

Technical Working Party on Automation and Computer Programs **TWC/38/7**

Thirty-Eighth Session
Alexandria, United States of America, September 21 to 23, 2020

Original: English
Date: August 20, 2020

PRESENTATION OF PATHOSTAT APPLICATION

Document prepared by an expert from France

Disclaimer: this document does not represent UPOV policies or guidance

This document contains a copy of a presentation prepared by an expert from France, to be made at the thirty-eighth session of the Technical Working Party on Automation and Computer Programs (TWC), as follows:

- Annex I: Presentation on "Presentation of PATHOSTAT application"
- Annex II: Pathostat user guide

[Annexes follow]

Presentation of PATHOSTAT application

Technical Working Party On Automation And
Computer Programs



Presentation of application

- Decision support system
- R Shiny app → Interactive web application
⇒ <https://pathostat.geves.fr>

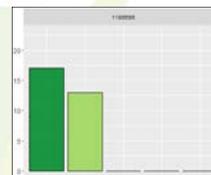
- Pathostat-Veg Project



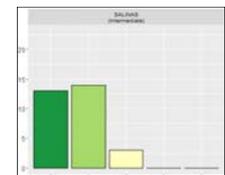
- Objectives :

- Support expert analysis with use of resistance test to diseases

LIMONE	1	3	7	0	0	0
	2	5	5	0	0	0
	3	3	6	0	0	0
PATRIOT (Susceptible)	1	0	0	3	4	3
	2	0	0	1	2	7
	3	0	0	3	1	7
NATEXIS (Intermédiaire)	1	0	4	2	3	0
	2	0	4	3	3	0
	3	0	5	4	1	0
SALINAS (Intermédiaire)	1	4	6	0	0	0
	2	4	4	2	0	0
	3	5	4	1	0	0
COSTARICA (Resistant)	1	7	3	0	0	0
	2	6	4	0	0	0
	3	10	0	0	0	0



More or less susceptible ?



Groupe d'Étude et de contrôle
des Variétés Et des Semences

Presentation of application

- Objectives :
 - Select statistical tests applicable to the quantitative results of bio-tests
 - For DUS:
 - Identify if varieties are distinct based on non qualitative data
 - Provide proof that it is uniform/homogeneous
- Purpose :
 - Improvement and harmonisation of decision rules on the interpretation of biotest results
 - Improvement of variety judgement
 - Better correlation between applicants' claims and official tests

Presentation of application



- Free hosting and open access on GEVES website

⇒ <https://pathostat.geves.fr>



- Application is secure

⇒ Data are not saved within application



- User manual

⇒ Exists in English and French



- Application has 2 functions :

⇒ The respect of the regulatory framework and official protocol

⇒ Adaptability for methodology or research



Statistical Methods

Monday, 10 August 2020

WWW.GEVES.FR



Statistical Methods

● Test on the minimum number

Variety	Number \geq threshold	Class 0	Class 1	Class 2	Class 3	Class 4
VARIETY A	✓	4	11	8	6	1
VARIETY B	X	22	4	0	1	0
CASAQUE ROUGE (Susceptible)	✓	0	0	0	3	27
ANAHU X CASAQUE ROUGE (Resistant)	✓	23	7	0	0	0
CAMPEON (Intermediate)	✓	7	3	6	15	0
TYONIC (Intermediate)	✓	4	10	9	7	0

- Minimum threshold number depends on pathosystem (here 30 plants)
- Reference threshold number is automatically selected
- Option to change the value of the minimum threshold

Statistical Methods

● Comparison of distribution between repetitions

– Stacked histogram

- To compare graphically distribution of symptoms between repetitions for each variety
- Highlights differences between repetitions

Distribution of symptoms for repetition 1 of variety 2019.13 appears different from others



Statistical Methods

● Comparison of distribution between repetitions

– Homogeneity test of repetition

- To study the distribution of symptoms between repetitions
- All varieties are pooled together
- Highlights any problems in experimental protocol

➔ Chi-squared test (χ^2 test)

- Condition for the validity:
 - Expected value ≥ 5 in at least 80 % of cases
- Hypothesis of test :
 - H_0 : homogeneity of the distribution of symptoms between repetitions
 - H_1 : heterogeneity of the distribution of symptoms between repetitions

• Chi-squared test

Pearson's Chi-squared test
data: pooled data
X-squared = 3.0944, df = 6, p-value = 0.7969



Statistical Methods

● Comparison of distribution between repetitions

– Homogeneity test of repetition per variety

➔ Fisher's exact test

- To compare distribution between repetitions of each variety
- Preferred to χ^2 test because:
 - The small number of each repetition
 - The predictable imbalance of the expected value
- Hypothesis of test:
 - H_0 : homogeneity of the distribution of symptoms between repetitions
 - H_1 : heterogeneity of the distribution of symptoms between repetitions

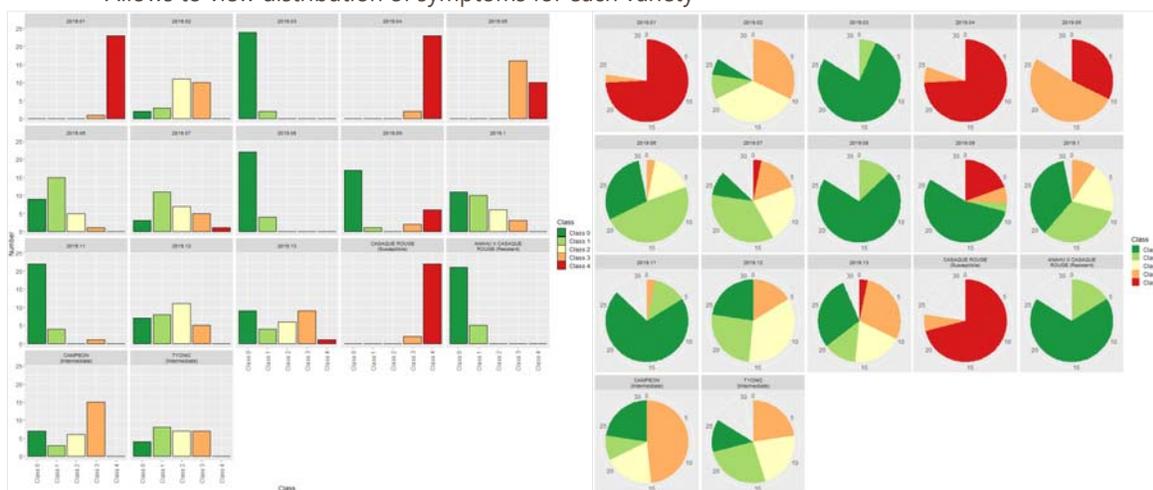
Variety	P-value	Result
2019.01	1.000	Homogeneity of repetitions
2019.02	0.062	Homogeneity of repetitions
2019.03	0.508	Homogeneity of repetitions
2019.04	1.000	Homogeneity of repetitions
2019.05	1.000	Homogeneity of repetitions
2019.06	0.095	Homogeneity of repetitions
2019.07	0.180	Homogeneity of repetitions
2019.08	0.335	Homogeneity of repetitions
2019.09	0.513	Homogeneity of repetitions
2019.1	0.262	Homogeneity of repetitions
2019.11	0.141	Homogeneity of repetitions
2019.12	0.749	Homogeneity of repetitions
2019.13	0.020	Heterogeneity of repetitions
CASAQUE ROUGE (Susceptible)	1.000	Homogeneity of repetitions
ANAHU X CASAQUE ROUGE (Resistant)	1.000	Homogeneity of repetitions
CAMPEON (Intermediate)	0.118	Homogeneity of repetitions
TYONIC (Intermediate)	0.699	Homogeneity of repetitions

Statistical Methods

● Study of the susceptibility/resistance of varieties

– Histogram – Pie chart

- Allows to view distribution of symptoms for each variety



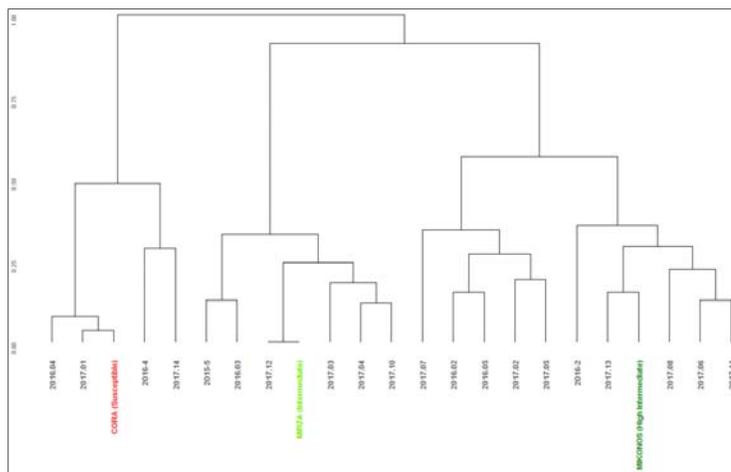
Statistical Methods

● Study of the susceptibility/resistance of varieties

– Hierarchical Cluster Analysis

- Principle of the method :
 - intraclass homogeneity
 - interclass heterogeneity
- Criterion of similarity : relative frequency of affiliation to each class of symptom
- Based on Euclidean distance

