Human Brain Project

Project Title:	Human Brain Project
Sub-Project Title:	Medical Informatics Platform (SP8)
Document Title:	**Web UI – Analyses Flow - description**
Summary:	Description of the steps for exploring and launching analysis. This document describes the specification for EE (Epidemiological Exploration) and IA (Interactive Analysis) part II. Part I (PoC) was delivered by Virtua V*II* on 29th September 2015, and part II should be an enhancement of the initial functionality.

Versions, changes and contributors:

Authors	Version	Date	Status	Change details
Mihaela Damian, CHUV	v1.3	27/10/2015	Draft 1	
Mihaela Damian, CHUV	v5	08/12/2015	Draft 2	Modified process workflow
Mihaela Damian, CHUV	v6	06/01/2016	Final	Added missing information Pg 9 – stats will be on the side, not within the variable

TABLE OF CONTENTS

1 Background

The Medical Informatics Platform (MIP) is one of the six platforms to be developed as part of the Human Brain Project, and will provide full services to end users (clinicians, researchers, statisticians, health managers, general public) for investigating data correlations, distributions and interactions of disease processes and epidemiological factors.

MIP platform will provide means of accessing and interrogating remotely stored data and visualising results (via a Web UI), back-end software framework and micro-services to manage the distribution of statistical algorithms, run computations and merge locally calculated results.

2 Purpose of Document

The current document describes the specification for the Web UI (SOW2). SOW1 was delivered by Virtua *VII* and SOW2 is an improvement and further development of the functionality already provided.

Full code of SOW1 will be available.
Technologies used:
- Front-end: AngularJS, Highcharts - http://www.highcharts.com/
- For web design - http://themeforest.net/item/avenxoiresponsiveiangularjsihtmliadmin/11660185 with Grunt, Bower

3 Introduction

The end-user (statistician, scientist, clinician, public) will want to generate a data model. The data model will be finally represented through a plot, a design matrix and statistical data associated.

The user should have the possibility to change the parameters of the model and regenerate it.

Main actions of user :

- **Explore** and Select the variables and/or variables' group

- **Configure** their variables and the query to be run (the scientific method)

- **Display:** user may study the displayed result (=model)

    At this point s/he can:

- Potentially refine the model (back to the 1st or 2nd step)

- Potentially choose another plot visualisation

- Save or delete the model

Please see below the above steps in more detail.

4 Curent solution - Interactive Analyses (IA)

4.1 Plot creation: select parameters

The end-user (statistician, scientist, clinician, public) will want to generate a plot over existing data.

S/He selects the desired parameters, which are organised in: *variables*, *groupings*, *covariables* and *filters*, at least one of each type.

The parameters are organised in a tree structure. Once they are selected, the user clicks <Execute> to display the plot.
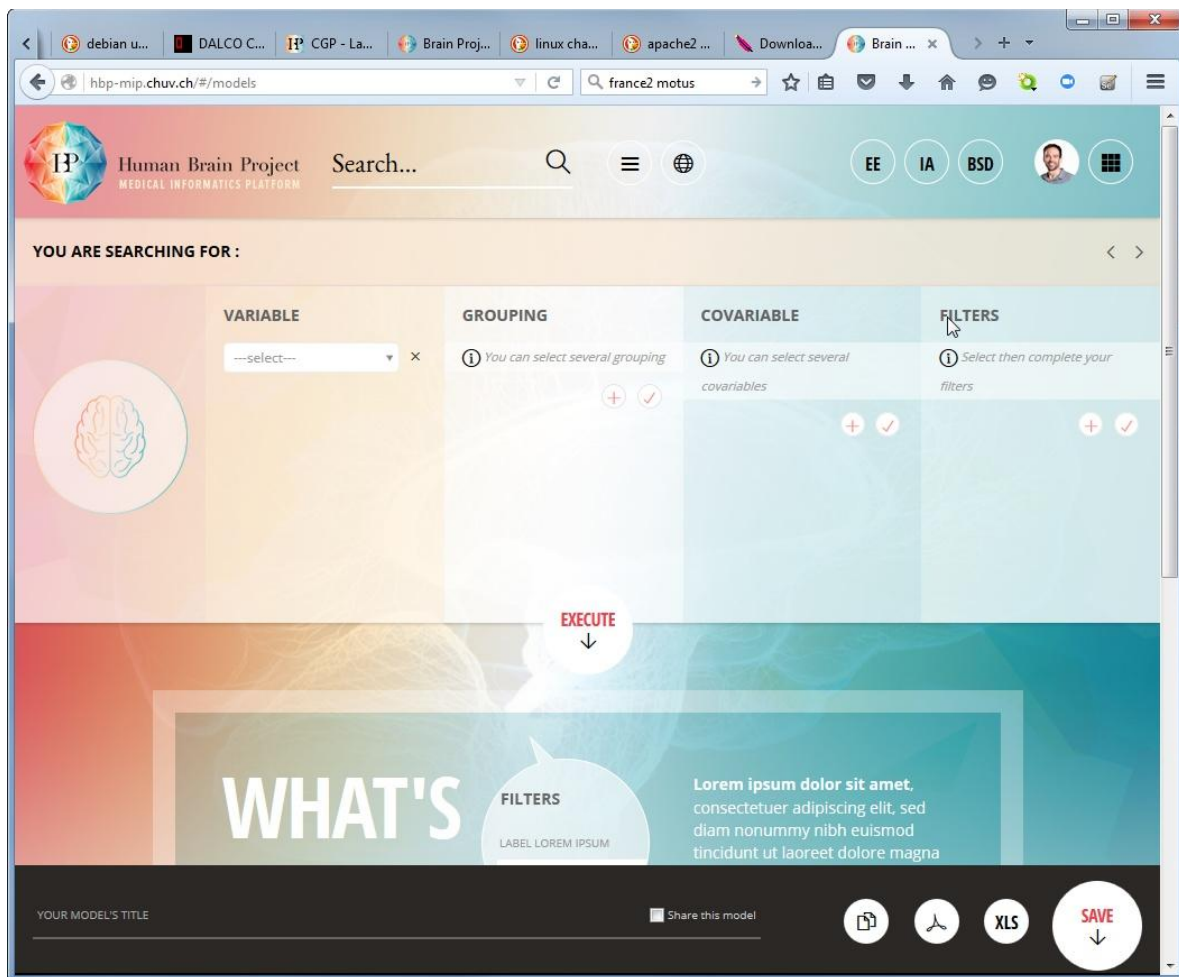
4.2 Plot creation: "Execute"

