PD
P006	24	10-07-2025	67.57	44.73	23-01-2025	11-12-2025	46	Completed	Trt A	PD

Note: Synthetic data aids visualization and privacy but lacks real-world complexity, so it is not clinical evidence.

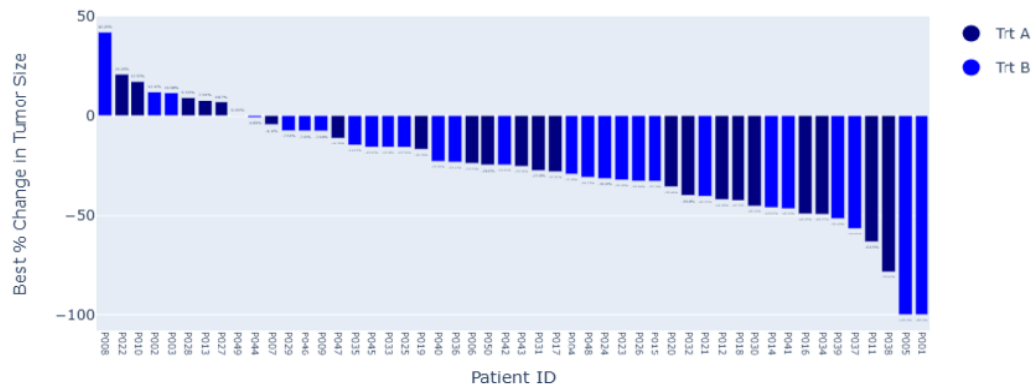
Key Variables Description

Variable Name	Description
Patient	Unique identifier for each patient
Visit	Week number of the clinical visit relative to treatment start
VisitDate	Actual calendar date of the visit
TumorSize	Measured size of the tumor (in mm)
PercentChange	Percentage change in tumor size from baseline (Week 0)
TreatmentStartDate	Date when treatment began
TreatmentEndDate	Date when treatment ended or was last recorded
TreatmentDurationWeeks	Total duration of treatment in weeks
TreatmentStatus	Indicates whether treatment is ongoing or completed or Discontinued.
Treatment	Type or arm of treatment assigned (e.g., Trt A, Trt B)
OverallResponse	Investigator-assessed response (CR, PR, SD, PD)

Waterfall Plot - Visuals

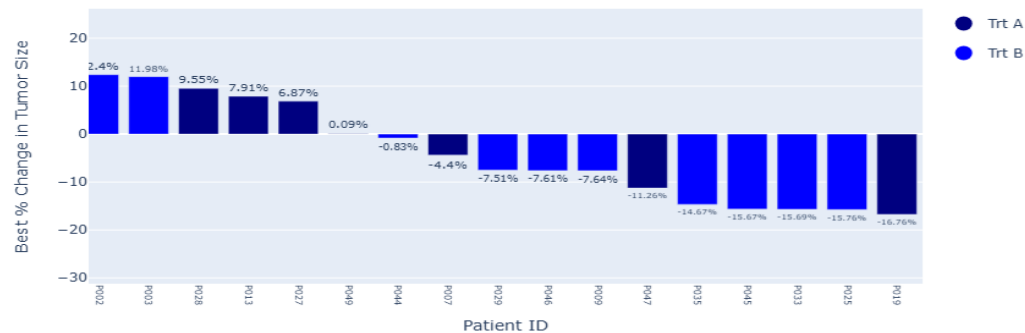
Overall view

Waterfall Plot of Best Percent Change per Patient by Treatment



Closed view

Waterfall Plot of Best Percent Change per Patient by Treatment



Interpretation

Majority of patients show tumor shrinkage, with several achieving deep reductions (up to $\sim -100\%$).



High variability

High variability in response across patients, indicating a heterogeneous population.



Mixed response

Both treatments (Trt A & Trt B) exhibit mixed responses, with no clear dominance. Mind further investigate on subgroups like Disease stages, Gender



Positive % changes

Presence of positive % changes (tumor growth) highlights cases of progression or resistance.



Extreme responders

Extreme responders (near -100%) indicate potential biomarkers worth investigating

Spider Plots



CONCEPTUAL OVERVIEW

WHY IT'S USED?