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$
- Dendrogram (classification tree)
- 5 agglomerative methods
- Cophenetic correlation coefficient



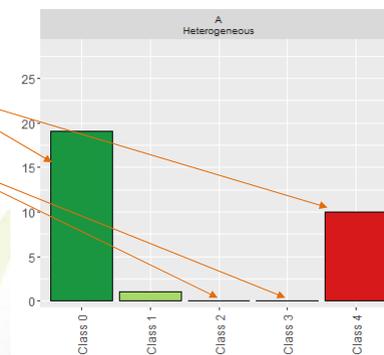
Statistical Methods

● Study of the susceptibility/resistance of varieties

– Decisions

➔ Heterogeneity rule

- Variety is proposed heterogeneous if there are:
 - Plants in both extreme classes within the limit of the tolerated number of off-types
 - No plant in at least one of the intermediate classes
- The tolerated number of off-types of extreme classes depends on:
 - Species
 - Number of plant in extreme classes of control



Statistical Methods

● Study of the susceptibility/resistance of varieties

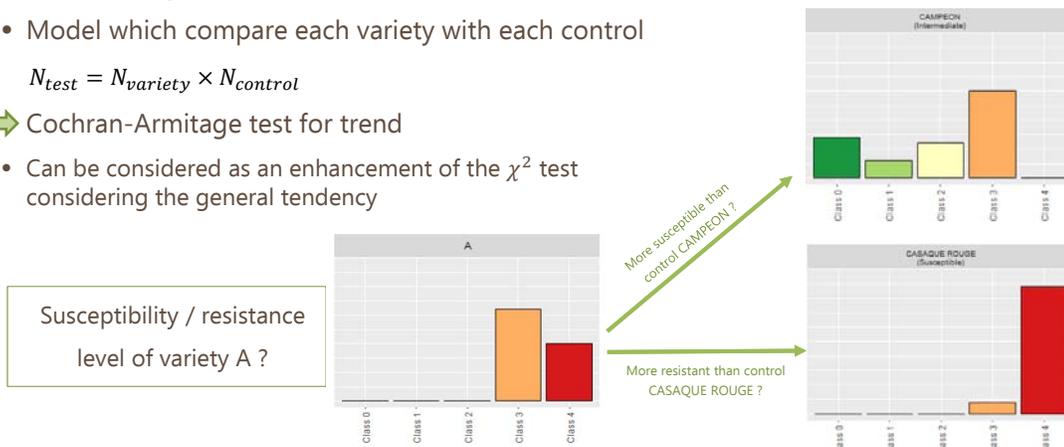
– Decisions

- Susceptibility / resistance level of varieties decided on comparison with control
- Model which compare each variety with each control

$$N_{test} = N_{variety} \times N_{control}$$

➡ Cochran-Armitage test for trend

- Can be considered as an enhancement of the χ^2 test considering the general tendency



Statistical Methods

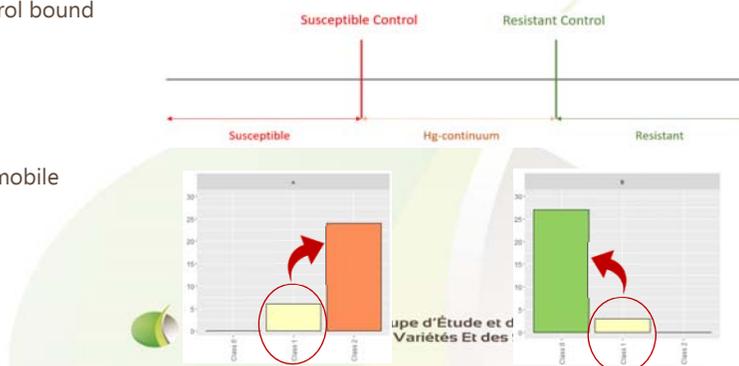
● Study of the susceptibility/resistance of varieties

– Decisions

➡ Cochran-Armitage test for trend

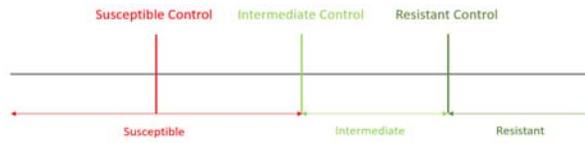
- No condition for the validity
- CPVO decision rules will determine interpretation of the variety in comparison with controls
 - No control bound

– Class 1 mobile



Statistical Methods

- IR control lower bound



- Resistant control lower bound

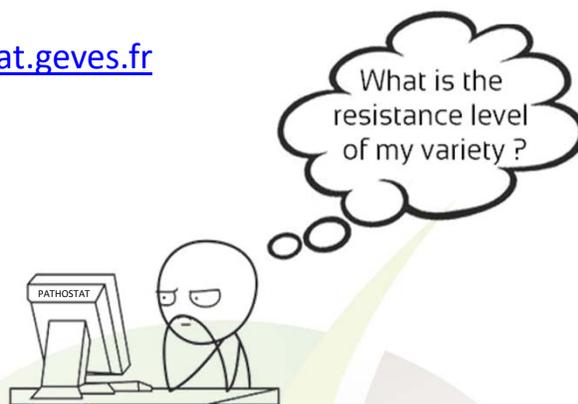


- Susceptible control higher bound



Demonstration of application

⇒ <https://pathostat.geves.fr>



Any Question ?

Thank your for listening !

*ANY
QUESTIONS
...*



So, it is over to you !

<https://pathostat.geves.fr>



 **GEVES** | Groupe d'Étude et de contrôle
des Variétés Et des Semences

[Annex II follows]

PATHOSTAT Application

USER MANUAL



QUEMAR Thibaud

GRUPE D'ÉTUDE ET DE CONTRÔLE DES VARIÉTÉS ET DES SEMENCES

Table of contents

Presentation of application	3
Selection and parametrization of the biotest	3
Selection of pathosystem	3
Entry spreadsheet	3
Loading of resistance test.....	4
Minimum threshold number	5
Selection of controls.....	7
Comparison of distribution between repetitions.....	8
Loading of a biotest without repetition	8
Merging of classes	9
Stacked histogram	10
Homogeneity regardless of the variety.....	11
Homogeneity for each variety.....	12
Study of the susceptibility / resistance of varieties	12
Class layout.....	13
Graphical representation	13
Dendrogram	15
Decisions.....	16
Save results.....	20
Report in PDF format.....	21
Report in Word format.....	27

Presentation of application

The Pathostat application, developed within the framework of the eponym project PATHOSTAT-Veg, is a decision support system which allow for statistics integration to the results of resistance test to diseases, for registration of vegetable species.

First, to use Pathostat application you just need to connect to the following address: <https://pathostat.geves.fr>

⚠ After a while of extended inactivity a connection loss toward application is possible.

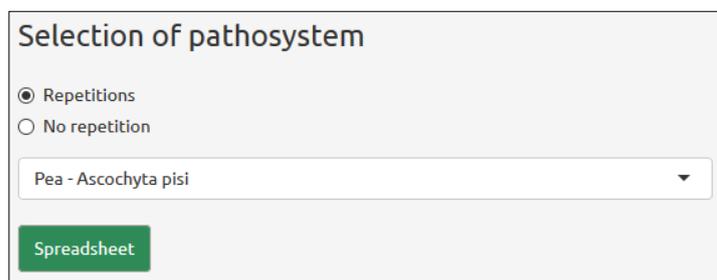
The latter is constituted of 3 tabs respectively entitled:

- Selection and parametrization of the biotest
- Comparison of distribution between repetitions
- Study of the susceptibility / resistance of varieties

Selection and parametrization of the biotest

First tab allows to load our results, as well as to configure statistical analysis corresponding to the data set, which will be perform.

Selection of pathosystem



The screenshot shows a web interface titled "Selection of pathosystem". It contains two radio buttons: "Repetitions" (which is selected) and "No repetition". Below the radio buttons is a dropdown menu with the text "Pea - Ascochyta pisi" and a downward arrow. At the bottom of the interface is a green button labeled "Spreadsheet".

First frame in the top left-hand corner of this tab is related to the choice of pathosystem. It has a dual role:

- To download an Excel spreadsheet (.xlsx). This part is optional if we already have a data set to load.
- To apply by default, parameters and statistical tests related to pathosystem selected in the drop-down list according to the CTPS decision rule applicable when application is used.

⚠ Application is already configured according to the CTPS decision rules applicable when application is used. Nevertheless, all parameters modification can lead to a bias in relation to the interpretation which can be done within the framework of an official test.

