

THE MEDICAL INFORMATICS PLATFORM

MIP User Guide

How to Select the Variables - 1

You can access the variables available in the MIP at any time by clicking on the "variables" button in the upper banner

1



The left panel presents all the variables available for the selected "use case", here the "use case dementia".

2

They are grouped by theme and each theme is represented by a circle. Here we have the circle which represents the anatomical data

3

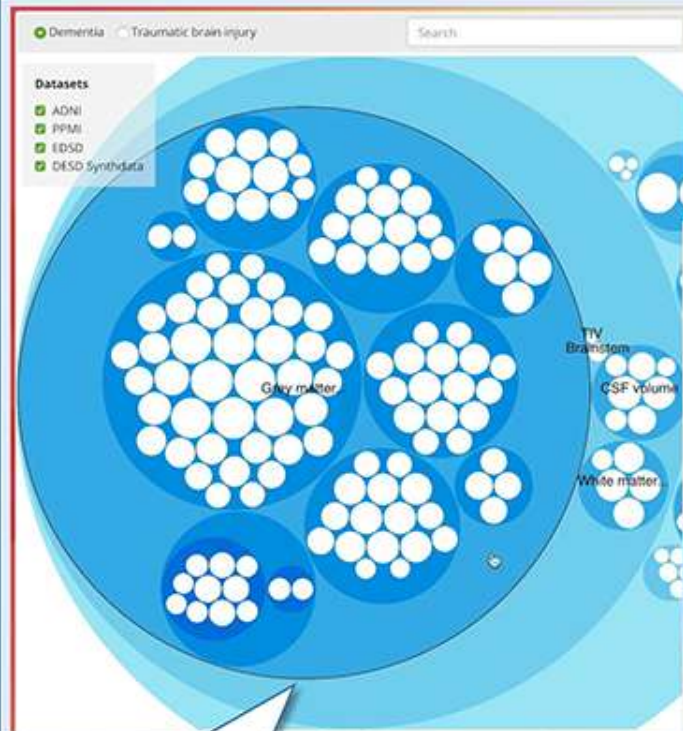
Demographic data

Diagnostic data

To move between these different themes, just click on a circle

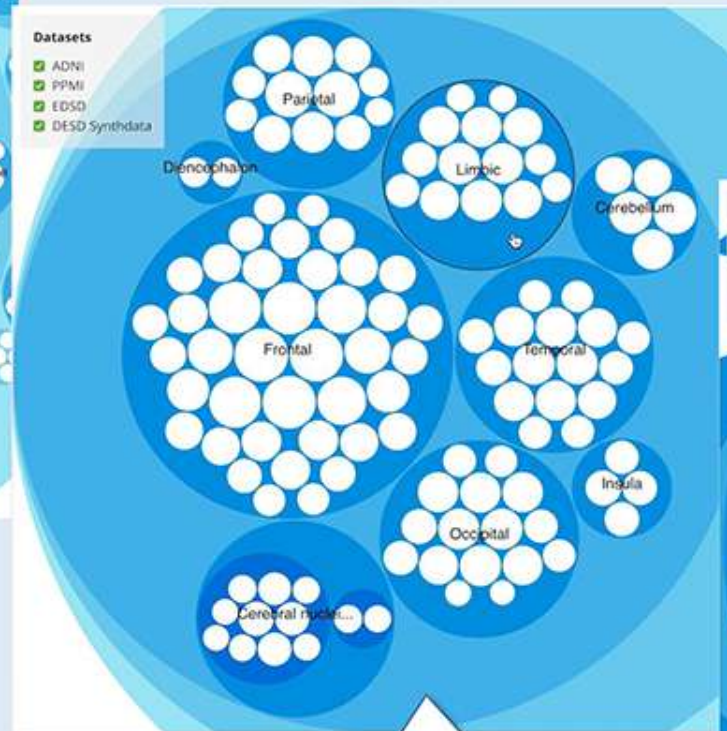


How to Select the Variables - 2



Here, anatomical data have been sub divided into gray matter and white matter

1



If you select "gray matter", the brain regions appear

2

And if you select a particular region, the limbic region for example, all the variables of this region appear