WHAT IT DISPLAY?

The “Trend Tracker”

- A line chart (also called spaghetti plot) used in oncology and longitudinal trials
- Each line = one patient, showing response trajectory over time
- Visual resembles a web or tangled spaghetti due to overlapping lines
- Useful for understanding patterns of response and variability among patients

Temporal Effect (How treatment impact evolves over time)

- Track individual patient responses over time
- Identify patterns of response and progression
- Compare treatment effects across subgroups
- Complement summary statistics with detailed longitudinal data
- Support regulatory submissions with patient-level insights

Interpretation

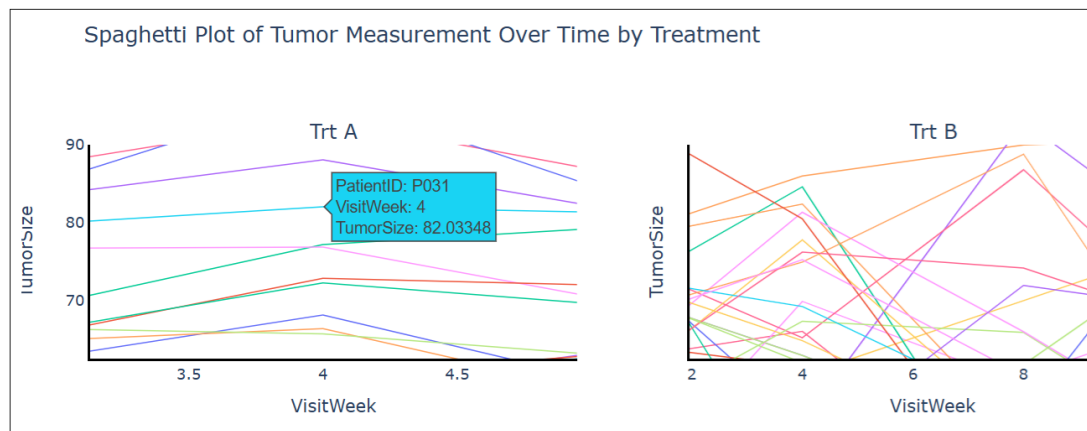
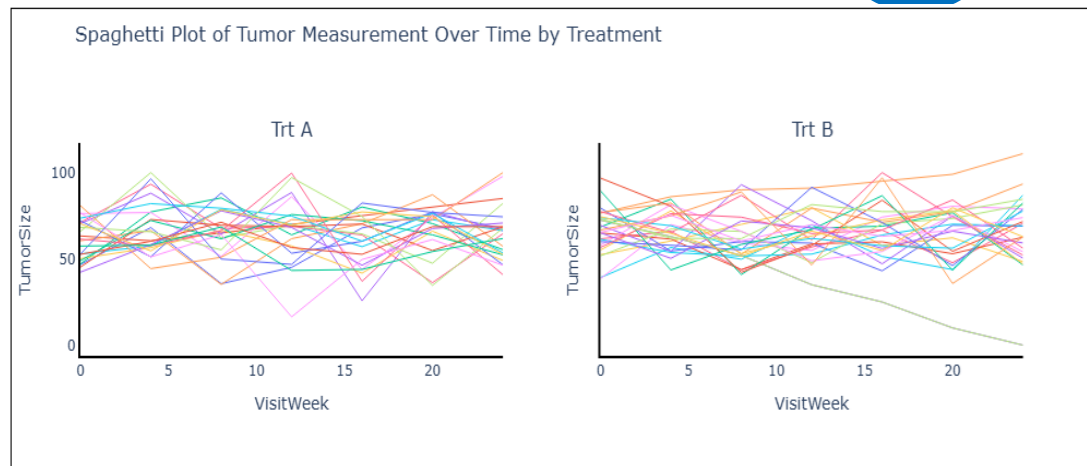
X-axis: Time (e.g., weeks or months since treatment start)

Y-axis: % change in tumor size (or another metric) from baseline

- Each line: Represents an individual patient's response trajectory
- Color coding (optional): Treatment arms, response categories or biomarker status
- Reference line at 0%: Indicates no change from baseline

Spider Plot - Visuals

Overall view



Interpretation

Spider/Spaghetti plot shows individual tumor size trajectories over time for Treatment A and B.



