

IDEO, an Idea for Interactive and Visual Representations of Oncology Summary Statistics

Charlotte Cheinin, Sanofi, Chilly-Mazarin, France

ABSTRACT

In the scope of efficiently delivering more interactive and user-friendly statistics to team members who interpret clinical trial analysis results, a web-based application can be a solution. In this context, Sanofi statistical programmers and statisticians have constructed Interactive DEcision Oncology (IDEO) as an RShiny application dedicated to support project team decision making process based on faster delivered statistical results in oncology.

INTRODUCTION

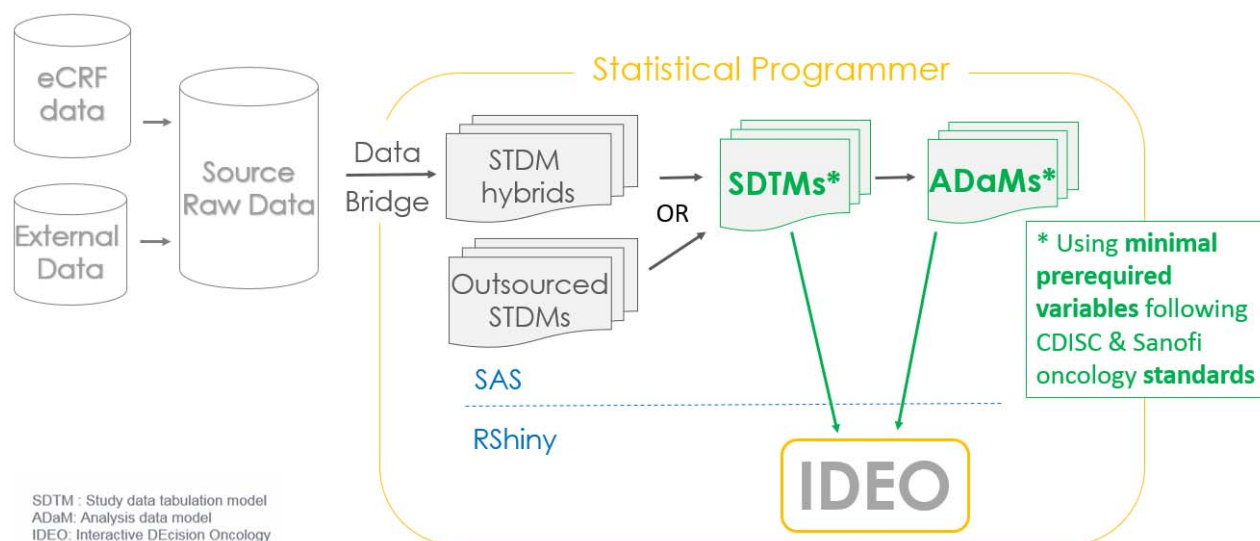
IDEO is an in-house RShiny application providing interactive and visual representations of summary statistics for decision takers (e.g. statisticians, clinicians, pharmacokinetics, management). The statistical programmers who develop the application use stable and validated R and RShiny code and Clinical Data Interchange Standards Consortium (CDISC) standardized data as source. It allows the 1st version of this application to be used on multiple studies in early oncology (phase I to IIa) with minimal adjustments.

IDEO, AN RSHINY APPLICATION FOR ONCOLOGY STUDIES

As an application fully integrated into Sanofi clinical data flow (Figure 1), the 1st release of IDEO is focused on a selection of descriptive statistics in:

- Safety for all oncology studies
 - o Lineplots and boxplots of laboratory values, percent change from baseline and corresponding grades derived based on National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE)
 - o Tables of treatment emergent adverse events (overview, by System Organ Class/Preferred Term with filtering options on seriousness, adverse event of interest, leading to drug withdrawal, leading to death, grade ≥ 3 ...)
- Efficacy for studies with Response Evaluation Criteria In Solid Tumors (RECIST)
 - o Swimmer plots of duration of treatment and best overall response
 - o Waterfall and spaghetti plots of relative change from baseline in tumor measurement.

Figure 1: IDEO in the Sanofi GxP dataflow



PROGRAMMING FOR MULTIPLE STUDIES

When considering using a single application accessing data of multiple studies to create Tables/Listings/Figures (TLFs), a right balance between standardization and flexibility is preferable.

STANDARDIZED SOURCE DATA

IDEO source data are standardized:

- CDISC compliant datasets and variables which are developed for clinical report purpose
- Additional standards are defined at the company level to complete current available CDISC requirements, e.g.:
 - o ADaM structure for RECIST evaluation of tumor assessments and responses
 - o controlled terms for the variable PARAMCD in the ADaM dedicated to exposure summary
ADES.PARAMCD="DURD" for a treatment duration in days:

```
# Add overall subject's treatment duration to RECIST results
data.src <- data %>%
  full_join(ADES %>%
    select(STUDYID, USUBJID, SUBJID, PARAMCD, AVAL, ESTRT) %>%
    filter(PARAMCD == "DURD", ...
```

To have first IDEO TLFs available as quickly as possible to the end users from 1st data entry in the Case Report Form (CRF), the selection of needed SDTMs and ADaMs datasets and variables is pre-identified. And instructions are available to the study statistical programmers to prioritize the programming. For instance, in a dose-escalade study, datasets used to produce statistics on safety are needed to be looked at earlier than results on efficacy, biomarkers and pharmacokinetics (PK).

