SANKEY DIAGRAMS TO TENDRIL PLOTS IN R : AN AID FOR CLINICAL DECISION MAKING

Najma Ashraf

Biostatistician, Novartis Healthcare Pvt Limited, Hyderabad, India

IASCT
Indian Association for Statistics in Clinical Trials

Introduction

Visualizations are important in providing key insights into the data. Graphical techniques are routinely employed in clinical studies as they are more easier to interpret than reading from tables. R is known to be a really powerful programming language when it comes to graphics and visualizations. Towards this, we utilize the myriad functionalities of R packages to create plots to aid in clinical decision making process.

We review the following plots, its usefulness and limitations :

- **Sankey diagrams** helps to study flows. Here, the nodes are represented by rectangles and they are connected by links. The width of the links is proportional to the rate of flow.
- **Spider plots** are useful for studying the changes in a continuous response over time per subject relative to the baseline value measurement.
- **Subpopulation Treatment Effect Pattern Plots(STEPP)** for studying treatment effects in overlapping patient subgroups defined by a continuous variable.
- **Tendril plots** for visualizing the incidence and time course of adverse event data. Here, each tendril represents a Medra PT and each point is an event. The distance between two any two points in a tendril is proportional to time between events. The tendril will tilt to right for an event occurring in placebo arm and to the left for an event occurring in active arm

Methods

Sankey Diagram
- Package used : *networkD3*
- Input data consists of nodes which corresponds to the treatments. The data is given as source node, target node and value which is the frequency count of patients from source to target node

Spider Plot
- Package used : *ggplot2*, *tumgr* and *plotly*
- Data consists of multiple observations for each subject at each time point with percentage change of tumor size from baseline
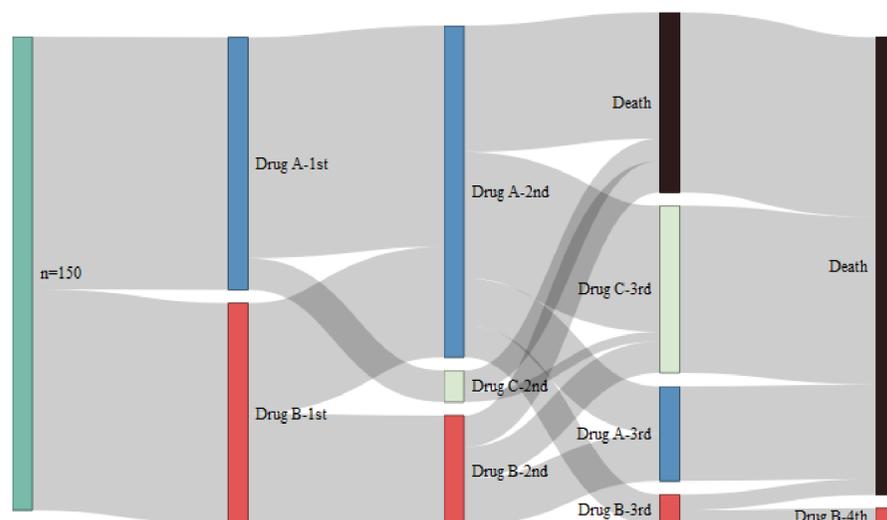
STEPP plot
- Package used : *subgrPlots*
- Data consisting of survival time, event status, treatment arm and age based on which subgroups are created
- The parameters, r_1(Subgroup sample size) is set to 40 and r_2(Overlap size) is set to 20
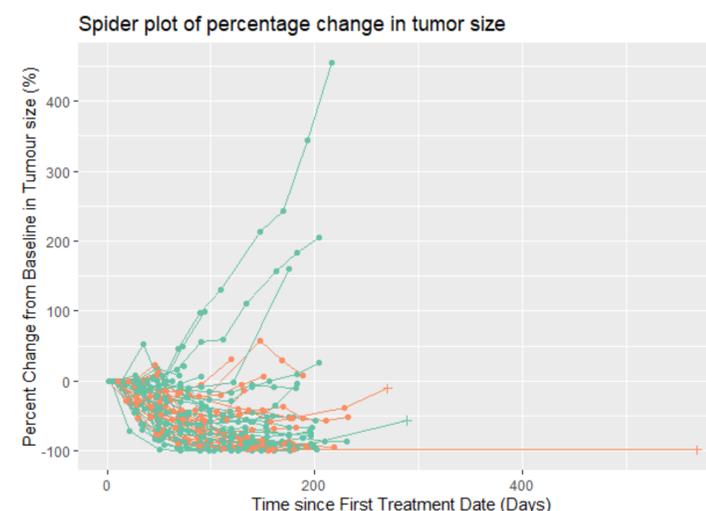
Tendril Plot
- Package used : *Tendril*
- Input dataset with treatment arm, number of days since treatment start and the preferred term
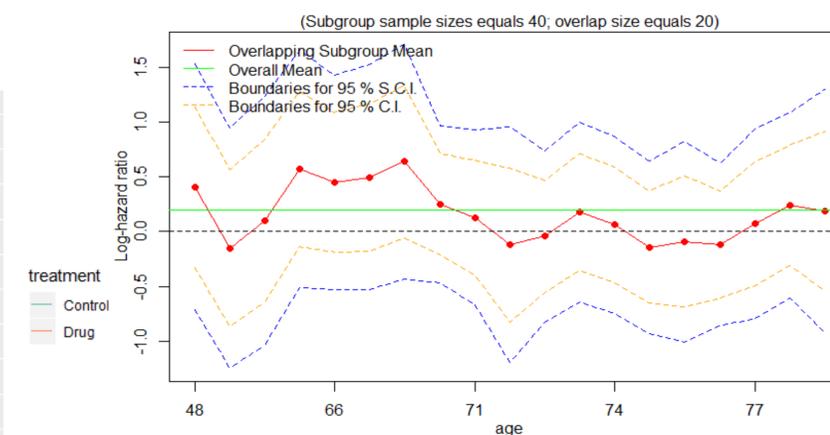
Results