Entry spreadsheet

Then two options are available for us : either we already have a resistance test model in the right format of the application (an Excel file with the right column names in English), or it's also possible to download a blank entry spreadsheet to fill oneself afterwards.

If we don't have an Excel file in the right format, we can download a blank entry spreadsheet. To achieve this, all we must do in a first time is to choose if we want to download a spreadsheet model with or without repetitions. It should be noted that option « Repetitions » or « No repetition » only affect the downloading of the spreadsheet, and in no case determine if comparison of distribution between repetition will be carried out. Therefore, it's possible to download a test with repetition even if option « No repetition » is ticked.

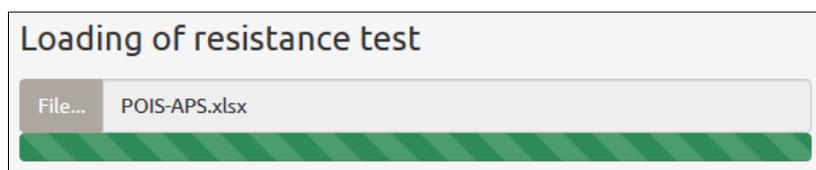
Then, once we have chosen the pathosystem related to our resistance test in the drop-down list (here a Pea – *Ascochyta pisi* resistance test), we can download a blank entry spreadsheet. The blank spreadsheet will be presented in that form:

	A	B	C	D	E	F
1	Variety	Repetition	Class 0	Class 1	Class 1 : some little superficial necrosis	Class 3
2		1				
3		2				
4	MDK (Susceptible)	1				
5		2				
6	MADONNA (Resistant)	1				
7		2				
8	NINA (Resistant)	1				
9		2				

The Excel spreadsheet only have one empty cell to add a variety, nevertheless it's possible to add manually as many rows as we have varieties. Once the spreadsheet is filled, it's necessary to save it on his computer and to load it within application in order to the integrate the results.

If we want to use our own resistance test file, it's primordial that the resistance test has a column « Variety » with all controls used for the test. « Repetition » column is optional, nevertheless it allows to determine logically if tests will be performed in the tab « Comparison of distribution between repetitions ». Columns related to various class of symptom are commented to get information as for definition of each class.

Loading of resistance test



The following step is the loading of the resistance test in the Excel format. This phase can be undertaken separately from the downloading of the spreadsheet. Nevertheless, it's essential to observe in the Excel table, columns « Variety » and « Repetition » (in the case of a test with repetition) because these strings then will be literally used. As soon as biotest is loaded, it comes out as it is in the right window:

Variety	Repetition	Class 0	Class 1	Class 2	Class 3
V1	1	3	5	2	0
V1	2	8	1	1	0
V2	1	0	0	1	9
V2	2	0	0	3	6
MDK (Susceptible)	1	0	0	0	10
MDK (Susceptible)	2	0	0	0	9
MADONNA (Resistant)	1	8	2	0	0
MADONNA (Resistant)	2	1	9	0	0
NINA (Resistant)	1	10	0	0	0
NINA (Resistant)	2	9	1	0	0

⚠ One must be careful to make sure than in the Excel file loaded, cells with various numbers are correctly interpreted as an integer and not as text. Moreover, it's necessary to be sure the most resistant symptom class is the first one of the loaded biotest. For instance, here class 0 is the most resistant, while class 3 is the most susceptible.

Minimum threshold number

Minimum threshold number

20
⌵

Don't apply any threshold
 Inform about varieties for which number < threshold
 Pull out varieties for which number < threshold

Then, you need to decide if you want to apply a minimum threshold number to your data set. The value of this threshold is defined by pathosystem selected in the drop-down list above. The default values of threshold defined according to the CTPS rules applicable when you use application are:

Pathosystem	Threshold
Bean - <i>Colletotrichum lindemuthianum</i> race 6	20
Bean - <i>Colletotrichum lindemuthianum</i> race Kappa	20
Bean - <i>Pseudomonas savastanoi</i> pv <i>phaseolicola</i> race 6	20
Cabbage - <i>Fusarium oxysporum</i> f. sp. <i>conglutinans</i>	20
Cucumber - <i>Cladosporium cucumerinum</i>	30
Cucumber - CMV	30
Cucumber - <i>Corynespora cassicola</i>	30
Cucumber – CVYV	15
Cucumber - <i>Podosphaera xanthii</i>	20
Lettuce - <i>Bremia lactucae</i>	20
Lettuce - <i>Fusarium oxysporum</i> f.sp. <i>lactucae</i>	20
Melon - <i>Fusarium oxysporum</i> f. sp. <i>melonis</i> race 0	30

Melon - <i>Fusarium oxysporum</i> f. sp. <i>melonis</i> race 1.2	30
Melon - <i>Fusarium oxysporum</i> f. sp. <i>melonis</i> race 1	30
Melon - <i>Fusarium oxysporum</i> f. sp. <i>melonis</i> race 2	30
Melon - MWMV	30
Melon - <i>Podosphaera xanthii</i> race 1, 2, 3, 5, 3-5	20
Melon - ZYMV	30
Pea - <i>Ascochyta pisi</i>	20
Pea - BYMV	20
Pea – <i>Erysiphe pisi</i>	20
Pea - <i>Fusarium oxysporum</i> f.sp <i>pisii</i> race 1	20
Pepper - <i>Meloidogyne incognita</i>	20
Tomato - <i>Fusarium oxysporum</i> f.sp. <i>lycopersici</i> race 0	20
Tomato - <i>Fusarium oxysporum</i> f.sp. <i>lycopersici</i> race 1	20
Tomato - <i>Fusarium oxysporum</i> f.sp. <i>lycopersici</i> race 2	20
Tomato - <i>Fusarium oxysporum</i> f.sp. <i>radicis lycopersici</i>	20
Tomato - <i>Meloidogyne incognita</i>	30
Tomato - <i>Pseudomonas syringae</i> pv. <i>tomato</i>	20
Tomato - <i>Pyrenochaeta lycopersici</i>	20
Tomato - <i>Stemphylium</i> spp.	20
Tomato - <i>Verticillium dahliae</i>	20
Zucchini – CMV	20
Zucchini – WMV	20
Zucchini – ZYMV	20

⚠ Some resistance tests exist only without repetitions ! It's especially the case for tests Cucumber – *Cladosporium cucumerinum*, Cucumber – CMV and Cucumber - *Corynespora cassiicola*.

This default value of the threshold can afterwards be changed, either typing directly a new value of threshold, or using upwards and downwards arrows.

3 options are presented to you:

- Don't apply any threshold: by default, application don't apply any threshold.
- Inform about varieties for which number < threshold: variety names which are under-staffed are only red-coloured. Moreover, a column called « Number ≥ threshold » comes out informing if variety has a sufficient number (✓) or not (X). Nevertheless, the test of number is not carried out per repetition, but on the whole of the variety. Then, if a variety has an

Variety	Number ≥ threshold	Repetition	Class 0	Class 1	Class 2	Class 3
V1	✓	1	3	5	2	0
V1	✓	2	8	1	1	0
V2	X	1	0	0	1	9
V2	X	2	0	0	3	6
MDK (Susceptible)	X	1	0	0	0	10
MDK (Susceptible)	X	2	0	0	0	9
MADONNA (Resistant)	✓	1	8	2	0	0
MADONNA (Resistant)	✓	2	1	9	0	0
NINA (Resistant)	✓	1	10	0	0	0
NINA (Resistant)	✓	2	9	1	0	0

insufficient number on each repetition, it's in spite of everything the total number which is too small.

- Pull out varieties for which number < threshold: under-staffed varieties are pulled out of the resistance test.

Variety	Repetition	Class 0	Class 1	Class 2	Class 3
V1	1	3	5	2	0
V1	2	8	1	1	0
MADONNA (Resistant)	1	8	2	0	0
MADONNA (Resistant)	2	1	9	0	0
NINA (Resistant)	1	10	0	0	0
NINA (Resistant)	2	9	1	0	0

⚠ Pull out varieties can also be equivalent to remove one or several controls (here for instance susceptible control MDK), well presence of these ones is essential to decide if a variety is susceptible or resistant. In this way, you should take care regarding this!

Selection of controls

Selection of controls

Susceptible controls

MDK (Susceptible)

Intermediate resistant controls

Strong intermediate resistant controls

Resistant controls

MADONNA (Resistant) NINA (Resistant)

Then, you need to determine what are the controls to use for the resistance test. This step is important because it's the susceptibility level of the controls which establish susceptibility / resistance of each variety. According to the resistance test, you don't need to use each kind of control. High intermediate resistant control and resistant control are interpreted in the same way: they are controls most resistant than the intermediate resistant control.

By default, controls are looking for among various varieties according to regular expressions. Thus, for each variety, if there is one of the following regular expression, variety will be judged as a control variety:

- Susceptible controls: regular expressions are « Susceptible » with or without majuscule and « SCn » with *n* an integer. In other words, all varieties with these terms will be automatically identified as a susceptible control by default.