3



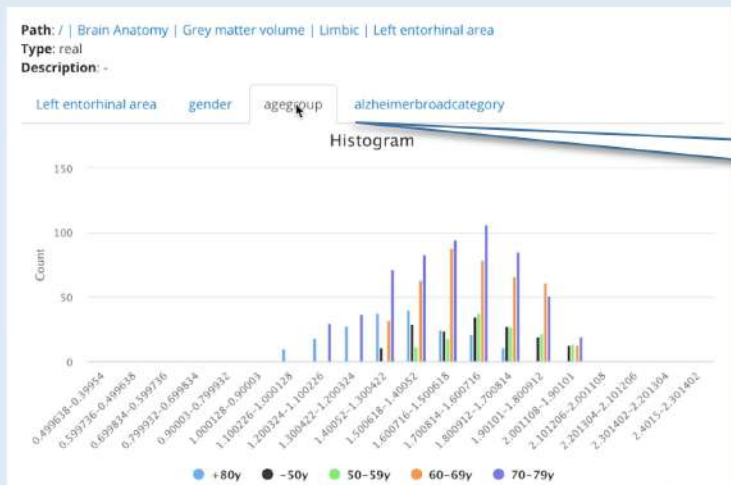
7

How to Select the Variables - 3



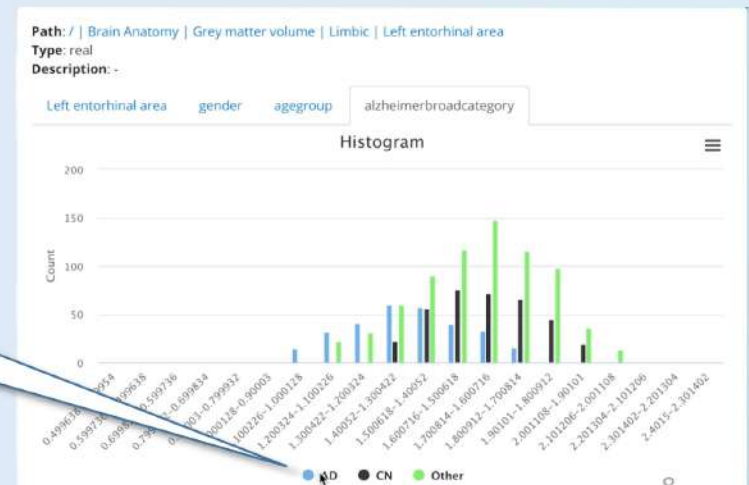
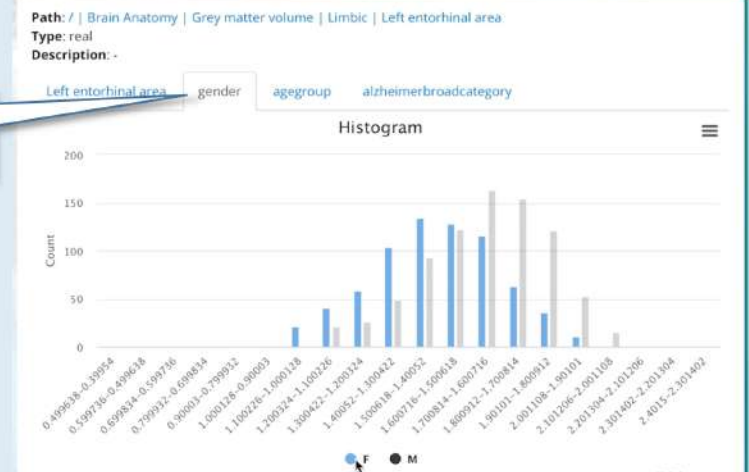
You can also have this distribution by gender

To select a variable, simply click on it. The panel at the bottom right updates immediately showing the distribution associated with the newly selected variable. By default, the distribution is based on all the data available in the selected data sets



Distribution by age group or according to Alzheimer's disease

Distribution of Alzheimer's patients, with no cognitive impairment, and who do not belong to the previous two categories.



