SHINY STATS
EXTENDING THE REACH OF STATISTICS THROUGH INTERACTIVE APPLICATIONS

Thomas Brown and Anastasiia Raievska, Veramed

1. Introduction

A common scenario for the clinical statistician is to be involved in discussions following a theme of 'innovative' and 'efficient' clinical trials. Such discussions are important and fundamental to the value we as statisticians can bring to the pharmaceutical industry. However, how can we leverage our expertise beyond the canonical design and analysis of clinical trials? In this poster we highlight how R Shiny can be used as a tool to enhance the value pharmaceutical statisticians bring to the product development cycle. Two case studies are presented. The first proposed R Shiny for enhancing a trial's screening criteria and the second involved the analytical platform for preclinical development.

2. Objectives

- Detail two case studies where R Shiny was implemented outside of clinical trial design and analysis to enhance pharmaceutical development

- Evaluate the successes of the developed applications and address roadblocks we encountered during the application development and deployment processes

- Provide strategic guidance to help others who are considering developing a shiny application in the pharmaceutical industry

3. Shiny

Shiny is an R package developed by R Studio that allows interactive web based applications ('apps') to be built using R code. Apps can be built to perform any type of statistical analysis or technique and can use numerous data structures. Once built, an app is hosted on a Shiny Server where it can be accessed from anywhere with an internet connection.

4. Case Study I – Enhancing Early Phase Screening Criteria

Recruitment difficulties during a Proof of Concept (PoC) study can arise from:
- The indication having slow disease progression
- Patients already responding to the current standard of care (SoC).

To help recruitment difficulties, we were required to identify suitable patients (i.e. fast progressors who do not respond to SoC) for the study using a surrogate biomarker.

The surrogate biomarker was assessing the decline in kidney function by looking at the rate of decline in estimated Glomerular Filtration Rate (eGFR) prior to enrolment. Patients were eligible for recruitment if they had at least a 10% decline in eGFR over the preceding 9 months calculated from a linear model fitted using patient eGFR data up to 2 years prior to screening. This analysis can be difficult for investigators without adequate statistical training. An R Shiny app was proposed to allow an analytical strategy designed by statisticians while enabling approximately 30 European sites to access the app. The cloud based structure means the only technical requirement is an internet connection and browser.

The study team agreed with this approach and a prototype was developed. However, issues associated with its deployment arose. The main challenges were IT-related:
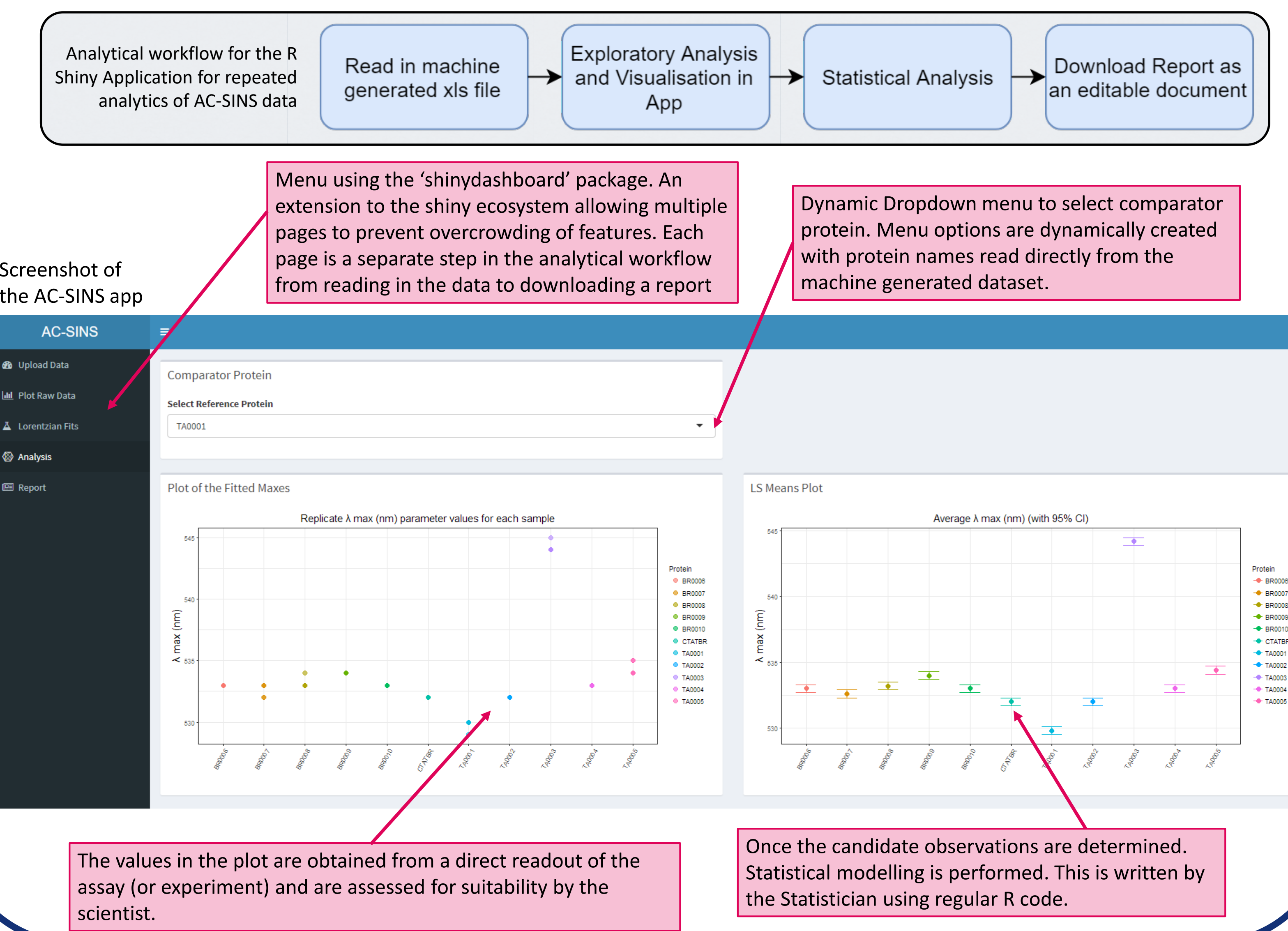
Goal:	Develop the tool to facilitate the assessment of key inclusion criterion
Timeline:	9 months