- Resistant controls: regular expressions are « Resistant » with or without majuscule and ended by a closing parenthesis «) » or at the end of the name of the variety. We meet also regular expression « RCn » with n an integer. In other words, all varieties with these terms will be automatically identified as a resistant control by default.
- High intermediate resistant controls: regular expressions are « High » with or without majuscule. We meet also regular expression « HIR » ended by a closing parenthesis «) », an integer n , or at the end of the variety name. In other words, all varieties with these terms will be automatically identified as a high intermediate resistant control by default.
- Intermediate resistant controls: regular expressions are « IR » ended by a closing parenthesis «) », an integer n , or at the end of the variety name. We also meet regular expression « Intermediate » with or without majuscule and ended by a closing parenthesis «) » or at the end of the variety name. This also applies to regular expression « Low » with or without majuscule (this regular expression is used for identification of a low intermediate resistant control). In other words, all varieties with these terms will be automatically identified as an intermediate resistant control by default. Furthermore, a variety is judged as an intermediate resistant control if it was not already found as resistant control or high intermediate resistant control.

⚠ Some varieties could be found wrongly as a control. Hence, it is important to be careful to what are controls identified by default! As an example, varieties called « AIR », « Highlander » and « DISC3 » will be respectively found intermediate, high intermediate and susceptible.

Then, it is possible to add or remove control varieties. In this way, we can remove varieties detected as control even though they are not and conversely adding as a control a variety unidentified as such. With this feature, we could also test several potential controls.

Comparison of distribution between repetitions

Once biotest is loaded and parametrization is achieved, we can move to the next tab entitled « Comparison of distribution between repetitions ». The aim of this tab is to test existence of the homogeneity of the distribution of symptom between each repetition. This tab is itself composed of 3 sub-tabs:

- Stacked histogram
- Homogeneity regardless of the variety
- Homogeneity for each variety

Loading of a biotest without repetition

You need to know that the whole of tests performed at this tab are optional. Indeed, if a resistance test without repetition (in other words which don't have a column « Repetition ») is loaded, no statistic test will be carried out. Then, the following error message will come out in the window of this tab:

⚠ Biotest loaded doesn't contain repetitions !

Merging of classes

No merging
 Merging of classes

Class 0 - Class 1

Apply

Graphic settings

Size labels

10

Text size of axis

12

Legend size

12

⚠ As long as biotest comprises a « Repetition » column, application will try to perform tests of homogeneity between repetitions. If you keep « Repetition » column, with empty cells for repetitions 2 and/or 3, application will try in vain to carry out tests of homogeneity by failing to interpret empty cells. It may be possible to replace empty cells by 0, nevertheless tests of homogeneity will be anyway performed even if it doesn't make sense. The best solution is therefore to remove « Repetition » column, as well as rows with empty cells.

Merging of classes

Merging of classes

No merging
 Merging of classes

Apply

Merging of classes

No merging
 Merging of classes

Class 0 - Class 1

No merging

Class 0 - Class 1

Class 1 - Class 2

Class 2 - Class 3

Class 0 - Class 1 - Class 2

Class 1 - Class 2 - Class 3

Class 0 - Class 1 ; Class 2 - Class 3

It's also possible for comparison of distribution between repetitions to arrange data by merging together several classes as one. This may be very relevant for instance when two classes of symptom are close.

Merging of classes allow to perform one or two gathering maximum. Every consecutive classes can be merged, insofar as rearranged biotest must have maximum two classes of symptom.

To perform a merging of classes, all you must do is to tick « Merging of classes », to select in the drop-down list desired merging and to press on button « Apply ».

This merging of classes will be effective for the whole of sub-tabs of « Comparison of distributions between repetitions ».

According to the choice of pathosystem in the drop-down list of the first tab, some associations are performed by default. For instance, if you have selected in the drop-down list a test Pea – *Ascochyta pisi*, then classes 0 and 1 will be merged together as one.

Stacked histogram

The first tab relates to the stacked histogram. This is a graphical representation method of the distribution of symptom classes per repetitions. Therefore, this method allows to visually compare if there is a difference between two repetitions in distribution of symptom classes.

Merging of symptom classes mentioned above is effective in the sub-tab « Stacked histogram ». Thereby, it is possible to directly view effect of the merging on the new distribution of symptoms for each repetition.

It's also possible to manually change graphic settings of histogram for a better display, to wit:

- Size labels: This is the name of each variety. Some varieties have name of important size, in this way it is possible to reduce size of labels, or in the other hand to increase it when labels are too small to read variety name. By default, label size is 10, nevertheless this text size can be modified either by typing directly a new value, or by using upwards and downwards arrows. Labels size are depicted in blue in the figure below.
- Text size of axis: It's also possible to modify text size of ordinate and abscissa, as well as their legend. By default, size of axis text is 12, nevertheless this text size can be modified either by typing directly a new value, or by using upwards and downwards arrows. Text size of axis is depicted in yellow in the figure below.
- Legend size: Text size of legend (framed in purple in the figure below) can also be changed. By default, this one is of 12, but it can be modified either by typing a new value, or by using upwards and downwards arrows.



Homogeneity regardless of the variety

First, we test if there is a difference between distribution of repetitions regardless of variety (in other words all varieties pooled together). That is why, a χ^2 test is performed to test if there is a difference between the distribution of symptom for each repetition. The advantage of this test is that it can be performed regardless of the number of class and regardless of the number of repetitions of biotest.

Hypothesis of Chi-squared test are:

- H_0 : Homogeneity of the distribution of symptom between repetitions (variables susceptibility and repetition are independent).
- H_1 : Heterogeneity of the distribution of symptom between repetitions (variables susceptibility and repetition are not independent).

In the display window, results of chi-squared test are divided in 3 parts:

- Distribution of symptom per repetition: A simple contingency table with in row various repetitions and in column classes of symptom.
- Chi-squared test: Results of the test strictly speaking with statistic of the χ^2 , the degree of freedom df and the p-value related to the test.
- Results of the test: Statistical decision associated to the results of the test. If p-value is higher than 0.05, then we cannot reject null hypothesis and we conclude to a homogeneity of the distribution of symptom between repetitions. Otherwise, null hypothesis is rejected, and we conclude to a heterogeneity of the distribution of symptom between repetitions.

Chi-squared test will be carried out on the biotest with merging of classes if the latter is performed. For instance, here results are observed on a test pea – *Ascochyta pisi* with merging of classes 0 and 1:

• Distribution of symptom per repetition

Repetition	Class 0-1	Class 2	Class 3
1	28	3	19
2	29	4	15

• Chi-squared test

Pearson's Chi-squared test

data: pooled data
X-squared = 0.59042, df = 2, p-value = 0.7444

• Results of the test

Homogeneity of the distribution of symptom between repetitions

Homogeneity for each variety

To compare distribution between repetitions of each variety, a test of homogeneity is performed this time for each variety. A Fisher's exact test is preferred here to chi-squared test for two reasons: the small number of each repetition and the predictable imbalance of the expected numbers.

Hypothesis of Fisher's exact test nevertheless remain the same:

- H_0 : Homogeneity of the distribution of symptom between repetitions (variables susceptibility and repetition are independent).
- H_1 : Heterogeneity of the distribution of symptom between repetitions (variables susceptibility and repetition are not independent).

This time, results are presented as a table with the name of the variety, the p-value of the Fisher's exact test for each variety and the statistical decision associated to the results of the test. In other words, if p-value is lower than 0.05, we reject H_0 and we conclude to a heterogeneity of the distribution of symptom between repetitions. Otherwise, we cannot reject null hypothesis, and we conclude to a homogeneity of the distribution of symptom.

Furthermore, in this table, names of homogeneous varieties are coloured green, while names of heterogeneous varieties are coloured red.

Variety	P-value	Result
V1	0.103	Homogeneity of repetitions
V2	0.303	Homogeneity of repetitions
MDK (Susceptible)	1.000	Homogeneity of repetitions
MADONNA (Resistant)	0.005	Heterogeneity of repetitions
NINA (Resistant)	1.000	Homogeneity of repetitions

Study of the susceptibility / resistance of varieties

The following tab concerns the study of the susceptibility / resistance of varieties. It is consisted itself of 3 sub-tabs:

- Graphical representation
- Dendrogram

- Decisions

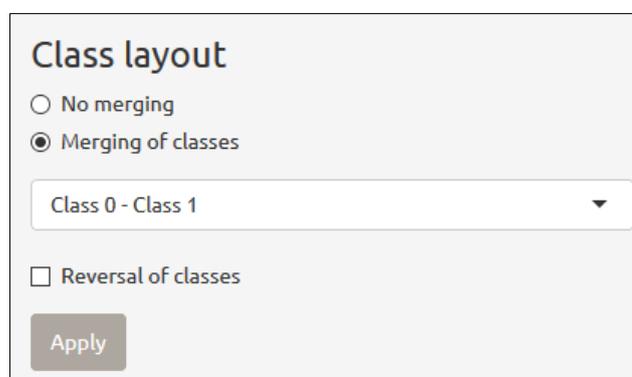
Class layout

As for the last tab, it is also possible to reorganise data by merging together several classes as one. This may be useful, for instance when two symptom classes are close.

