Paper DV01

ResponDEX – Response Data Exploration Dashboard

Sevak Kamalyan, Armat Analytics, Yerevan, Armenia

Alexander Mkrtchyan, Armat Analytics, Yerevan, Armenia

ABSTRACT

This paper introduces ResponDEX, an interactive R Shiny-based dashboard designed to facilitate comprehensive analysis of solid tumor response using RECIST criteria and industry-accepted clinical endpoints. Built on SDTM datasets, ResponDEX provides seamless access to essential metrics such as Best Overall Response (BOR), Objective Response Rate (ORR), Disease Control Rate (DCR), Progression-Free Survival (PFS), Overall Survival (OS), Duration of Response (DoR) and many others. The dashboard features a rich selection of interactive visualizations, including Kaplan-Meier survival curves, Tumor Diameter Change Waterfall Plots, Swimmer Plots, and Forest Plots for Hazard Ratios. Users can explore tumor progression over time with an innovative human body map visualization, dynamically tracking lesion locations and changes. Additionally, ResponDEX offers a comprehensive set of summary tables and listings, providing structured insights into patient response and study outcomes. All components are interconnected, allowing for intuitive filtering, drilldowns, and parameter adjustments to support flexible and exploratory analysis.

THE INITIAL VISION/GOALS

The main trigger for creating this system was the fact that a large portion of reported endpoints for studies within the same therapeutic area or disease are the same. This means there are common analysis methodologies shared by such studies, which makes it possible to automate the generation of such endpoints and create efficiency by removing the gap between generating SDTM-mapped datasets and having the first report on our study.

The system is designed with the following goals in mind:

- The system should cover a wide variety of endpoints that are common for specific therapeutic areas/diseases.
- The system should be flexible enough to accommodate requirements imposed by a specific study (e.g., censoring rules, imputation rules, etc.).
- The system should operate using SDTM data as input.
- The system should be scalable, meaning adding new types of endpoints or therapeutic areas/diseases should be feasible.
- The dashboard should be highly interactive and informative, providing insights in a clear and concise manner.

ARCHITECTURE

All the imposed requirements are addressed with the following system architecture.

Main Modules of the System:

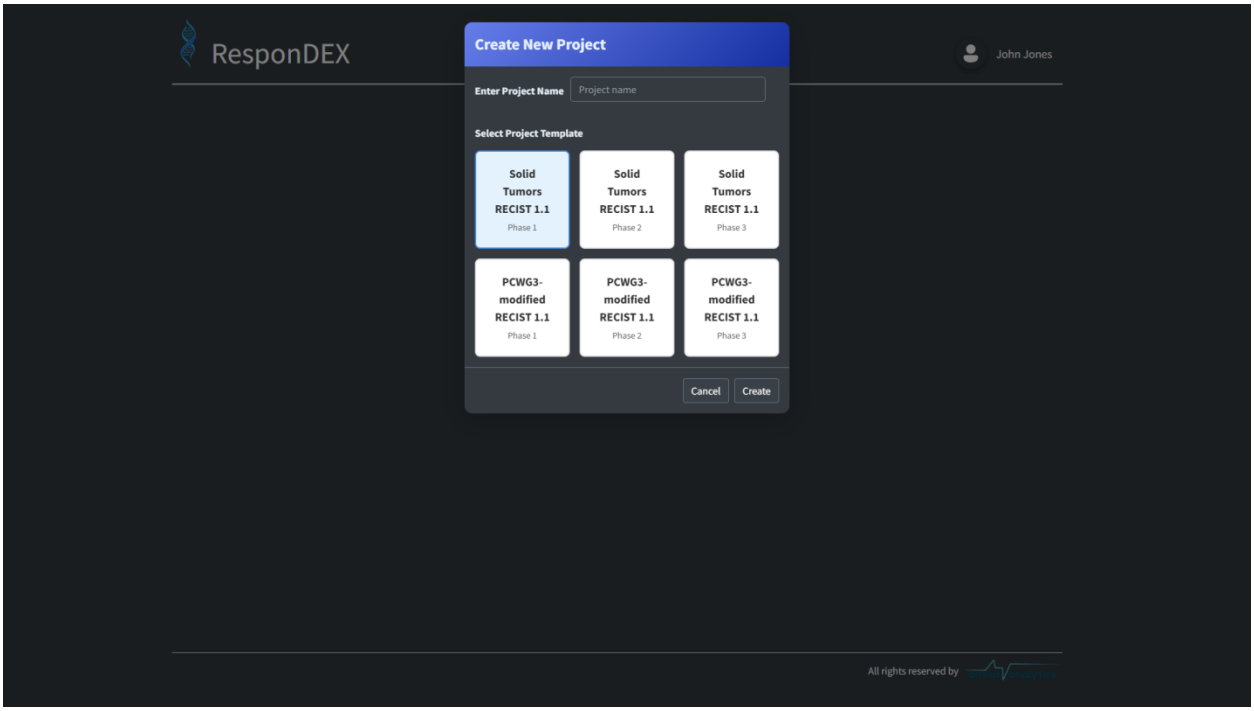
- System UI
- Data Processing Engine
- Dashboard Engine