Substantial variability

Highlights substantial variability in responses, reflecting heterogeneous treatment effects. Variability may stem from biological, clinical, or demographic factors.



Change from baseline

Change from baseline can be used for better assessment of relative treatment impact.



Usefulness

Plot alone cannot guide clinical decisions; it must be complemented with summary statistics and inferential analyses.

Useful for identifying individual-level patterns and outliers that aggregate data may mask.

Swimmer Plots



CONCEPTUAL OVERVIEW

WHY IT'S USED?

WHAT IT DISPLAY?

The “Marathon tracker”

- Visualizes individual patient journeys in clinical trials.
- Each horizontal bar = one patient (like a swimming lane).
- Shows treatment duration, response timing and key events.
- Ideal for oncology trials to assess treatment consistency and variability.

Duration (How long patients remain on effective therapy)

- Track treatment exposure and clinical events per patient.
- Identify patterns (early discontinuation, prolonged benefit).
- Assess consistency across treatment arms.
- Complement summary statistics with patient-level detail.
- Support regulatory submissions with clear visual evidence

Interpretation

X-axis: Time (e.g,months, weeks, days).

Y-axis: Patients (Subject IDs).

- Bars: Represent treatment duration for each patient; often color-coded by categories such as disease stage, treatment arm or biomarker status.
- Markers: Indicate key clinical events, e.g.: Start of response, End of response, Ongoing response (shown with arrows)
- Annotations: May include: Duration of response (numeric values), Patient status labels (e.g., Ongoing, Completed, Discontinued)

Implementation with Data

continued.....

Synthetic Data

For swimmer Plot

Subject ID	Treatment_Start_Date	Treatment_End_Date	Treatment_Start_Month	Treatment_End_Month	Disease_Stage	BOR	Response_Start_Date	Response_End_Date	Response_Start_Month	Response_end_month	Status
S001	23-01-2022	11-07-2024	1	30	Stage III	PR	18-04-2022	04-06-2022	2	9	Discontinued
S002	27-02-2022	08-11-2025	1	45	Stage II	SD	null	null	null	null	Discontinued
S003	13-02-2022	29-12-2022	1	35	Stage II	CR	23-06-2022	29-07-2022	4	6	Ongoing
S004	05-02-2022	20-04-2025	1	39	Stage II	CR	14-06-2022	10-08-2022	4	null	Ongoing
S005	10-01-2022	29-04-2025	1	28	Stage III	PD	null	null	null	null	Ongoing
S006	10-01-2022	24-01-2025	1	37	Stage II	SD	null	null	null	null	Discontinued
S007	04-01-2022	22-06-2024	1	30	Stage I	PD	null	null	null	null	Completed
S008	21-02-2022	10-07-2024	1	29	Stage II	CR	19-06-2022	01-08-2022	3	null	Ongoing
S009	06-02-2022	26-04-2024	1	27	Stage III	SD	null	null	null	null	Completed
S010	12-02-2022	26-06-2025	1	41	Stage I	SD	null	null	null	null	Completed