How to Select the Variables - 4

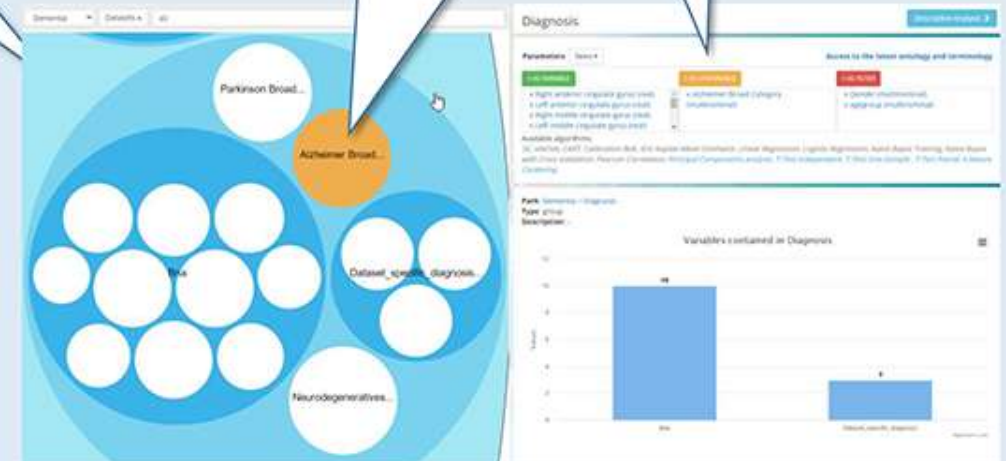
To perform an experiment, these variables are used in two ways: either as variables, that is to say observations or dependent variables; or we use them as co-variables, that is, predictors or independent variables

1 To add a variable to one of these two categories, just click on the associated button, for example here on the green "as variable" button. Automatically the circle of the selected variable changes color and becomes green

2 To select a variable from another theme, zoom out by clicking outside the circle

3 This allows you to select another circle. For example here the "diagnosis" part, and the variable "Alzheimer broad category"

4 This variable I will can be made a co- variable by simply clicking on the orange button



How to Analyse the Variables

At this stage of the study, you have selected the variables and have placed them either in the variables or covariables category

On the left we can see a summary of the selected datasets, here ADNI, CHUC CLM and BRESCIA. Below, we can see a summary of the selected variables and covariables. The variable is Left entorhinal area and covariables are Alzheimer Broad Category and Gender

In the center, a table presents a summary of the different variables by dataset, here ADNI, CLM and Brescia

For continuous variables like the volume of the left entorhinal area, we have the mean, the confidence interval as well as the standard deviation. For the Brescia dataset, the mean is 1.53, the confidence interval is 0.31 - 2.24, and the standard deviation is 0.23

The screenshot shows the MIP Interactive Analysis interface. On the left, there's a sidebar with 'Datasets' (ADNI, CHUC CLM, BRESCIA) and 'Variables' (Left entorhinal area, Alzheimer Broad Category, Gender). The main area displays a table with the following data:

VARIABLES	adni	clm	bfi
Left entorhinal area	1.53 (0.90-2.13) - std: 0.23	1.51 (0.56-2.06) - std: 0.23	1.53 (0.31-2.24) - std: 0.23
Alzheimer Broad Category	533	699	1784
Alzheimer's disease	111	59	151
Cognitively Normal	145	52	1240
Other	277	572	192
Mild cognitive impairment	0	36	201
Gender	533	699	1960
Male	288	344	768
Female	245	355	1194