SYSTEM UI

The System UI is responsible for the initial setup of the project and further navigation through it. One of the main responsibilities of the System UI is to capture input information from the user that dictates the analysis required and the design of the dashboard. This is done by introducing the concepts of study templates and study endpoints.

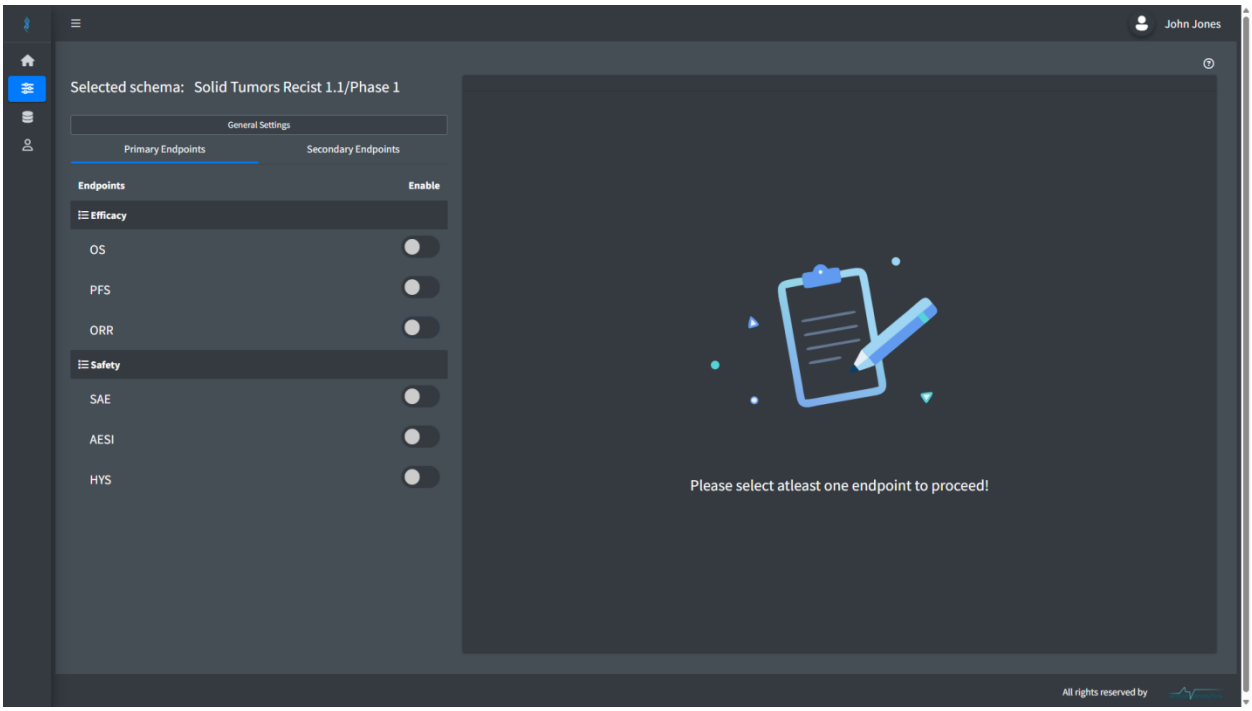
A template is a set of endpoints that is common for a specific type of study. For example:

- Solid Tumor RECIST 1.1
- Prostate Cancer (RECIST + PCWG)
- Custom

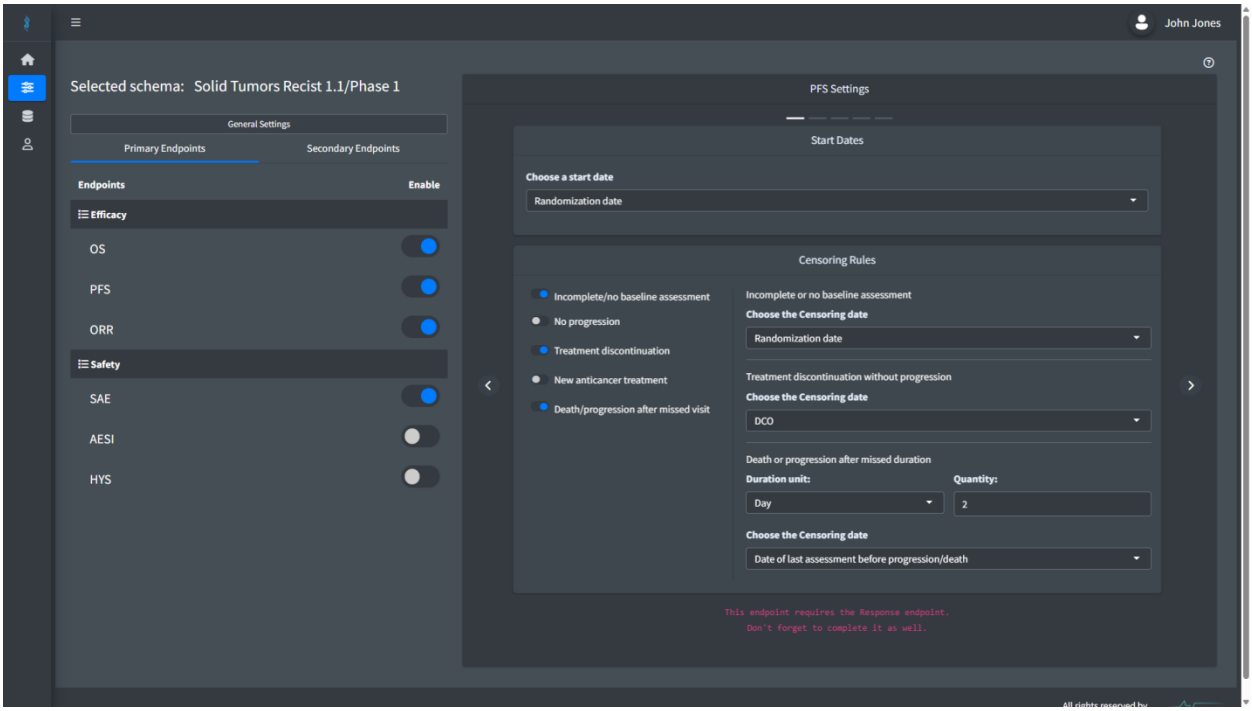


By selecting a template, the user effectively selects the pool of endpoints that will be available for designing the dashboard.

The available endpoints are used by the user to design the study analysis. The user can select any of the available endpoints and assign them to the corresponding category (e.g., Primary Efficacy, Primary Safety, Secondary Efficacy, etc.).



For each endpoint, the user maps source variables and rules that must be applied to them (e.g., “Censoring rules”).

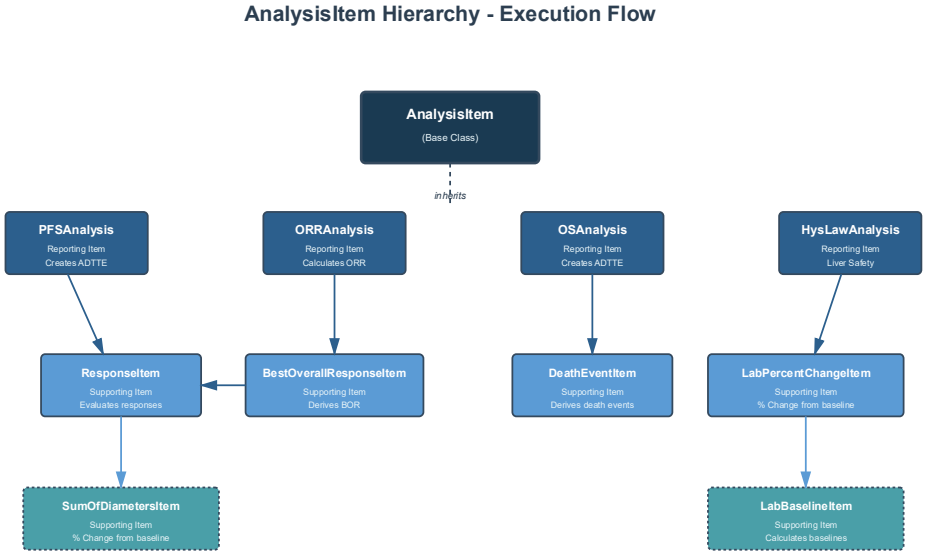


Once all the necessary input is collected, the Data Processing Engine comes into play.