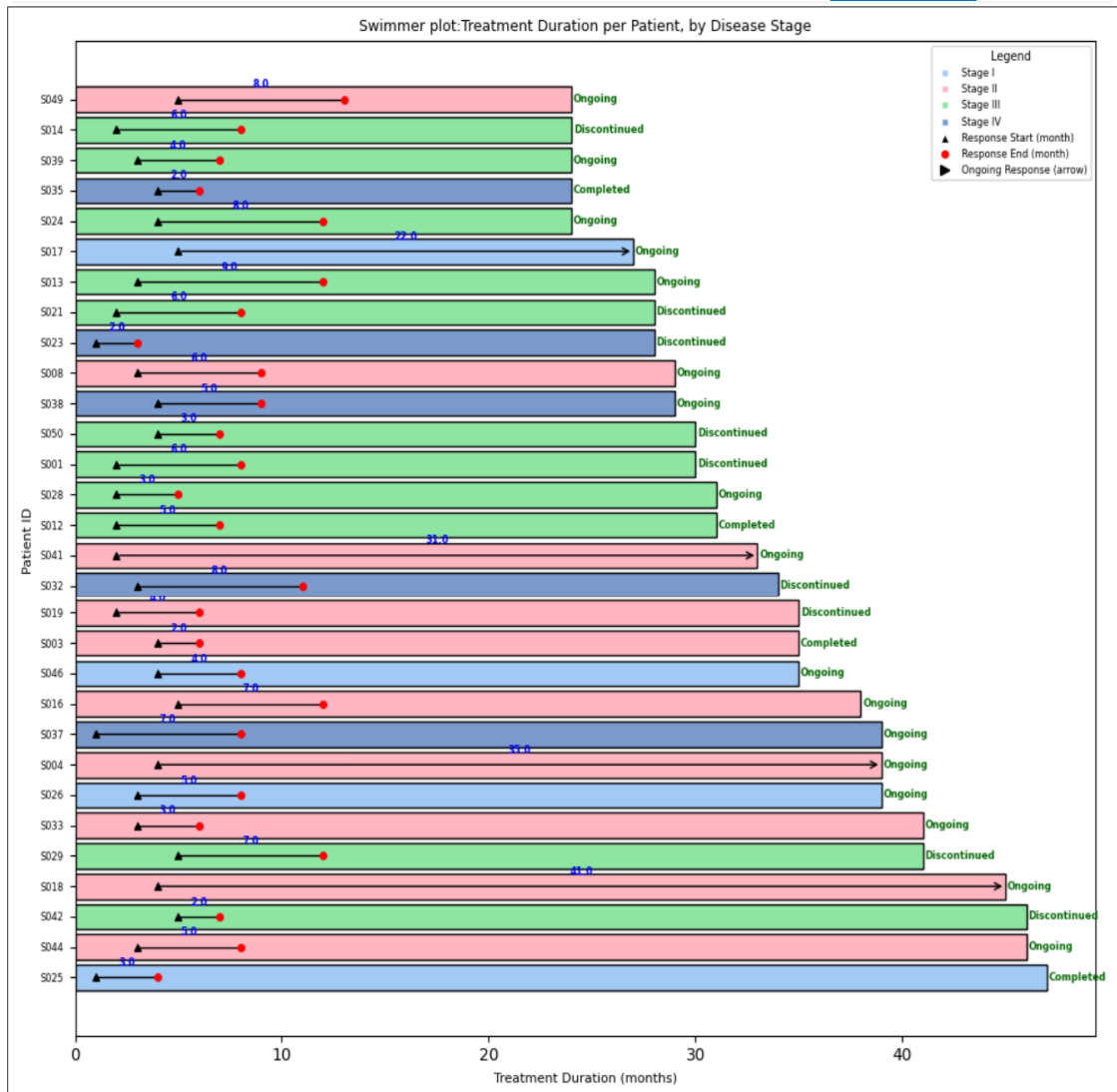
Note: Synthetic data aids visualization and privacy but lacks real-world complexity, so it is not clinical evidence.

Key Variables

Description

Variable Name	Description
SubjectID	Unique identifier for each patient
Treatment_Start_Date	Date when treatment began
Treatment_End_Date	Date when treatment ended
Treatment_Start_Month	Month number when treatment started
Treatment_End_Month	Month number when treatment ended
Disease_Stage	Stage of disease at baseline
BOR	Best overall response during treatment
Response_Start_Date	Date when response was first observed
Response_End_Date	Date when response ended
Response_Start_Month	Month number when response started
response_end_month	Month number when response ended
Status	Current status of the patient

Swimmer Plot - Visuals



Interpretation

Each horizontal bar represents treatment duration per patient, color-coded by disease stage (Stage I–IV).

Treatment duration varies widely across patients, ranging from a few months to over 40 months.



Responses

Ongoing responses (arrows) are observed in multiple patients, indicating continued benefit at data cutoff.

Start and end of response markers (black triangles and red circles) show variability in response timing.

Numeric annotations highlight duration of response, with some patients achieving long-term benefit (>20 months).

Patient status labels (Ongoing, Completed, Discontinued) reveal mixed treatment outcomes.