	Study team perspective	IT perspective
Cloud based **R Shiny**	- Easy access from any device with an internet connection - Quick development by a statistician	- Novel tool and requires external secure server - Expensive and requires extended validation
Local based **Excel tool**	- Requires excel and system specifications - Interface inferior to R Shiny	Traditional tool and easy to validate given the timeline

The key requirement for the R Shiny app was the existence of an external secure server to ensure data security while being processed online. Even with a 9 month timeframe, the sponsor's own imposed regulations related to server set up and validation made it unfeasible. A macro-enabled Excel spreadsheet was developed instead and sent individually to European sites.

5. Case Study II – Pre-Clinical Analytics

Affinity-capture self-interaction nanoparticle spectroscopy (AC-SINS) is a technique that allows assessment of colloidal stability of antibodies and can aid therapeutic candidate selection. The large number of antibody variants in the early discovery pipeline makes data analysis a complicated and laborious task. Our client's pre-clinical research statistics team identified the potential for several methodological improvements in the analytical procedure which would enhance the screening of candidate antibodies. However, given the repetitive nature of the task, substantial resource is required by the statistician to ensure the analysis is performed appropriately. Secondly, input is also required from the scientist following an initial exploratory sequence of data visualisations and descriptive results to guide the direction of subsequent analytical processes.

R Shiny was proposed as a tool to overcome these two issues. An interactive app would allow scientists to use their domain expertise in directing the analysis while writing the application in R would mean the statistician could control and ensure appropriate methods and techniques were being followed consistently.

Analytical workflow for the R Shiny Application for repeated analytics of AC-SINS data → Read in machine generated xls file → Exploratory Analysis and Visualisation in App → Statistical Analysis → Download Report as an editable document

Screenshot of the AC-SINS app

Menu using the 'shinydashboard' package. An extension to the shiny ecosystem allowing multiple pages to prevent overcrowding of features. Each page is a separate step in the analytical workflow from reading in the data to downloading a report

Dynamic Dropdown menu to select comparator protein. Menu options are dynamically created with protein names read directly from the machine generated dataset.

The values in the plot are obtained from a direct readout of the assay (or experiment) and are assessed for suitability by the scientist.

Once the candidate observations are determined. Statistical modelling is performed. This is written by the Statistician using regular R code.

6. Recommendations and Conclusions

Scope → **Coding** → **Testing** → **Deployment** → **Considerations beyond deployment**

Scope: Agree on the scope of the application's functionality as early in the development process as possible.

A simpler application in R Shiny is usually easier to adopt. Additionally, it will help prevent code from bloating, which makes it difficult to share and iterate on.

Coding: Write general reactive functions to do certain analytical functions and data manipulations that will be used multiple times in the app.

If more features are added at a later stage, the code will be easier to edit and less prone to producing errors.

Testing: Stress testing the app is critical to ensure it behaves as expected in a wide range of scenarios. Given the user base will extend beyond those developing the application, it is difficult to predict how a user will interact with the app.

Deployment: Deploying an app on a server can lengthen timelines depending on the organisation's structure and IT infrastructure.

To get users to engage with the app as early as possible, send out code they can run on their own installation of R or on a corporate wide R Server. This will help get feedback as early as possible.

Considerations beyond deployment:

Internal / External App: Additional time and cost if app is used externally to the Sponsor's organisation → Corporate Branding and setting up external access to the R Shiny Server

Handling Sensitive Data: Steps need to be made to ensure GDPR compliance. There may be additional costs associated → Have a secure external server and consider adding in user authentication to the app

GxP: Time needs to be given to ensure GxP compliance → Additional validation of the app's handling of data and analytical procedures

Management and Oversight: If a CRO is involved, factor in the required additional time and costs → Consider CRO training requirements to effectively engage with the app

Conclusions for using R Shiny
- Real-time analysis
- Friendly interface
- Automation and efficiency
- Enhanced collaboration between scientists and statisticians
- Requires buy-in from IT