DATA PROCESSING ENGINE

The ResponDEX Processing Engine is built on an object-oriented design that enables any analysis to be executed through a unified model. Based on user-selected endpoints and data mapping specifications, a set of Analysis Items is created. Each Analysis Item acts as a contract between the user interface and the analytical engine.

The main concept that ensures the high scalability of the system is the idea that an Analysis Item is self-sufficient - it knows how and what it should calculate. Another key concept is that each Analysis Item can trigger the generation of child Analysis Items (e.g., PFS triggers Response derivation), which creates a tree of Analysis Items that execute hierarchically, each contributing its input to the development of the final ADaM package.



This design ensures that a theoretically infinite number of new endpoints, custom logics, study templates, and therapeutic areas can be added to the system without the need to update the main engine. Each new element is a self-sufficient module that connects to the system based on predefined policies.

DASHBOARD ENGINE

The ResponDEX dashboard is the interactive layer that transforms processed data into exploratory analytics. Built in R Shiny, it provides both project-level and subject-level perspectives, integrating multiple visualization types into a single analytical environment.

The main concepts that are fundamental in the design of the Dashboard Engine are:

- Both project-level and subject-level visualizations
- Highly interactive design, ensuring:
    - Dynamic filtering
    - Transfers from project-level dashboard to the study level (e.g., by clicking on a subject in a swimmer plot, you transfer to the dashboard of that subject)
    - Interdependence of context (filtering one visualization triggers filtering in others where appropriate)

Each line, bar, and point in visualizations has some type of information connected to it. For example: treatment arm, subject ID. By clicking on a corresponding element, the desired information is taken from one visualization and used to filter the datasets that other visualizations depend on. Apart from just filtering, any other operation can be connected to the click event, and one of them is to switch from one tab to another, such as the subject-level dashboard tab.

One of the main visualizations at the subject-level dashboard - and one that is particularly interesting from mathematical, algorithmic, and visualization technique perspectives - is the Human Body Map.

The paper will further discuss in detail how the Human Body Map is designed, along with its requirements and main challenges.

HUMAN BODY MAP

The Human Body Map is a highly complex visualization that aims to present tumor growth dynamics across timepoints in anatomically correct regions of the human body, based on the collected SDTM data.

The Human Body Map design is based on the following goals.

- Tumor visibility: All tumors should be visible on the map.
- Anatomically correct placement: Tumor visualization should be placed on the organs/body parts they belong to. Tumors should not cross the boundaries of their respective regions.
- No overlap: Tumors that belong to the same region should not overlap, ensuring a high level of interpretability of the plot.
- Proportionality:
    - Tumor visualizations should always stay proportionate to each other - larger tumors must appear larger than smaller ones.
    - Tumor visualizations should be proportionate to the region they belong to - the area a tumor occupies should not significantly differ from what an actual tumor of the same size would look like on a real body part.
- Event handling: The visualization should handle merge and split events of tumors.
- All the above-mentioned requirements should be maintained across all time points.
- Traceability: Tumor visualizations should be traceable through time - the same tumor should always appear in the same place.
- Size dynamics: Changes in tumor sizes should be visible.

All the requirements above can be formalized into the following problem:

THE PROBLEM

We are dealing with a known mathematical problem with a set of complications. Specifically:

Circle Packing Problem with the following additional complications:

1. Circle proportionality to the region. Tumor size must remain proportional to both its anatomical region's area and to each other, while meeting a minimum visibility threshold.
2. Spatial consistency through time. Each tumor's visual center must remain fixed across all time points.
3. Visibility requirement on absolute value and change. Visualization must ensure that both small and large tumors remain distinguishable.
4. Special logic for handling merge and split cases. When tumors merge or split, their spatial transitions should appear smooth: before merging, individual tumors are placed next to one another; after merging, the new lesion occupies the intermediate position; after a split, the resulting tumors are arranged around the original site.

SOLUTION

To tackle the Circle packing problem, let's start with the known solution and add complications one-by-one.

PACKING CIRCLES IN ONE TIMEPOINT

A fixed center c_i is assigned to each tumor i. Since we are presenting each tumor as a circle on our silhouette, initially the longest diameter of the tumor r_i is selected to be the radius of the corresponding circle. Here we face the first problem of trying to use the original size as is (in mm).