Once displayed, the plot type may be changed and its axis and colours can be customised.
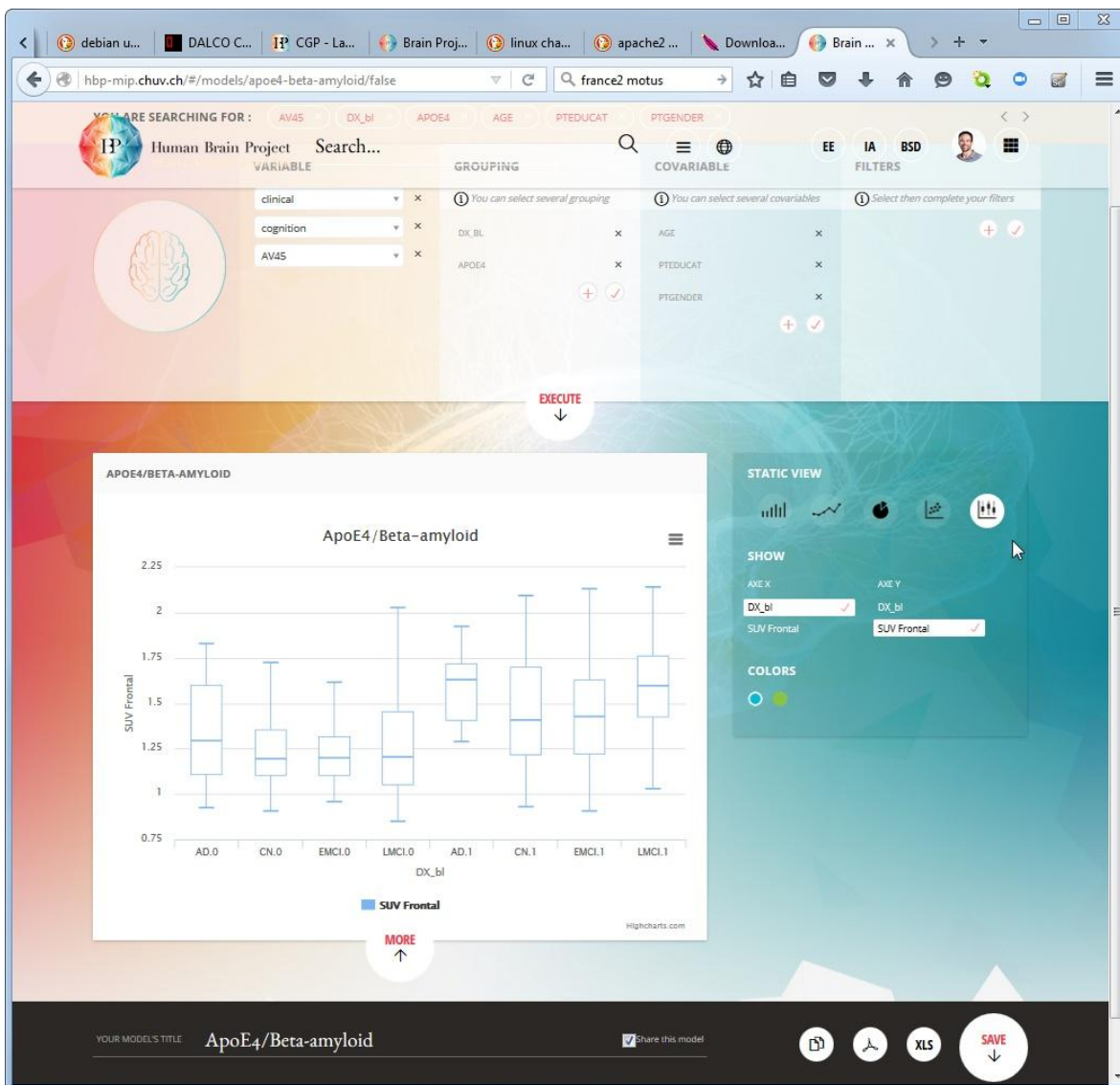
5 New Solution

5.1 Summary of Changes

C1: The selection of data/parameters will be done through a Circle Packing. Here, the user **EXPLORES** the ontology (via Circle Pack) and selects the parameters of interest.

(Ontology = repository of variables, compartmentalised and with established inter-relationships. E.g. they may belong to a tree-class structure, would be related to other variables or classes and may have specific attributes)

The ontology variables will be sent to the Web UI from the backend.

C2: The categorisation of the parameters selected in C1 (as *variables*, *groupings*, *covariables* and *filters*), will now be done in a separate step.
Here, the user **CONFIGURES** the Scientific Method(s) by categorising the parameters and selecting the Scientific Methods to be run. One or more methods can be selected.