The merging of classes allows to carry out one or two gathering maximum. Every consecutive classes can be merged together, insofar as rearranged biotest must have maximum two classes of symptom.

To carry out a merging of classes, it is enough to tick « Merging of classes », to select in the drop-down list desired merging and to press on button « Apply ».

It is also possible to modify position of symptom classes. Indeed, first class of symptom is expected to be the most resistant. If this is not, we can tick option « Reversal of classes » and press on button « Apply » to ensure that the first class of symptom become the most susceptible.



The screenshot shows a settings panel titled "Class layout". It contains two radio buttons: "No merging" (unselected) and "Merging of classes" (selected). Below the radio buttons is a dropdown menu with the text "Class 0 - Class 1" and a downward arrow. Underneath the dropdown is a checkbox labeled "Reversal of classes" which is currently unchecked. At the bottom of the panel is a button labeled "Apply".

According to the choice of the pathosystem in the drop-down list of the first tab, some merging or reversal of classes are undertaken by default. For example, if we have selected in the drop-down list a test Pea – *Ascochyta pisi*, then classes 0 and 1 will be merged together. In the other hand, if we have selected a test Cucumber – *Corynespora cassiicola* application will carry out by default, a reversal of classes.

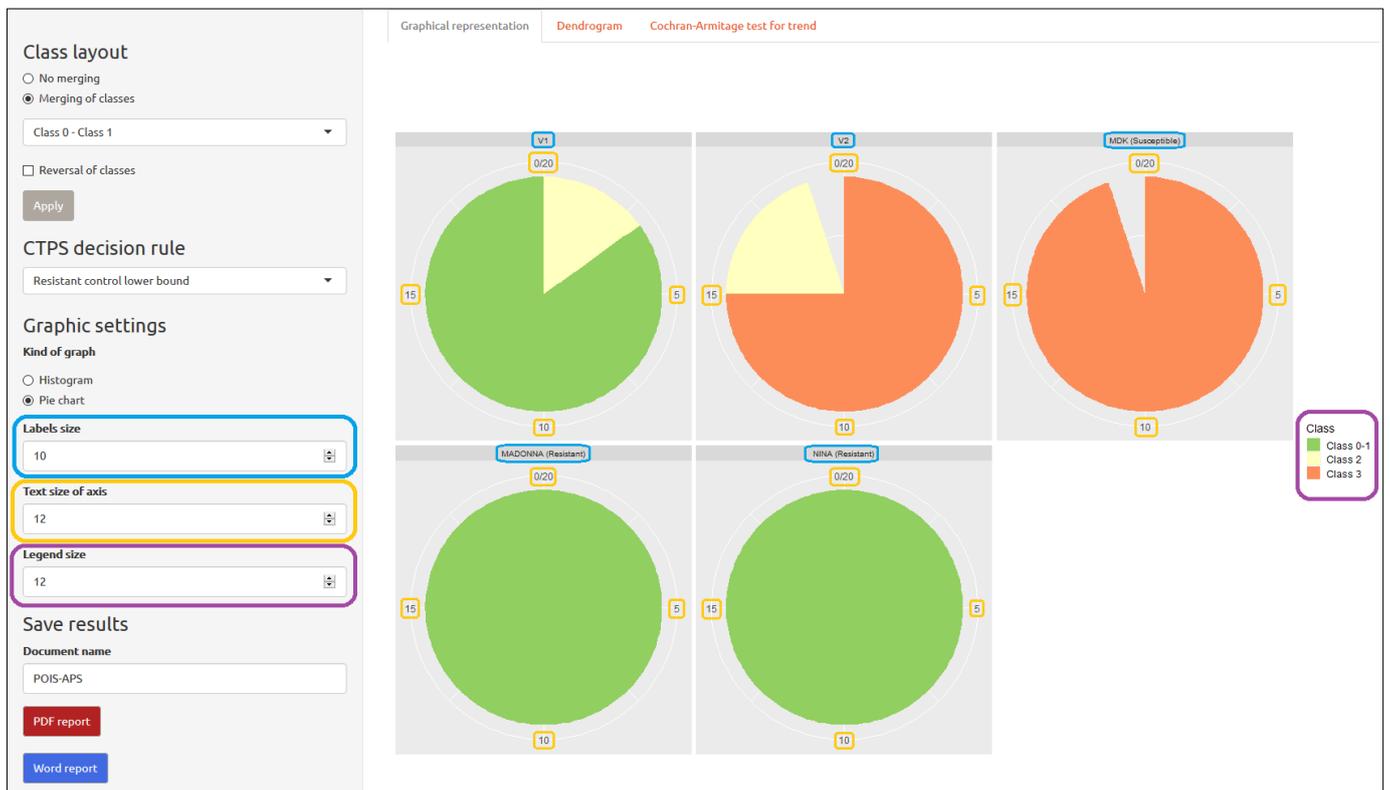
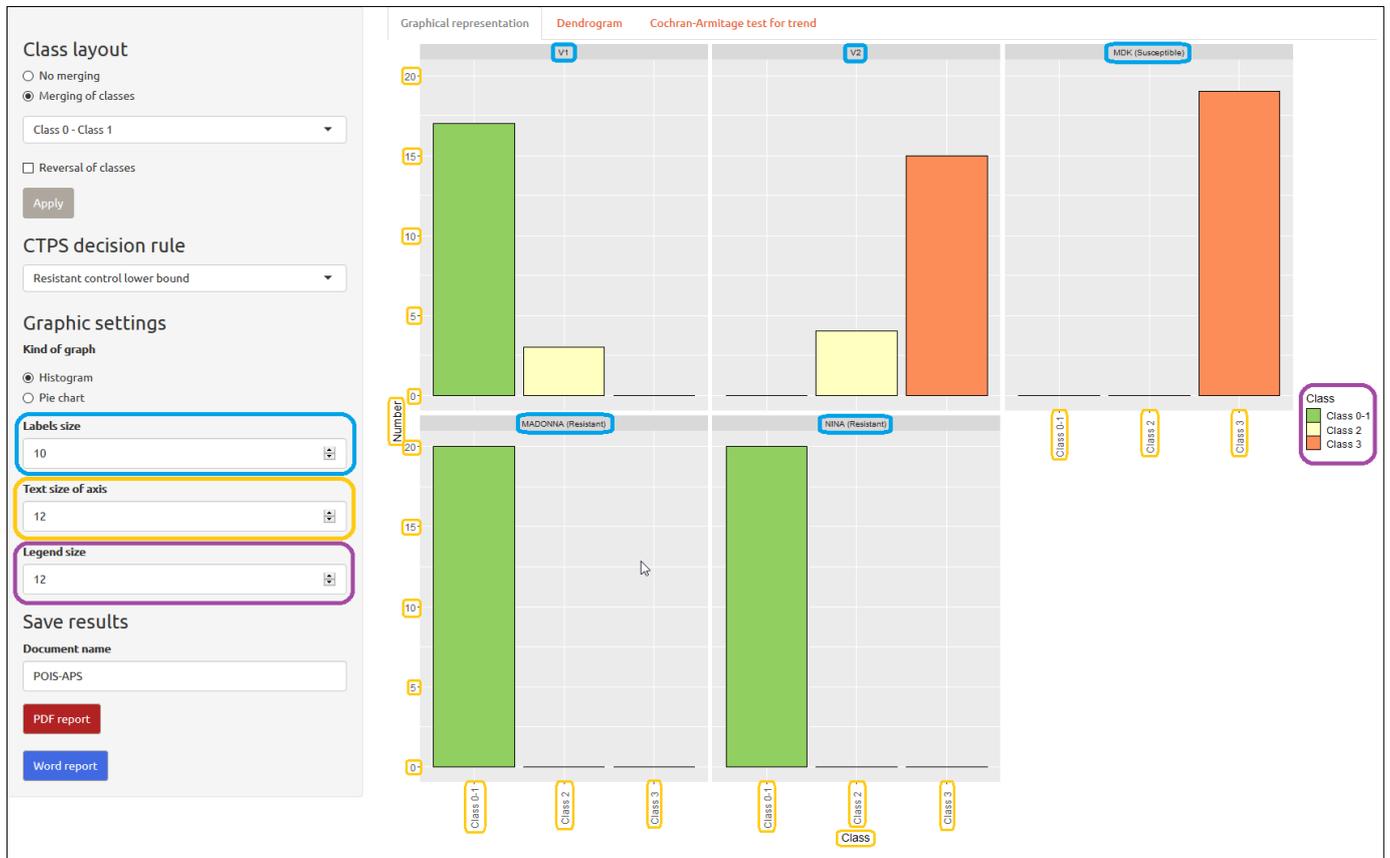
Graphical representation

The first sub-tab entitled « Graphical representation » allows as the name suggests, to view distribution of symptom classes. Within the option « Graphic settings », it is firstly possible to choose the kind of graph. In this way, to display data, we have a choice between a histogram or a pie chart. Both kinds of graph are presented below.