The size of the body is not fixed; we cannot use absolute size for tumors and need to make tumor size proportional to the illustrated human body while maintaining minimum visibility size. Another important consideration is ensuring the proportionality of tumor sizes within each other, i.e. pre- and post- transformation size differences should have the same ratio. New radii are defined using a scaling function and a global minimum visible radius:

$$r_i' = \max(r_i \cdot s_R, r_{min}) \quad (1)$$

where s_R (the scaling function) provides a coefficient based on the anatomical region area, and r_{min} is the minimum visible tumor radius which is based on the display size of the human body. This ensures minimal visual distortion.

APPLYING TUMOR PACKING THROUGH TIMEPOINTS

Since the sizes of tumors are changing over time, tumors that do not overlap at one timepoint may overlap in a different one. To address this, two algorithms are used: Conservative and Greedy.

Conservative algorithm selects the maximum size across all measurements for each tumor and tries to pack them inside the organs using the standard constraints below:

$$\left|\left|c_i - c_j\right|\right| \geq r_i^{max} + r_j^{max} + \epsilon \quad (2)$$

$$dist(c_i, \partial P_R) \geq r_i' \quad (3)$$

The first constraint ensures that two tumors do not overlap. Here ϵ is a minimal gap margin that ensures the circles do not touch. The second constraint ensures that tumors do not cross organ boundaries. Here ∂P_R is the organ polygon boundary.

After initial placement, we want to spread all tumors within an organ and make sure that they are not accumulated at any subregion. For that purpose, Lloyd Relaxation was used to smooth out local spacing.

In case the adjusted area of the tumors at maximum size is bigger than the area of the organ as illustrated we switch to Greedy algorithm. That allows us to cover the drawbacks of the Conservative algorithm.

The **Greedy** algorithm uses iterative optimization to determine the optimal center coordinates while ensuring that no conditions lead to the problems described above.

Procedure includes several steps:

1. Initialize all centers c_i by random sampling within the organ
2. An energy function was defined as

$$E = \sum_{i<j} \phi\left(\left||c_i - c_j|\right| - r_i' - r_j'\right) + \sum_i \psi(dist(c_i, \partial P_R) - r_i') \quad (4)$$

3. Iteratively minimize E with gradient descent until convergence
4. Re-sample positions if any c_i violate the abovementioned constraints after convergence

HANDLING MERGE AND SPLIT CASES

After making sure that tumors do not overlap, we should consider the special cases of tumors splitting or merging. That is addressed by using a weaker constraint (5) instead of (2), which allows splitting and merging tumors to be tangent:

$$\left||c_i - c_j|\right| \approx r_i^{max} + r_j^{max} \quad (5)$$

Additionally, in case any tumor A and B merge to form tumor M, A and B centers are initially positioned close to each other using the given formula:

$$\left||c_A - c_B|\right| \approx r_A' + r_B' + \epsilon \quad (6)$$

where ϵ is a small gap margin (almost touching). After the merge, the center of tumor M is determined by the formula

$$c_M = \frac{\omega_A c_A + \omega_B c_B}{\omega_A + \omega_B} \quad (7)$$

Here ω_A, ω_B are weight coefficients proportional to the areas and radii of the tumors A and B.

If tumor S splits to form tumors S_1 and S_2, the centers of the split tumors are dispersed outward from the tumor S center with the formula:

$$c_{S_j} = c_S + (r_S' + \delta)\hat{u}_j \quad (8)$$

Where $\hat{u}_j$ are evenly spaced unit vectors around the parent center and δ is a small gap to avoid overlap.

HUMAN BODY MAP FINALIZATION

By applying the abovementioned logics in the main algorithm that creates body visualization and by further enhancing the visualization with dynamic annotation and cross-filtering functionality we end up with highly informative patient profile dashboard, which merges into the main concept of ResponDEX dashboard.

CONCLUSION

ResponDEX has the potential of transforming the study analysis review process. Study setting features and built-in automated tools significantly reduce the time spent on programming while allowing more time for getting into the essence of data. The extremely modular design allows to infinitely scale up the system and add new endpoints, therapeutic areas, study templates, etc. This means that ResponDEX will continuously grow, covering more and more use cases. With the extra flexibility of allowing one to interactively review results on both study and subject levels, ResponDEX has all the prerequisites of becoming a key part of any clinical trial's life cycle, being an extra validator for analysis, and replacing programming completely in the future.

CONTACT INFORMATION

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

Author Name: Sevak Kamalyan
Company: Armat Analytics
Email: sevak.kamalyan@armatanalytics.com

Author Name: Alexander Mkrtchyan
Company: Armat Analytics
Email: alexander.mkrtchyan@armatanalytics.com

Address: 2 Adonts St, 0014, Eraz Business Center
Work Phone: +374 98 545222
Email: info@armatanalytics.com
Website: www.armatanalytics.com