C3: Extend plot display.
The UI will now **DISPLAY** 3 visualisation types**:**

1) plot - 1st/default

2) full stats

3) Design Matrix

Default page: Design Matrix
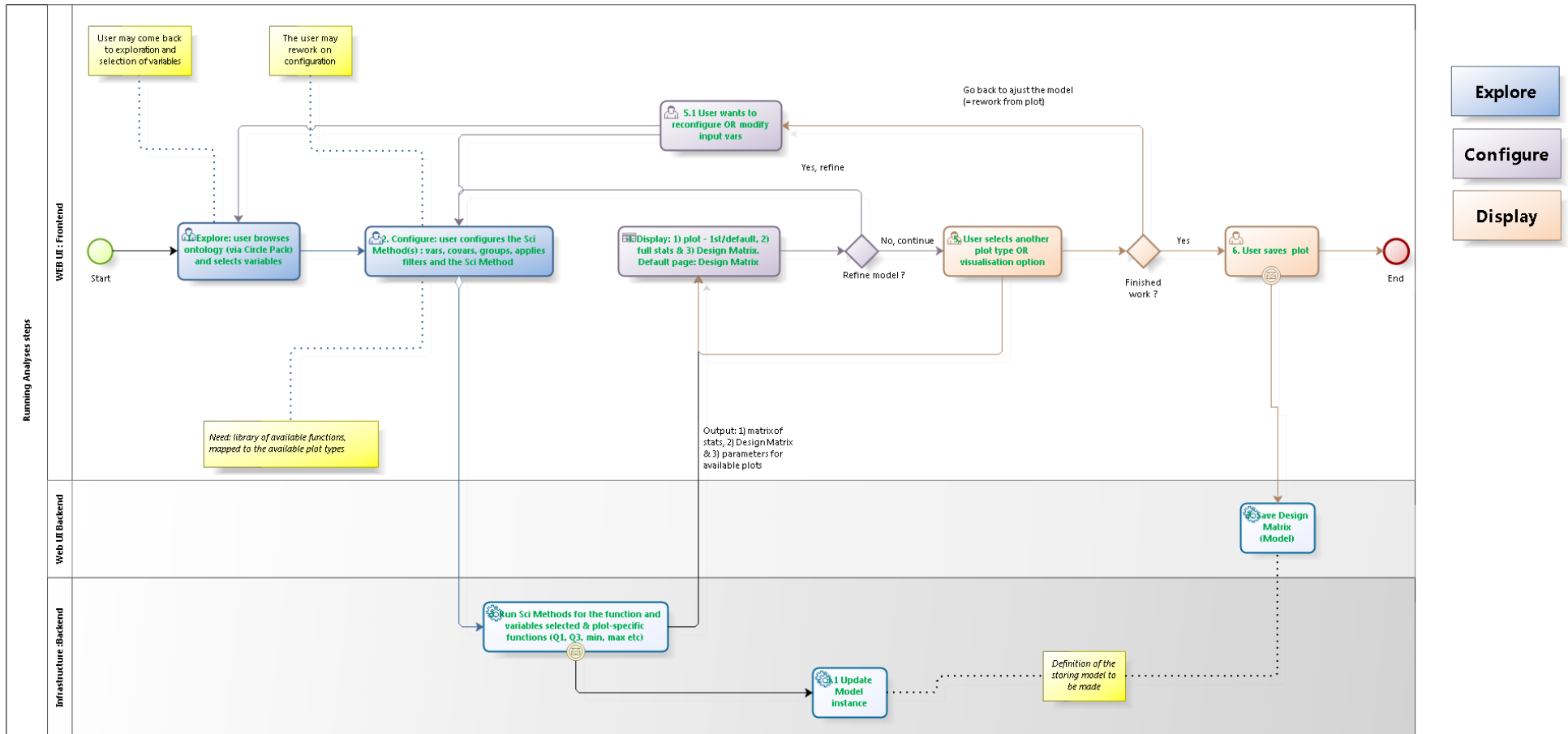
The **EXPLORE, CONFIGURE and DISPLAY** steps will be organised sequentially, the process stepping from Explore to Configure, and from Configure to Display.

The UI will behave in an "accordion" manner: the current step will expand, whereas the other two will be retracted.

Most of the interactions will be on the frontend, but some will access the backend through JSon Restful services

5.2 Process Flow / Business Process

User may come back to exploration and selection of variables

The user may rework on configuration

Go back to ajust the model (= rework from plot)

5.1 User wants to reconfigure OR modify input vars

Yes, refine

Explore: user browses ontology (via Circle Pack) and selects variables

2. Configure: user configures the Sci Method(s) : vars, covars, groups, applies filters and the Sci Method

Display: 1) plot - 1st/default, 2) full stats & 3) Design Matrix. Default page: Design Matrix

No, continue

User selects another plot type OR visualisation option

Yes

6. User saves plot

Start

Refine model ?

Finished work ?