Observation

Stage III and IV patients appear more frequently, suggesting higher representation of advanced disease.

Note: This example serves an illustrative purpose only, and conclusions drawn from synthetic data should not be interpreted as clinical evidence.

Python Libraries Behind the Visuals



Aspect	Waterfall Plot	Spider Plot	Swimmer Plot
Library / Package Used	Pandas, plotly.graph_objects, kaleido	Pandas, plotly.graph_objects, plotly.subplots	Pyspark.sql.functions, pandas, matplotlib.pyplot, seaborn
Type of Plot	Dynamic (interactive bar chart)	Dynamic (interactive line plot)	Static (annotated horizontal bars with markers and arrows)
Reason for Package Choice	Plotly for interactivity; Kaleido for export	Plotly for multi-panel and hover details	Matplotlib for layout control; Seaborn for color
Limitations for Dynamic Use	Manual layering for labels and legends	Subplot and legend customization is complex	No native support for arrows or mixed markers

The Full Picture: A Unified Patient Story



Waterfall answers

Was the drug effective at its peak?

Swimmer answers

For how long did patients benefit before an event occurred?

Spider answers

How quickly did the effect manifest and change over time?.



Integration of Efficacy, Kinetics and Duration

We move from three individual snapshots to one coherent narrative.

MARS⁷

www.marsgroupinc.com
USA | INDIA | EUROPE

Summary



Patient-level visualizations (swimmer, waterfall, spider plots) reveal individual treatment journeys beyond aggregate summaries.



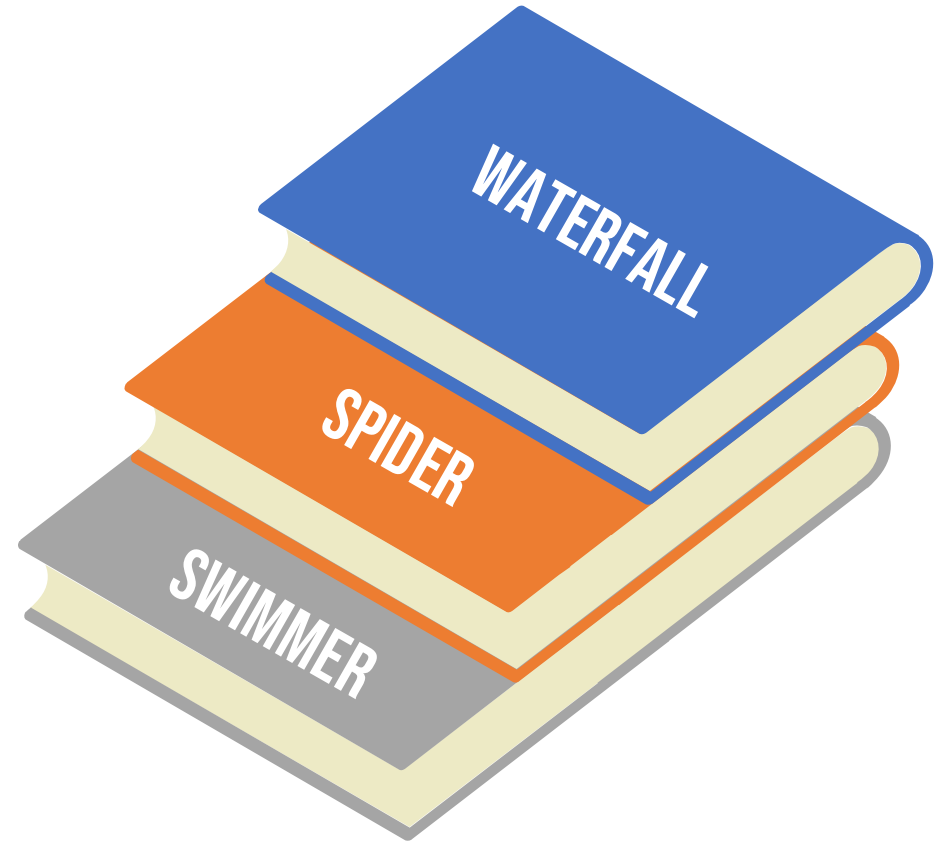
These plots highlight variability, response dynamics and outliers, improving clinical interpretation.