FLEXIBLE PROGRAMMING AND USER-FRIENDLY INTERFACE

An objective of IDEO is to be more efficient and gain time. Thus, the programming needs to be as flexible as possible to diminish if not delete any study-specific programming steps:

- IDEO programming structure has been thought as modules and functions called in from programs `global.r` and `server.r`. This way, efforts to deactivate one TLF section from the user interface for a study are minimized (e.g. efficacy RECIST section for a non-solid tumor cancer).
- Some programming tricks were also used to have a more flexible program and a more user-friendly interface, e.g.:
 - o A drop-down list for treatment arms is likely to be sorted by alphanumeric order (Figure 2) when it does not make sense in the context of a clinical trial. Indeed, we would more likely see these arms sorted with the corresponding order they are tested in a trial (Figure 3) using a flexible programming approach:

```
# Sort arms according to numerical arm variables
trt <- as.character(unique(ADSL$TRT01A[order(ADSL$TRT01AN)]))
```

Figure 2: List of arms by alphanumeric order

Choose an actual treatment group:

SAR123456 1 mg/kg + DrugX 3 mg/kg Q2W, SAR123456 3 mg/kg Q2W ▾

Select All	Deselect All
SAR123456 1 mg/kg + DrugX 3 mg/kg Q2W	✓
SAR123456 1 mg/kg Q2W	✓
SAR123456 10 mg/kg Q2W	✓
SAR123456 3 mg/kg + DrugX 3 mg/kg Q2W	✓
SAR123456 3 mg/kg Q2W	✓

Figure 3: List of arms aligned with to the study design

Choose an actual treatment group:

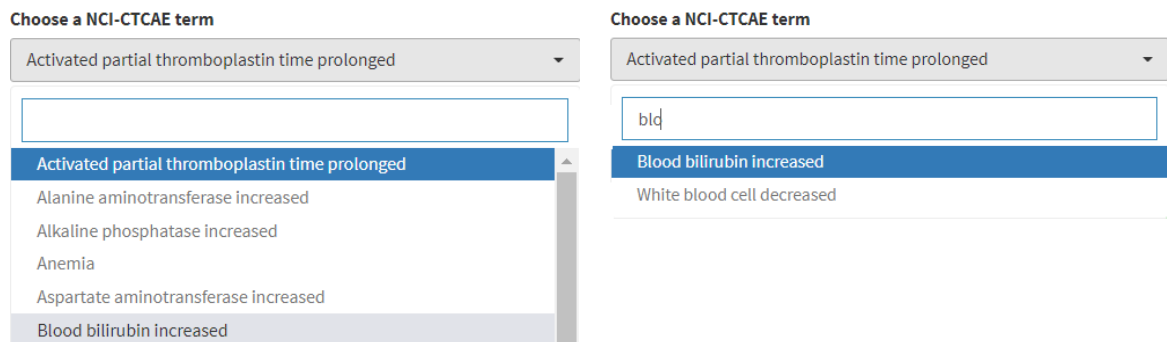
SAR123456 1 mg/kg Q2W, SAR123456 3 mg/kg Q2W ▾

Select All	Deselect All
SAR123456 1 mg/kg Q2W	✓
SAR123456 3 mg/kg Q2W	✓
SAR123456 10 mg/kg Q2W	✓
SAR123456 1 mg/kg + DrugX 3 mg/kg Q2W	✓
SAR123456 3 mg/kg + DrugX 3 mg/kg Q2W	✓

- o The 'live-search' option of `pickerInput{shinyWidgets}` has been activated to allow a user to look for a specific term by typing a few characters in a long drop-down list (Figure 4):

```
column(pickerInput(inputId="ctcae",
                    label="Choose a NCI-CTCAE term",
                    choices=ctcae,
                    options=list('live-search' = TRUE))
)
```

Figure 4: Typing characters in the search area to restrict the length of a drop-down list



- o A yellow background mouse-hover has been added to facilitate the reading of a single row in the large tables of adverse events on a screen (Figure 5):

- In the customized stylesheet `custom.css`, a yellow background for mouse hover is defined for the dashboard container identified by `tabName="ae"`

```
#shiny-tab-ae tr:hover { background-color: #ffff99; }
```