End

Need: library of available functions, mapped to the available plot types

Output: 1) matrix of stats, 2) Design Matrix & 3) parameters for available plots

Save Design Matrix (Model)

Run Sci Methods for the function and variables selected & plot-specific functions (Q1, Q3, min, max etc)

1 Update Model instance

Definition of the storing model to be made

Explore

Configure

Display

Running Analyses steps

WEB UI : Frontend

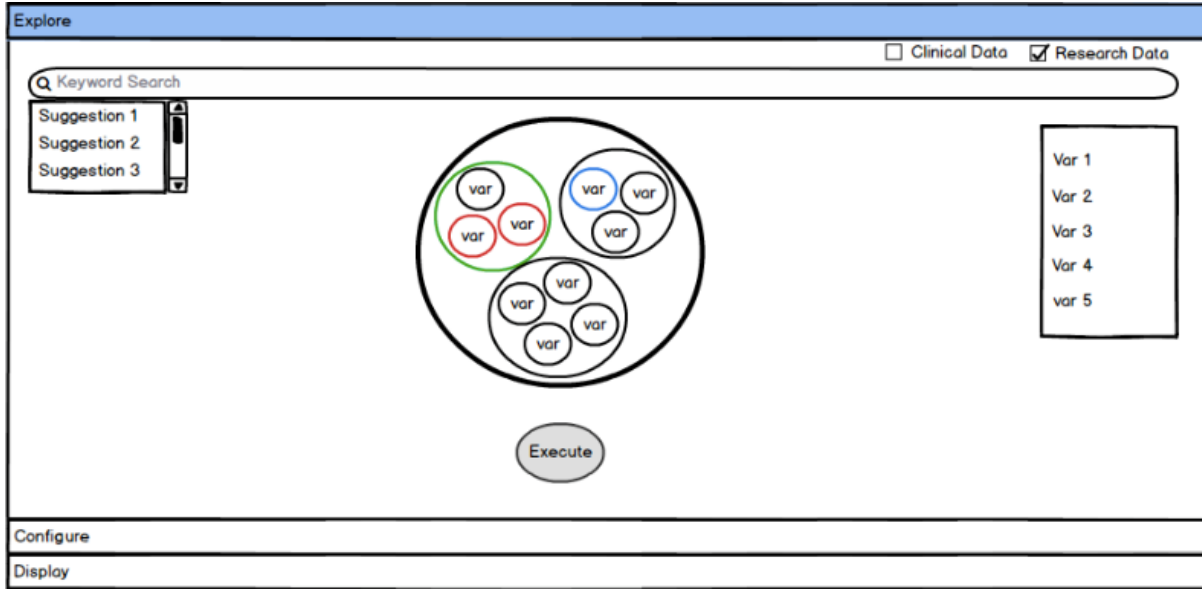Web UI :Backend

Infrastructure :Backend

5.3 Running Analyses: steps explained

5.3.1 Process Elements

5.3.1.1 ⬤ *Start*

5.3.1.2 🖾*1. EXPLORE: user browses ontology (via Circle Pack) and selects variables*

Above - a mockup suggested by CHUV team summarising the Exploration step.

User browses the ontology and selects the variables he wants to use. A « bubble browser » (circle pack) graphical view for displaying and browsing the groups and their variables is proposed.

Examples:

http://www.visualcinnamon.com/occupations, http://bl.ocks.org/mbostock/7607535
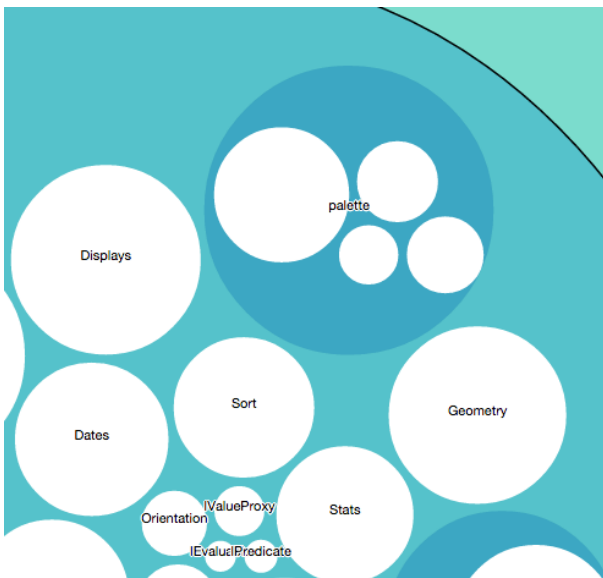
There will be multiple levels of groupings (max 3 at the moment: Domain, Construct and Sub-construct – a sample will be provided for info), on the 4st level (the last) being the variable itself.

The user should be able to browse the groups and variables, zoom in and out.

He should also be able select variables and/or groups. In the next step, CONFIGURE, the user should also be able to configure/organise/prepare the variables for being used by the statistical methods (categorise in variable, co-variables, groups, methods to be applied, any configuration)

To facilitate an easier browsing for users who are familiar with the circle pack and the ontology, a <keyword search> box should be available. When, for example, the user searches for a group or a variable name, the « bubble selector » will focus on the appropriate circle.

A suggestion dropdown will appear to help the user on the variables selection.

The selected variables will appear as a list on a side.

Summary of possible actions within a bubble:
- user zooms in and out to look-up variables

- user may select variables of interest. When a whole class is selected, all the variables contained are considered

  Variables will be represented as belonging to classes, on several levels, reflecting the way the Ontology is organised. There will be as many levels as described in the Ontology.