As for the stacked histogram from the last tab, we could manually modify the various graphic settings:

- Labels size: This is the name of each variety. By default, label size is 10, nevertheless this text size can be modified either by typing directly a new value, or by using upwards and downwards arrows. Labels size are depicted in blue in figures below.
- Text size of axis: It's also possible to modify text size of ordinate and abscissa, as well as their legend. By default, size of axis text is 12, nevertheless this text size can be modified either by typing directly a new value, or by using upwards and downwards arrows. Text size of axis is depicted in yellow in the figure below.

- **Legend size:** Text size of legend (framed in purple in the figure below) can also be modified. By default, this one is of 12, but it can be changed either by typing a new value, or by using upwards and downwards arrows.



Dendrogram

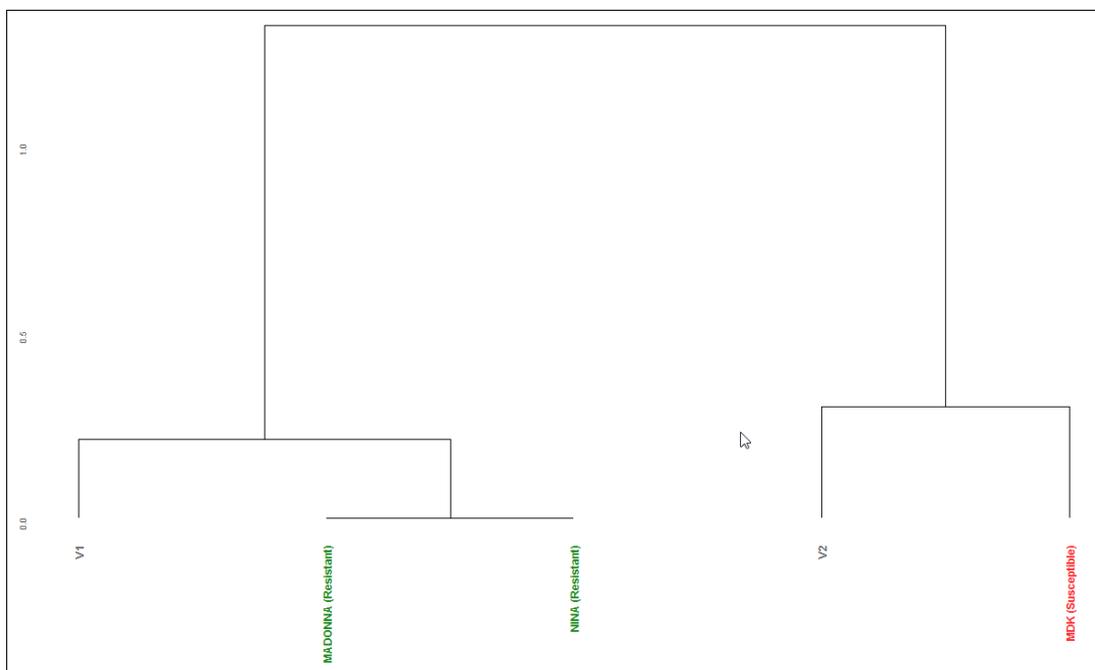
In the following sub-tab, a Hierarchical Cluster Analysis (HCA) is performed. HCA is an iterative clustering method based on a simple principle: we want individuals gathered in a same class (intra-class homogeneity) to be the most similar, while we want classes to be the most dissimilar (inter-class heterogeneity).

Thus, the concept is to gather individuals according a similarity criterion. We will use here as similarity criterion, distribution of symptom classes of each variety. That's why in a first time, distance matrix based on Euclidian distance has been built from distribution of symptom class for each variety. Furthermore, this distance matrix is based on the relative frequency of affiliation to each class rather than real numbers. In this way, two varieties with similar distribution pattern but different numbers are alike. This distance matrix is a square matrix including all distances two by two between each variety. This distance is a measurement from similarity / dissimilarity between two varieties.

From this matrix we carry out a hierarchical cluster analysis. The concept is to create a dendrogram (classification tree) where all varieties with similar pattern are gathered. 5 agglomerative methods are tested for implementation of the dendrogram: *single-linkage*, *complete-linkage*, *ward.D2*, *UPGMA* and *UPGMC*. Among these 5 trees, we want to select the best of all. That's why the cophenetic correlation coefficient is computed. This coefficient is an indicator of the quality of a hierarchical clustering. Thus, application will select tree with the higher coefficient.

Furthermore, for a best visibility, especially when we observe many varieties, controls variety names have been coloured: red for susceptible controls, light green for intermediate controls and low intermediate controls, and dark green for resistant controls and high intermediate controls.

It's also possible to increase or reduce size of the variety names. For this, you simply need, as for previous tab, to modify text size of the axis in the graphic setting.



⚠ The hierarchical cluster analysis doesn't consider position of symptom class but only the distribution of symptom. Thus, for example according to dendrogram a variety with all plants in class

0 will be found closer to a variety with only one plant in class 0 and all other in class 2, rather than a variety with all plants in class 1.

Decisions

Susceptibility / resistance level of various varieties is always decided on the comparison with control varieties. That's why a model has been built considering this.

The purpose of this model is to compare each variety with control varieties. For each variety, a Cochran-Armitage test for trend has been performed to study if its susceptibility level is lower or higher than every control variety. This test can be considered as an enhancement of the χ^2 test considering the general tendency between two variables. Thus, the model carries out for each variety as many Cochran-Armitage tests for trend as there are controls. However, according to the kind of control, hypotheses of the test are different:

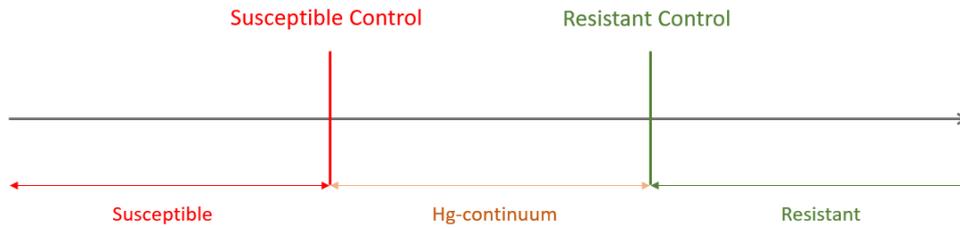
- Susceptible control:
 - H_0 : the susceptibility level of the variety is higher or equal to the control
 - H_1 : the susceptibility level of the variety is lower than the control
- Intermediate control:
 - H_0 : the resistance level of the variety is higher or equal to the control
 - H_1 : the resistance level of the variety is lower than the control
- High intermediate control:
 - H_0 : the resistance level of the variety is higher or equal to the control
 - H_1 : the resistance level of the variety is lower than the control
- Resistant control:
 - H_0 : the resistance level of the variety is higher or equal to the control
 - H_1 : the resistance level of the variety is lower than the control

As a reminder, when a p-value is lower than significance level α , we can reject null hypothesis H_0 . Value of the significance level α , for Cochran-Armitage test for trend is set arbitrarily to 0.05. The value of this significance level is increased to 0.1 for resistance test with two symptom classes (usually presence / absence of symptoms) to agree with rules of tolerated off-types.

⚠ Susceptibility level of control varieties must not be illogical (a resistant control variety must be more resistant than a susceptible control variety). Otherwise, results of the test will be logically aberrant. It's to the user to be vigilant as for susceptibility level of his control variety.

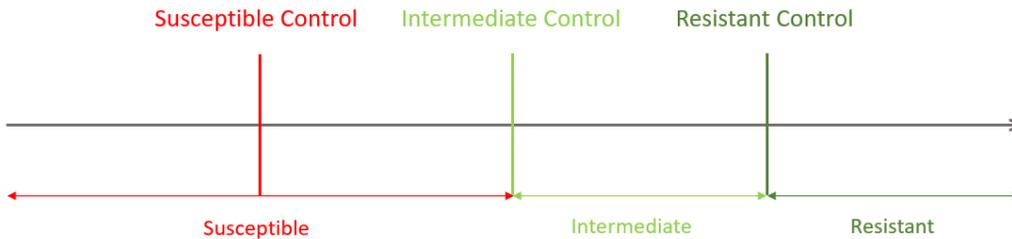
Moreover, there are 5 CTPS decision rules which allow various interpretation of the variety. Each decision rule has its own assessment method of the susceptibility / resistance:

- No control bound :
 - If the variety has a susceptibility level higher or equal to susceptible control, then variety will be susceptible.
 - If the variety has a susceptibility level lower or equal to resistant control, then variety will be resistant.
 - If the variety has a susceptibility level between two controls, then it will be heterogeneous, or it will present a fluctuation of expression of the genotype according to a continuum. The decision will be on expert judgement on the base of the comparison of variety with controls.