If several categorical variables have been selected, the breakdown by level is made for each variable, independently of the other variables. In our example, for the ADNI dataset, we see that the 533 participants are distributed between the different levels of the variable Alzheimer Broad Category and the variable Gender

For categorical variables, the interface presents the workforce for each level of the variable. Thus for the CLM dataset and the Gender variable, we see that there are 699 participants, 344 men and 355 women

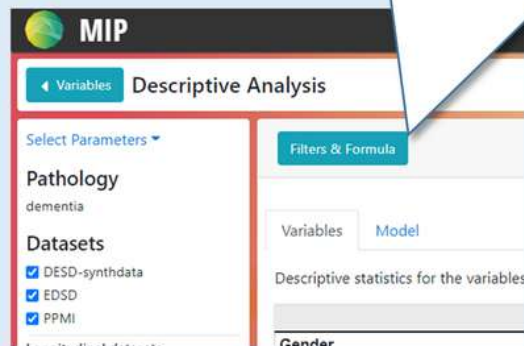
6

Finally at this stage, it is possible to modify the datasets used, for example by adding EDSD. The central table will be updated to include the information related to this new dataset

How to Filter & Save a Variable

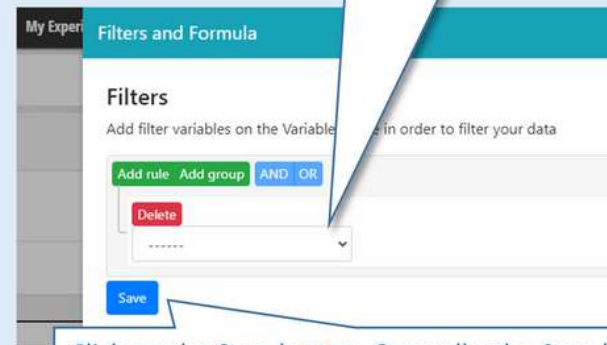
We are now going to focus on data filtering. For this you must click the button Filter & Formula, top left of the central panel. A new panel appears, it will allow to apply filters on the variables

1



To work only on participants over the age of 60, click on the empty field.

2



Click on the Save button. Generally, the Save button allows you to take into account the changes in the filter. We then see that the description of the variables is updated to consider the new filter

5

Filters
Add filter variables on the Variables page in order to filter your data

Add rule Add group AND OR

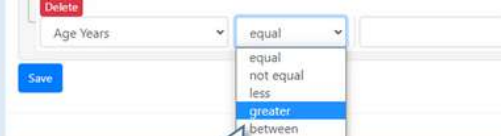


A drop-down menu appears where we find our variables. We select Age Years

3

Filters
Add filter variables on the Variables page in order to filter your data

Add rule Add group AND OR



In the next field select Greater. In the last field type 60

4

Variables
Left entorhinal area (real)

CoVariables
Alzheimer Broad Category
(polynomial)

Filters
Age Years > 60

Result. The filter applied is also noted under the variables

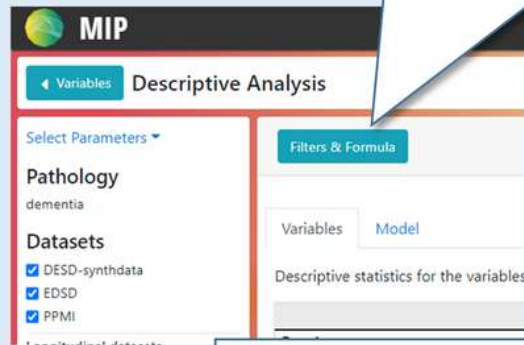
6

How to use Formula and Interactions

We are now going to focus on formulas. For this you must click the button Filters & Formula, top left of the central panel. A new panel appears, it will allow to apply transformation on a variable or add interaction between variables

In the next steps we will apply a logarithm transformation to the variable Left entorhinal area as an example so click on the field Variable...