Summary of information to display in a bubble:

- Variable name (*Note:* here, *the "variable" is used as a general name for the parameters selected during EXPLORE. In CONFIGURE some variables will become co-variables or groupings*)

- Variable short description (may be a tool-tip on mouse over – TBC the exact means of display)

- each variable will have one of the attributes below (TBC how these will be displayed):

Behaviour
Genes
Molecules
Cells
Circuits –Anatomy
Circuits – Function

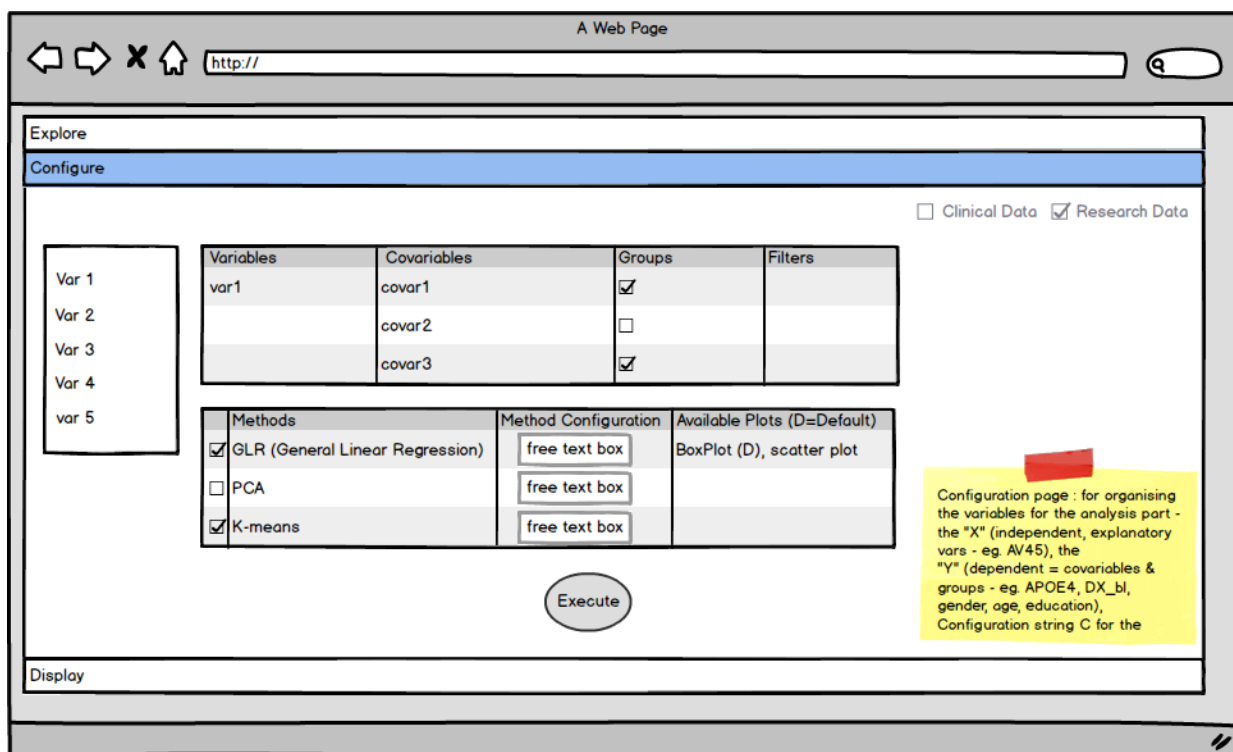**Summary of information to display outside the Circle pack:**

- **Stats:** For each element of the circle pack (group or variable), stats will be displayed. Stats will show various simple computations, i.e. number of cases (COUNT) for that variable or group, MIN, MAX, AVERAGE etc, all in rapport to covariables: age, education, gender and pathology. The stats will be clearly displayed for each of the data sources of the system (5 at the moment). For simplicity, we

Once all the variables are selected, click on the <Execute> button to go to the next step, CONFIGURE.

5.3.1.3 📇 *2. <u>CONFIGURE:</u> user configures the Scientific Method(s) : variables, covariables, groups, applies filters and the Method*

Above - a mockup suggested by the CHUV for the Configuration step.

Description

After the exploration and selection of variables, the user will prepare the parameters for the scientific methods (setup the query).

"Setup a query" refers to choosing the query parameters to call the JSon Restful service. The parameters are the variables, grouping, covariables, filters, the scientific methods and their *Configuration*.

1. Variables

a. Will choose which ones will be considered as a variables (can be one or several) and as covariables (can be one or several).

b. Out of the selected covariables, the user chooses which ones he wants to have as a « grouping » element. Can have more than one grouping.

c. Also for each covariable, the user can choose an interval of min and max value (filter).

2. Scientific Methods (/Functions/Algorithms)

a. The user can choose to apply one or many algorithms on the query (for example a linear regression and/or a K-means method).

b. The Scientific Methods may have configuration parameters – *Method Configuration* (e.g. the name of the function/method, the filters to be applied to the dataset etc). A way of inputting these parameters should be provided – currently displayed as "free text box", but TBC with the developer.

The selected variables, methods and their configuration , will be sent to the MIP backend (as a query JSon Restful), and the results will be pushed back up to the Web UI to be displayed (see Display step - next).

5.3.1.4 3. Run Scientific Methods for the function and variables selected & plot-specific functions

This step will be done at MIP backend. The variables and configuration details (the query) specified at the Configuration step, will be sent to the MIP back-end where there will be computed. The results will be sent back to the Web UI through JSon Restful services.

5.3.1.5 3.1 Update Model instance

This will take place at MIP backend (CHUV team).

5.3.1.6 4. Display: 1) plot - 1st/default, 2) full stats & 3) Design Matrix. Default page: Design Matrix

The result is composed of three views, organised as tabs. Each view can potentially be exported in a file (format has to be define). The default view will be the Design Matrix.