• IR control lower bound :

- If the variety has a susceptibility level higher or equal to susceptible control, then variety will be susceptible.
- If the variety has a susceptibility level higher than intermediate resistant control, then variety will be susceptible.
- If the variety has a susceptibility level lower or equal to intermediate resistant control, then variety will be intermediate resistant.
- If the variety has a susceptibility level lower or equal to resistant control, then variety will be resistant.



• Resistant control lower bound :

- If the variety has a susceptibility level higher or equal to susceptible control, then variety will be susceptible.
- If the variety has a susceptibility level higher than resistant control, then the variety will be susceptible.
- If the variety has a susceptibility level lower or equal to resistant control, then the variety will be resistant.



• Susceptible control higher bound :

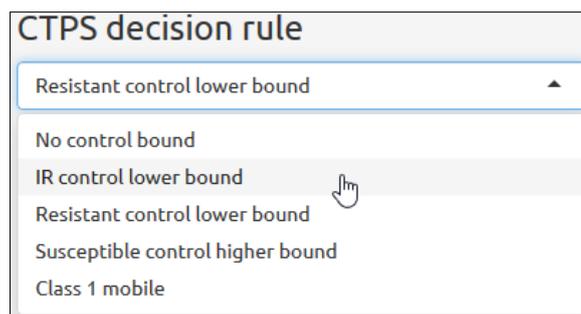
- If the variety has a susceptibility level higher or equal to susceptible control, then variety will be susceptible.
- If the variety has a susceptibility level lower than susceptible control, then the variety will be resistant.
- If the variety has a susceptibility level lower or equal to resistant control, then the variety will be resistant.



- Class 1 mobile :
 - If the variety has a susceptibility level higher or equal to susceptible control, then variety will be susceptible.
 - If the variety has a susceptibility level higher than resistant control, then the variety will be susceptible.
 - If the variety has a susceptibility level lower or equal to resistant control, then the variety will be resistant.
 - Class 1 is mobile : If plants are shared on classes 0 and 1, plants of class 1 would be moved in class 0 (considered as resistant) ; If plants are shared on classes 1 and 2, plants of class 1 would be moved in class 2 (considered as susceptible).

When the test uses two controls of same level, it will always be the less strict control which will be used as bound. In this way, when biotest will present two susceptible controls, it will be the less susceptible of both which will be used as reference. When resistance test will consist of two intermediate resistant controls, it will be the less intermediate resistant of both which will be used as bound. And finally, when we have two resistant controls or two high intermediate resistant controls, it will be the less resistant of both which will be used as reference bound.

We can choose a decision rule among these 5 rules in the section « CTPS decision rule » to the left of the screen:



According to the pathosystem selected in the drop-down list of the tab « Selection and parametrization of the biotest », the right CTPS decision rule will be adopted. Nevertheless, it's always possible to change it afterwards. Here is for each pathosystem, the decision rule associated by default :

Pathosystem	CTPS decision rule
Bean - <i>Colletotrichum lindemuthianum</i> race 6	Resistant control lower bound
Bean - <i>Colletotrichum lindemuthianum</i> race Kappa	Resistant control lower bound
Bean - <i>Pseudomonas savastanoi</i> pv <i>phaseolicola</i> race 6	IR control lower bound
Cabbage - <i>Fusarium oxysporum</i> f. sp. <i>conglutinans</i>	IR control lower bound
Cucumber - <i>Cladosporium cucumerinum</i>	No control bound
Cucumber - CMV	IR control lower bound
Cucumber - <i>Corynespora cassiicola</i>	No control bound
Cucumber – CVYV	IR control lower bound
Cucumber - <i>Podosphaera xanthii</i>	IR control lower bound
Lettuce - <i>Bremia lactucae</i>	No control bound
Lettuce - <i>Fusarium oxysporum</i> f. sp. <i>lactucae</i>	IR control lower bound
Melon - <i>Fusarium oxysporum</i> f. sp. <i>melonis</i> race 0	No control bound
Melon - <i>Fusarium oxysporum</i> f. sp. <i>melonis</i> race 1.2	IR control lower bound
Melon - <i>Fusarium oxysporum</i> f. sp. <i>melonis</i> race 1	No control bound
Melon - <i>Fusarium oxysporum</i> f. sp. <i>melonis</i> race 2	No control bound

Melon - MWMV	Resistant control lower bound
Melon - <i>Podosphaera xanthii</i> race 1, 2, 3, 5, 3-5	IR control lower bound
Melon - ZYMV	Resistant control lower bound
Pea - <i>Ascochyta pisi</i>	Resistant control lower bound
Pea – BYMV	No control bound
Pea – <i>Erysiphe pisi</i>	No control bound
Pea - <i>Fusarium oxysporum</i> f. sp <i>pisii</i> race 1	Resistant control lower bound
Pepper - <i>Meloidogyne incognita</i>	Susceptible control higher bound
Tomato - <i>Fusarium oxysporum</i> f. sp. <i>lycopersici</i> race 0	Resistant control lower bound
Tomato - <i>Fusarium oxysporum</i> f. sp. <i>lycopersici</i> race 1	Resistant control lower bound
Tomato - <i>Fusarium oxysporum</i> f. sp. <i>lycopersici</i> race 2	Resistant control lower bound
Tomato - <i>Fusarium oxysporum</i> f. sp. <i>radicis lycopersici</i>	No control bound
Tomato - <i>Meloidogyne incognita</i>	IR control lower bound
Tomato - <i>Pseudomonas syringae</i> pv. <i>tomato</i>	No control bound
Tomato - <i>Pyrenochaeta lycopersici</i>	IR control lower bound
Tomato - <i>Stemphylium</i> spp.	No control bound
Tomato - <i>Verticillium dahliae</i>	Class 1 mobile
Zucchini – CMV	IR control lower bound
Zucchini – WMV	IR control lower bound
Zucchini – ZYMV	IR control lower bound

These decision rules are all defined according to CTPS protocol. They all depend on control varieties used. If the resistance test differs from protocol, it's to the user to adapt the right decision rule in accordance with his own control varieties.

Besides the comparison of varieties with controls, a test for heterogeneity is also realised. Are heterogeneous all varieties which meet both conditions:

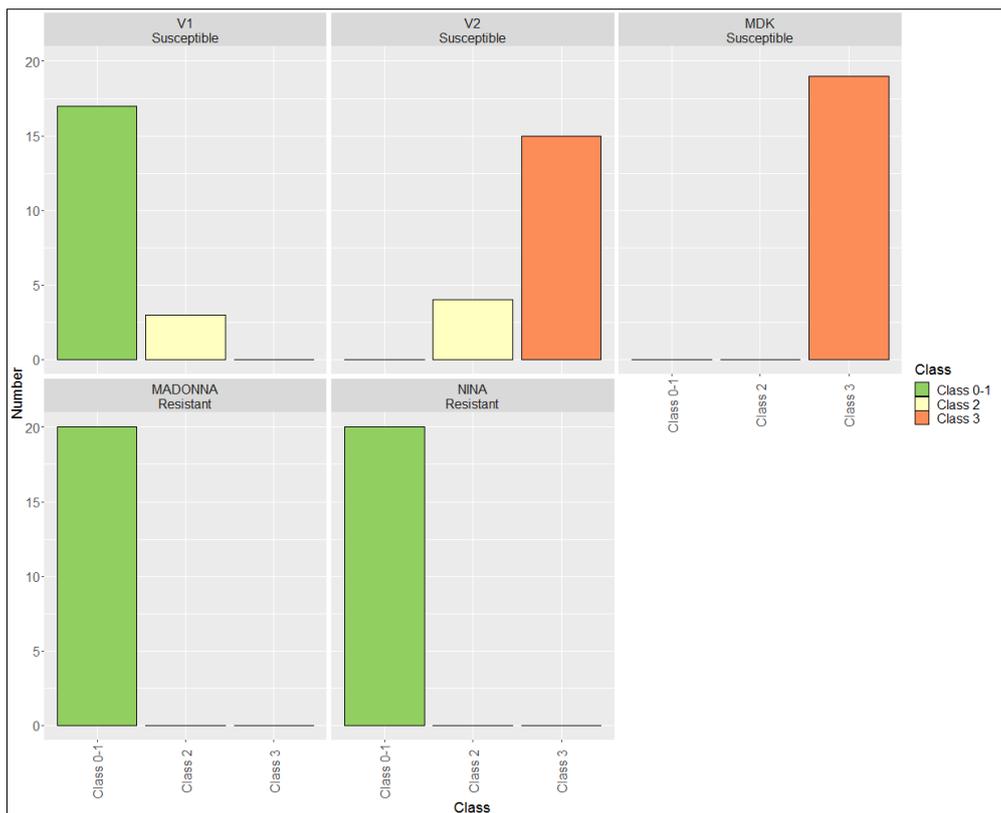
- A number in both most extreme classes (the most susceptible and the most resistance classes), within the limits of a tolerated number of off-types based on number of control varieties in the same classes.
- No plant in at least one of the intermediate classes.