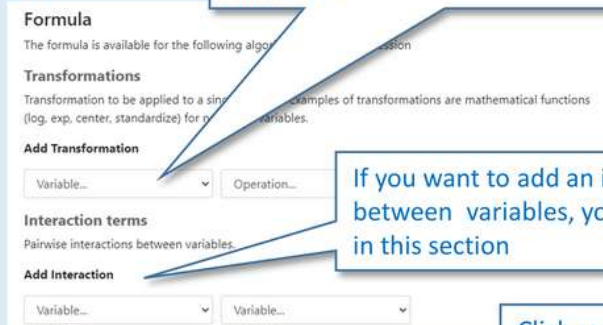
Finally, just click on the create experiment button to go to the next step



1

2

8



If you want to add an interaction between variables, you can do it in this section

Click on the tab Model to see the result of transformations and interactions

6

Formula

The formula is available for the following algorithm

Transformations

Transformation to be applied to a single variable. Examples of transformations are mathematical functions (log, exp, center, standardize) for numerical variables.

Add Transformation

Left entorhinal area

In the next field select the transformation you want to apply, for example log

4

A drop-down menu appears where we find our variable. We select Left entorhinal area

3

Click on the + button to save and apply the transformation

5

Intersection table for the variables of interest as it appears in the experiment.

	ppmi
log(leftentorhinalarea)	426
Datapoints	426
Nulls	0
std	0.1
max	0.74
min	0.22
mean	0.51

(nominal)
Gender (nominal)

Filters

Age Years > 60

Formula

Transformations

log: leftentorhinalarea

The summary of transformations and interactions applied are also noted under Filters

7

How to Perform a T-Test - 1

After clicking Run Experiment this screen is presented

At this stage of the study, we have selected the variables, and we have finished the filtering stage . The summary of the previous steps can be found in the panel to the left of the interface

1

The screenshot shows the MIP web interface. The top navigation bar includes 'Variables > Analysis > Experiment', 'My Experiments', 'Workflow', 'Articles', 'Profile', and 'Help'. The main header has a 'Create Experiment' button and a 'Run Experiment' button. The left sidebar contains sections for 'Datasets' (ADNI, PPMI, EDS, CHUV CLM, BREF), 'Variables' (Left entorhinal area (real)), 'CoVariables' (Alzheimer Broad Category (polynomial)), and 'Filters' (Age Years > 60 AND Alzheimer Broad Category not_equal MCI). The main content area has a 'Your algorithm' section with a prompt to select an algorithm from the 'Available Algorithms' panel. The 'Available Algorithms' panel lists: LINEAR_REGRESSION, ID3, LOGISTIC_REGRESSION, PEARSON_CORRELATION, TTEST_INDEPENDENT, ANOVA, and Naive Bayes with Cross Validation. The 'TTEST_INDEPENDENT' algorithm is highlighted in blue.

To carry out an independent T-test, the dependent variable must be of real type, and the independent variable of categorical type. Here, the dependent variable is Left entorhinal area , and the independent variable is Alzheimer Broad Category

2

The right panel presents the algorithms compatible with these variables in blue, the others are grayed out. To carry out a T-test, click on TTest_independant

3

How to Perform a T-Test - 2

← Review

Create Experiment

RELATED EXPERIMENTS ▾

Experiment name

Run Experiment ▶

Datasets
ADNI PPMI EDSO
CHUV CLM BRESCIA

mdi (4) ▾
Variables
Left entorhinal area (real)
CoVariables
Alzheimer Broad Category (polynomial)
Filters
Age Years > 60
AND
Alzheimer Broad Category not_equal MCI

Algorithm About Running Experiments

TTEST_INDEPENDENT
Student's Independent samples t-test
Parameters
The two levels of the grouping variable (Order matters):
xlevels

'different', 'greaterthan' or 'lessthan', the alternative hypothesis: group 1 different to group 2, group hypothesis