Each method selected by the user in the Configuration step will have a default plot associated (for example, a Linear Regression makes most sense to be displayed in a BoxPlot). The user must be able to change the type of plot at the current "Display" step.
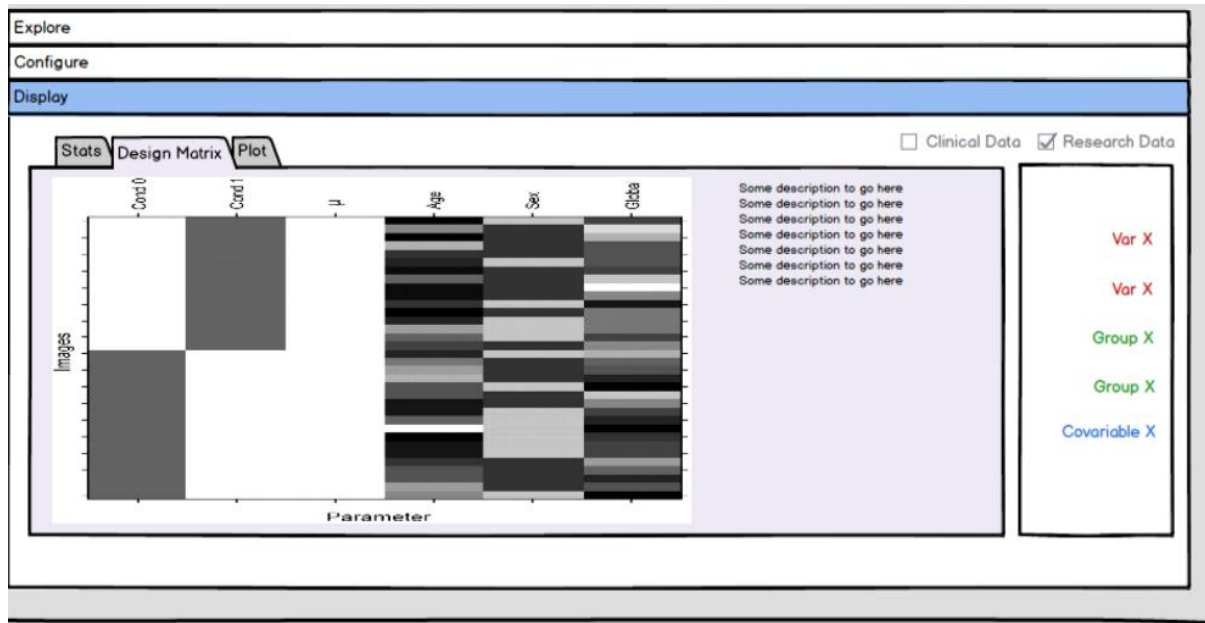
Description

Plots, Full stats and Design Matrices may be displayed for the different user types:

1) <u>Plots:</u> large public users, health management, clinicians or health writers
2) <u>Full stats:</u> statisticians (see table in process step 3)
3) <u>Design Matrix:</u> neuroscientists, scientists, clinicians

All these 3 visualisation options will be displayed in a tabular manner (as above). The default tab will be Design Matrix.

A visualisation set (plot + full stats + design matrix) will be displayed for each variable selected in the CONFIGURATION step, one above the other, i.e. for 2 variables selected there will be 2 visualisation sets. Nevertheless, for more methods and more data sources selected, there will be a larger number of display sets. For example, 2 variables, 2 functions and 5 data sources selected will mean 20 display sets. As this might be too much visually, a solution will be to allow user to hide (declutter) the screen by hiding results. Best method to be discussed with the developer.
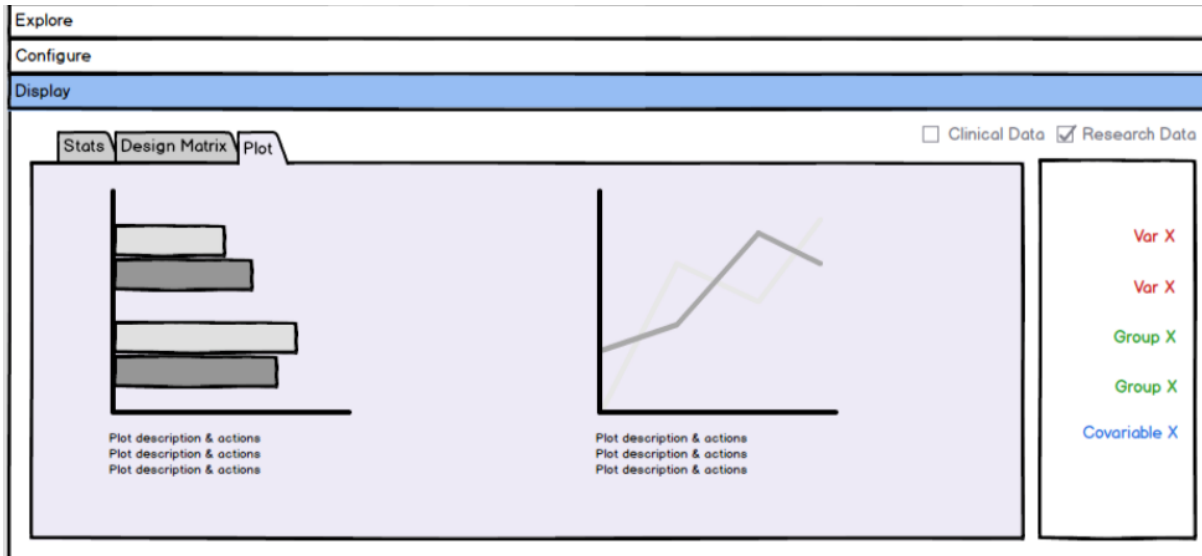
Actions at this step:

1. Switch between different views/tabs (plot, matrix, stats)
2. on the Plot tab, switch between plot types
3. Open Configuration accordion to further apply filters on the current variables
4. Open Exploration accordion to reselect variables
5. Retract/Hide visualisation sets

1) <u>**Plots**</u>

According to the configuration set by the user (especially the method/plot choice and its configuration), the plot is one of the result views.
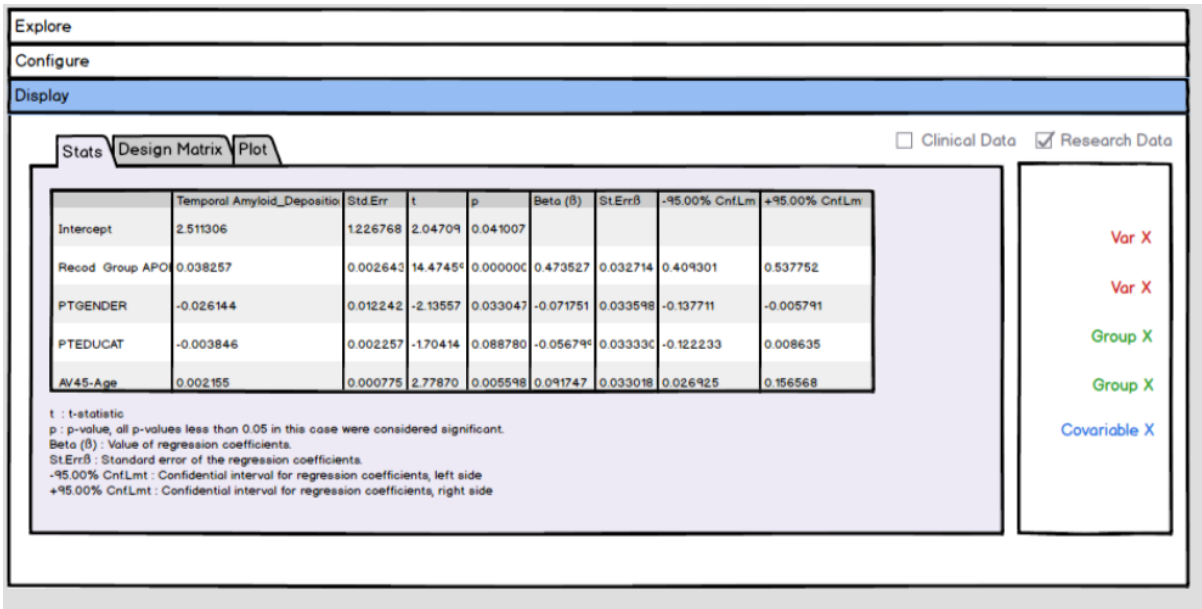
Concretely, the API returns a numerical matrix that the UI processes to present the data in the plot.

This part is already done for the POC, it will be reused here.

2) Full stats matrix
The resulting data matrix is also accessible.

3) Design matrix

This is the default view.

The Design Matrix = Model = representing the distribution of all the variables that influence the result of the study, i.e. age, education, gender, sleep etc., also called covariable or covariates.

The design matrix will be read only – so, it may even be an image generated at MIP backend (TBC). The matrix is based on the data displayed in the « Full stats matrix » screen.

Design matrix reference: https://en.wikipedia.org/wiki/Design_matrix
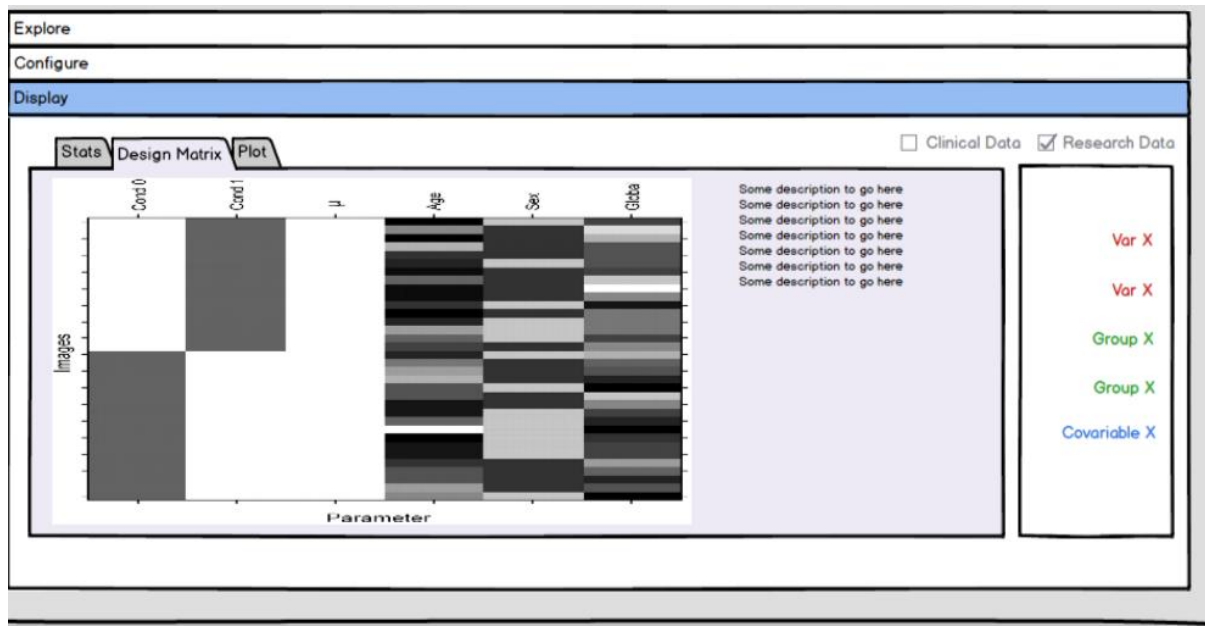
Example: for the use case of Beta amyloid deposition (AV45), all the covariables and groups will be displayed in rapport to the variable chosen (AV45), as columns. There will be no correlation between the covariables. The order of the variables selected matters (confirmed by BD: 28 Oct 2015), in the sense that the first one influences the ones following.