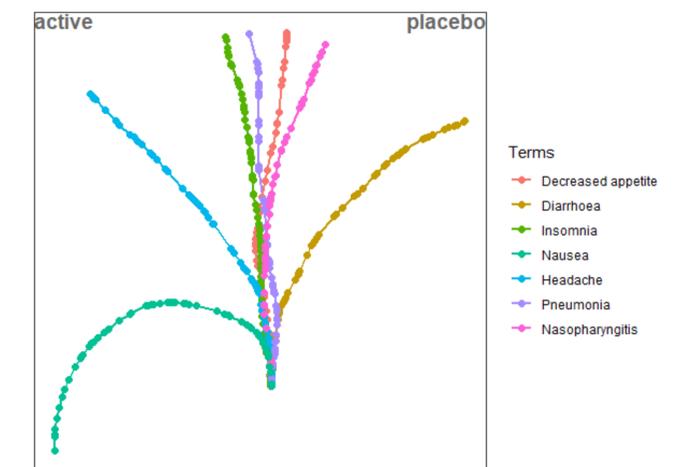**A.Sankey Diagram : To visualize treatment switch pattern**

B. Spider Plot : To study pattern of tumor response over time

C.STEPP plot :To identify heterogeneity in treatment groups

D.Tendril plot :To visualize incidence and time course of AE

References

1.Karpefors, Martin, and James Weatherall. *Journal of the American Medical Informatics Association* 25.8 (2018): 1069-1073.

2. Yip, Wai-Ki, et al. *Clinical Trials* 13.4 (2016): 382-390.

3. Mercier, Francois, et al. *Pharmaceutical statistics* 18.5 (2019): 526-532

4. Okita, Kazutaka, et al. *Clinical genitourinary cancer* 18.2 (2020): e103-e111