1 or 0, provide Cohen's d effect sizes
effectsize

Select... ✓
Alzheimer's disease
Cognitively Normal ✓
Other ✓
Mild cognitive impairment

1 ✓
Not blank, Integer

1 ✓
Not blank, Integer

1 ✓
Not blank, Integer

Available Algorithms
LINEAR_REGRESSION
ID3
LOGISTIC_REGRESSION
PEARSON_CORRELATION
TTEST_INDEPENDENT
ANOVA
Naive Bayes with Cross Validation

The central panel now presents the parameters of the algorithm. The first asks to enter the two levels of the Alzheimer Broad Category variable to compare. By clicking in the empty field, a drop-down menu appears with the levels of the variable. If we want to compare the volumes of the entorhinal area for the supposed Alzheimer participants and the patients without cognitive impairment. We will therefore select the two modalities Alzheimer's disease and Cognitively Normal

14

How to Perform a T-Test - 3

← Review Create Experiment

Datasets
☒ ADNI ☐ PPMI ☒ EDSO
☒ CHUV CLM ☒ BRESCIA

Variables
Left entorhinal area (real)

CoVariables
Alzheimer Broad Category (polynomial)

Filters
Age Years > 60
AND
Alzheimer Broad Category not_equal MCI

mdl (4)

Algorithm About Running Experiments

TTEST_INDEPENDENT
Student's independent samples t-test

Parameters

The two levels of the grouping variable (Order matters)
xlevels

Alzheimer's disease x Cognitively Normal x ✓
Not blank

'different', 'greaterthan' or 'lessthan', the alternative hypothesis: group 1 different to group 2, group 1 greater than group 2, and group 1 less than group 2 respectively
hypothesis

different ✓
Not blank

1 or 0, provide Cohen's d effect sizes
effectsize

1 ✓
Not blank, Integer

1 or 0, provide confidence intervals for the mean difference
ci

1 ✓
Not blank, Integer

1 or 0, provide means and standard deviations
meandiff

1 ✓
Not blank, Integer

Once the algorithm has been configured, you must name it to be able to execute it 3

RELATED EXPERIMENTS = ttest Run Experiment >

Available Algorithms
LINEAR_REGRESSION
ID3
LOGISTIC_REGRESSION
PEARSON_CORRELATION
TTEST_INDEPENDENT
ANOVA
Naive Bayes with Cross Validation

The second parameter allows you to select the type of comparison: greater than, different or less than. By default, the value is set to different 1

The three parameters regarding not displaying results, a value of 1 (default) displays the results 2

How to Perform a T-Test - 4

Results of Experiment **ttest** on **mdl-4**

Created a few seconds ago by phenixt

[SHARE EXPERIMENT](#)[RELATED EXPERIMENTS](#)

Datasets

Training datasets

adni
clm
fbf
edsd

mdl (4)

Variables

Left entorhinal area (real)

CoVariables

Alzheimer Broad Category
(polynomial)

Filters

Age Years > 60
AND
Alzheimer Broad Category
not_equal MCI

Algorithm

TTEST_INDEPENDENT

Parameters

xlevels: AD,CN
hypothesis: different
effectsize: 1
ci: 1
meandiff: 1

TTEST_INDEPENDENT

INDEPENDENT_TEST_TABLE

colname	statistics	df	Hypothesis	Meandifference	SSEdifference	Lower	Upper	Cohens_d
leftentorhinalarea	-16.818	955	0	-0.227	0.014	-0.254	-0.201	-1.121

The results presented are the standard results of this type of analysis: statistics, degrees of freedom, mean difference as well as its confidence interval and Cohen's distance

The lower left panel presents a summary of the parameters used in the algorithm

MIP displays this message whilst the experiment is running

TTEST_INDEPENDENT

Your experiment is currently running...

Please check back in a few minutes. This page will automatically refresh once your experiment has finished executing.