5.3.1.7 ◇ Refine model?

5.3.1.8 ▨ 5. User selects another plot type OR visualisation option
Description

The parameters of each possible plot are precalculated in the backend at the time when the user selects the method, and all are sent to the Web UI. Hence, the user can easily choose and switch between plot types at this step.

5.3.1.9 ▨ 5.1 User wants to reconfigure OR modify input variables
Description

On mouse-over, user gets information about the data displayed in the plot.
S/He may decide to refine the model/variables by db-clicking on the plot. Then follow step 10.2 of the process and go back to 1.

5.3.1.10 ◇ *Finished work?*

If "Yes", then:

5.3.1.11 ⬚ *6. User saves plot*

5.3.1.12 ⬚ *7. Save Design Matrix (Model)*

If the user wants to keep his generated model, he can save it. Everything is kept, the variables chosen, the variable and plot configuration.

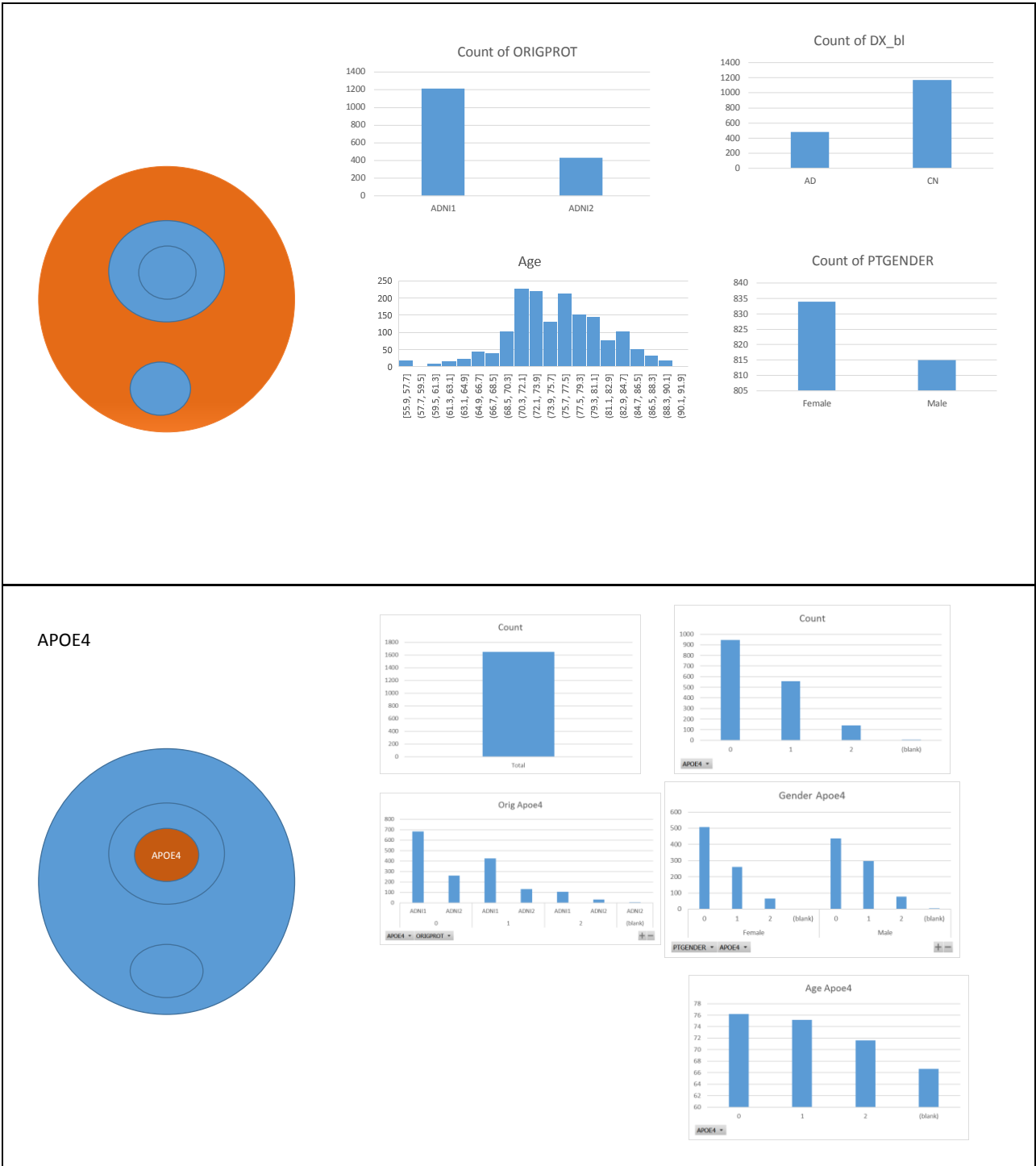He can reopen the model later to consult or to edit it.

If "No", then:

Go back to adjust the model (=rework from plot)

5.3.1.13 ⬤ *End*

Orange = element selected

APOE4

ADAS