- In program `ui.R`, the stylesheet `custom.css` is imported to the dashboard body and the background mouse-hover color is applied to the `tabItem{shinydashboard}` using `tabName="ae"`

```
body <- dashboardBody(
  tags$head(
    tags$link(rel="stylesheet", type="text/css", href="custom.css")
    ...
  ),
  tabItems(
    tabItem("home", homeUI(id="home")),
    tabItem("help", helpUI(id="help")),
    ...
    tabItem(tabName="ae" ...),
    ...
  )
)
```

Figure 5: Yellow highlighting mouse hover of a row to facilitate reading in a large table

PRIMARY SYSTEM ORGAN CLASS	Preferred Term n(%)	1 mg/kg Q2W (N = 3)		3 mg/kg Q2W (N = 10)		10 mg/kg Q2W (N = 4)		1 mg/kg + DrugX 3 mg/kg Q2 (N = 3)	
		All grades	Grade ≥ 3	All grades	Grade ≥ 3	All grades	Grade ≥ 3	All grades	Grade ≥ 3
Any class		3 (100.0)	0 (0.0)	9 (90.0)	6 (60.0)	4 (100.0)	3 (75.0)	3 (100.0)	3 (100.0)
Infections and infestations		0 (0.0)	0 (0.0)	1 (10.0)	0 (0.0)	1 (25.0)	0 (0.0)	1 (33.3)	0 (0.0)
	Urinary tract infection	0 (0.0)	0 (0.0)	1 (10.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (33.3)	0 (0.0)
	Oesophageal candidiasis	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (25.0)	0 (0.0)	0 (0.0)	0 (0.0)
	Respiratory moniliasis	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

STANDARDIZED ANALYSIS TLFs

The TLFs included in IDEO mimic commonly used TLFs in clinical reports. They are company standards which support the decision-making process in a more graphical and interactive way. After the proof-of-concept of the capability and feasibility of IDEO in the commonly used TLFs, the future development will generate more comprehensive statistical analyses to leverage the interactive feature to deep dive into the clinical data and support efficient decision making.

MAINTENANCE AND VALIDATION

The results contained in the 1st release of IDEO were validated with dual programming in SAS®.

The development team brainstorms about different validation approaches for the future:

- Maintain reproducibility of results over time when R community is very active
 - o Use stable versions of R software
 - o Stabilize the environment with docker
 - o Process to evaluate eventual impacts when updating to newer versions of software or package
- Validate the what and how of the application
 - o Use R base validated packages or widely used low-risk packages whenever possible
 - o Look at implementing R packages that facilitates testing
 - o Write adapted validation documentation.

ACCESS MANAGEMENT

The entire clinical dataflow (raw data, SDTMs, ADaMs) and IDEO application are stored in Sanofi GxP area. No clinical data are transferred nor stored outside this GxP environment. And the users log in to IDEO using their network identifier and password. Thus, access to the data and study results is as much protected as per our usual approach with deliverable of fixed TLFs in rtf/pdf files.

UTILISATION OF STATISTICAL RESULTS

The project and study statisticians remain at the center of any decision and interpretation of statistical results. They are accountable for the integrity of results presented into IDEO. In addition, trainings were put in place to educate the end-users on the utilization of the application prior to be granted access for a study. End-users acknowledge to not communicate any results presented into IDEO interface: it is forbidden to screenshot results and RShiny package options to download the graphics have been deactivated. Graphics and other statistical results included into IDEO and needed outside IDEO can still be provided by the study statistician and study statistical programmer.

CONCLUSION

Currently, we collect suggestions from the pilot study team which uses the 1st version of IDEO. Next releases are planned to:

- Integrate users' feedbacks
- Make sure that IDEO is a unique door to access any study linked to IDEO. There will not be one application by study. To do so, programming background work is needed for IDEO and a new feature will be added in the interface for the user to select the:
 - o Study, e.g. a project statistician needs to access all the studies of the compound into IDEO while a study statistician needs to only access to his/her study
 - o Study database extraction, e.g. a user can generate TLFs from a previous and cleaned extraction while the study continues to go over next intermediate analyses up to the final lock
- Activate options depending on the user's role, e.g. result download might be authorized for study statistician and statistical programmer
- Be extended to demography and disposition, exposure, biomarkers and PK results, additional filters and options
- Extend the scope of the application to a selection of non-solid tumor cancers for the efficacy section.

ACKNOWLEDGMENTS

For the 1st release of IDEO, big thanks to Paulin Charliquant, Jason Deng, Alice Gosselin, Jenny Hou, Zhanyun Hu, Ji Lin, Martine Madec, Cindy Marabotti, Jean-Christophe Quirot, Nils Ternes, Bin Wei, Ting Zhang.

CONTACT INFORMATION

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

Charlotte Cheinin

Sanofi

Biostatistics & Programming, B10

1 avenue Pierre Brossolette

91385 Chilly-Mazarin, France

Email: charlotte.cheinin@sanofi.com

Brand and product names are trademarks of their respective companies.