Synthetic data was used for demonstration; real-world data adds complexity (irregular patterns, missing values, adverse events).



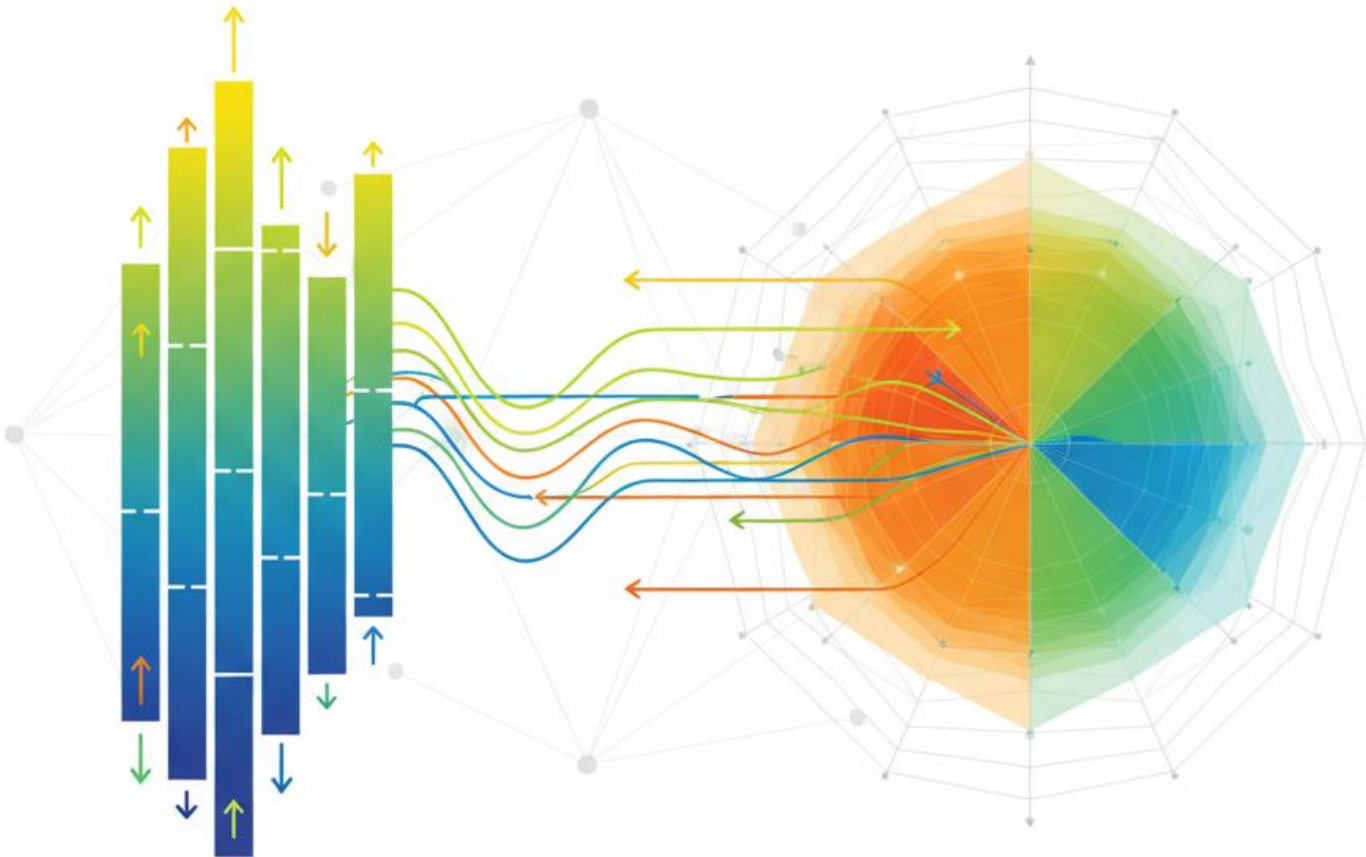
These visualizations should complement not replace, aggregate statistics for robust conclusions.





ANY
QUESTIONS?

Thank You!



Varsha Patil
Varshapatil@marsgroupinc.com