The boundary between a heterogeneous and an intermediate variety is based on expert judgement with the help of statistical test. Indeed, the viewing of symptoms , as well as test results don't always allow to identify if a variety is constituted of a mix of susceptible and resistant plants or if variety is intermediate. Furthermore, number of plant for each varieties and controls need to be similar in order that test be viable.

This last tab « Cochran-Armitage test for trend » is a copy of the tab « Graphical representation » except that results of the model as for susceptibility / resistance of each variety are displayed below their names. Level of control varieties is also indicated below their names as other varieties.

Just like graphical representation, it's possible to choose in the left section between to display results in the form of a histogram or a pie chart. Font size of labels, of axis text, and of the legend can also be modified to adapt to the window.

For this example, varieties V1 and V2 are both proposed susceptible because they are of a susceptibility level lower than resistant control MADONNA and NINA (Resistant control lower bound).



Save results

Once we get results of application, the following step is to save and export our results. Two kinds of output are conceivable : export results in PDF format or in Word format (docx).

For this, we firstly need to choose a document name. By default, application reuse the Excel file name including results of resistance test loaded at the first tab. Once document name modified, the following step is to press on the red button « PDF Report » or on the blue button « Word Report » to download output file at the desired format.

Save results

Document name

PDF report

Word report

PDF or Word output files will show the same results. Nevertheless, PDF output will have more graphical possibilities than a Word output. Indeed, options of document page layout are more interesting for a PDF output. Moreover, application take more time to build document in Word format than in PDF format because Word require an image processing with higher quality to get a correct rendering.

Despite its drawbacks Word output presents the benefits to be able to reuse results afterwards (especially images) and to be able to modify document page layout once it has been downloaded.

Report in PDF format

PDF output document is composed as follows. Document will bear the title selected formerly (here for instance POIS-APS). Under this title, we will observe today's date in *dd month yyyy* format. Directly below this date, we encounter our table of resistance test with our control varieties recently appointed. As inside application, according to option « Minimum threshold number » applied, varieties with a number lower than threshold could be ignored (« Don't apply any threshold »), could be indicated (« Inform about varieties for which a number < threshold ») [names red-coloured and arrival of column « Number ≥ threshold »], or could be clearly removed of resistance test (« Pull out varieties for which number < threshold »). In our example below, we have used option « Inform about varieties for which a number < threshold »:

POIS-APS						
<i>16 December 2019</i>						
Variety	Number ≥ threshold	Repetition	Class 0	Class 1	Class 2	Class 3
V1	✓	1 2	3 8	5 1	2 1	0 0
V2	✗	1 2	0 0	0 0	1 3	9 6
MDK (Susceptible)	✗	1 2	0 0	0 0	0 0	10 9
MADONNA (Resistant)	✓	1 2	8 1	2 9	0 0	0 0
NINA (Resistant)	✓	1 2	10 9	0 1	0 0	0 0

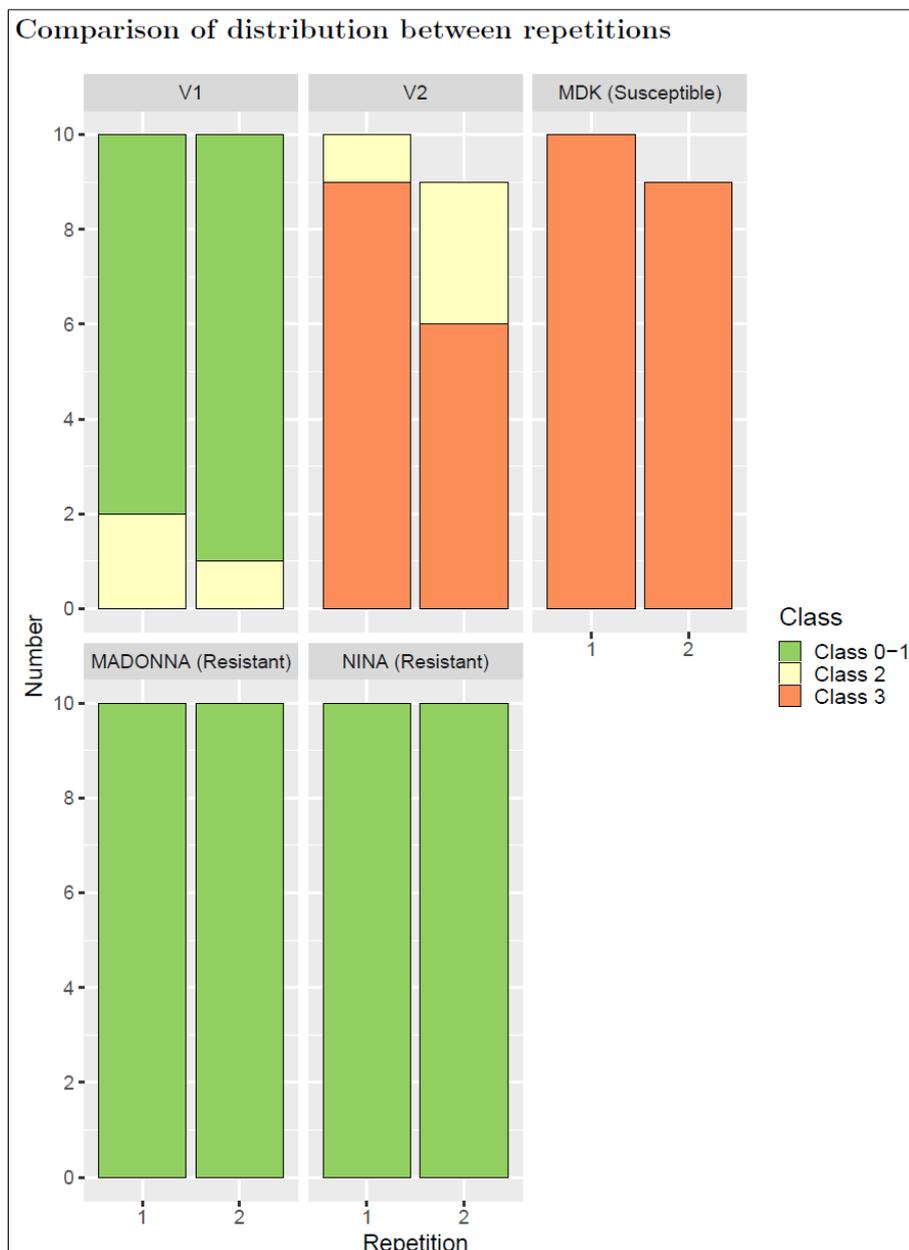
Directly below the table of the resistance test, we have another summarising table of the features of the resistance test. This table is informative as for pathosystem selected in the drop-down list, option of threshold chosen and value of the latter, various controls selected and CTPS decision rule for which we opted:

Resistance test features	
Parameter	Value
Species	Pea
Pathogen	Ascochyta pisi
Use of threshold	Inform about varieties for which number < threshold
Number	20
Susceptible control	MDK
Intermediate control	
High intermediate control	
Resistant control	MADONNA, NINA
CTPS decision rule	Resistant control lower bound

In the following part « Comparison of distribution between repetitions », a first page is only devoted to stacked histogram.

⚠ To keep a similar proportionality relationship between font sizes (labels, axis and legend) of application and PDF document, font size assigned within application is revised downwards when we compile the document:

- Labels size : it is reduced of 6 in the PDF document. In other words, if we assign a size of label lower than 7 within application, this one will be null or negative in the PDF document and labels won't be visible.
- Text size of axis : it is reduced of 6 in PDF document. In other words, if we assign a size of axis lower than 7 within application, this one will be null or negative in PDF document and axis and their title won't be visible.
- Legend size : it is reduced of 5 in PDF document. In other words, if we assign a size of legend lower than 6 within application, this one will be null or negative in PDF document and legend won't be visible.



The next page is at the same time devoted to the test of homogeneity regardless of the variety and to the test of homogeneity for each variety. Results are presented according to the same architecture that for corresponding tabs of the application. Names of the variety which are homogeneous are also coloured in green, just like the name of the variety heterogeneous are coloured in red.

Homogeneity regardless of the variety

- Distribution of symptom per repetition:

Repetition	Class 0	Class 1	Class 2	Class 3
1	21	7	3	19
2	18	11	4	15

- Chi-squared test:

Table 3: Pearson's Chi-squared test: pooled data

Test statistic	df	P value
1.693	3	0.6385

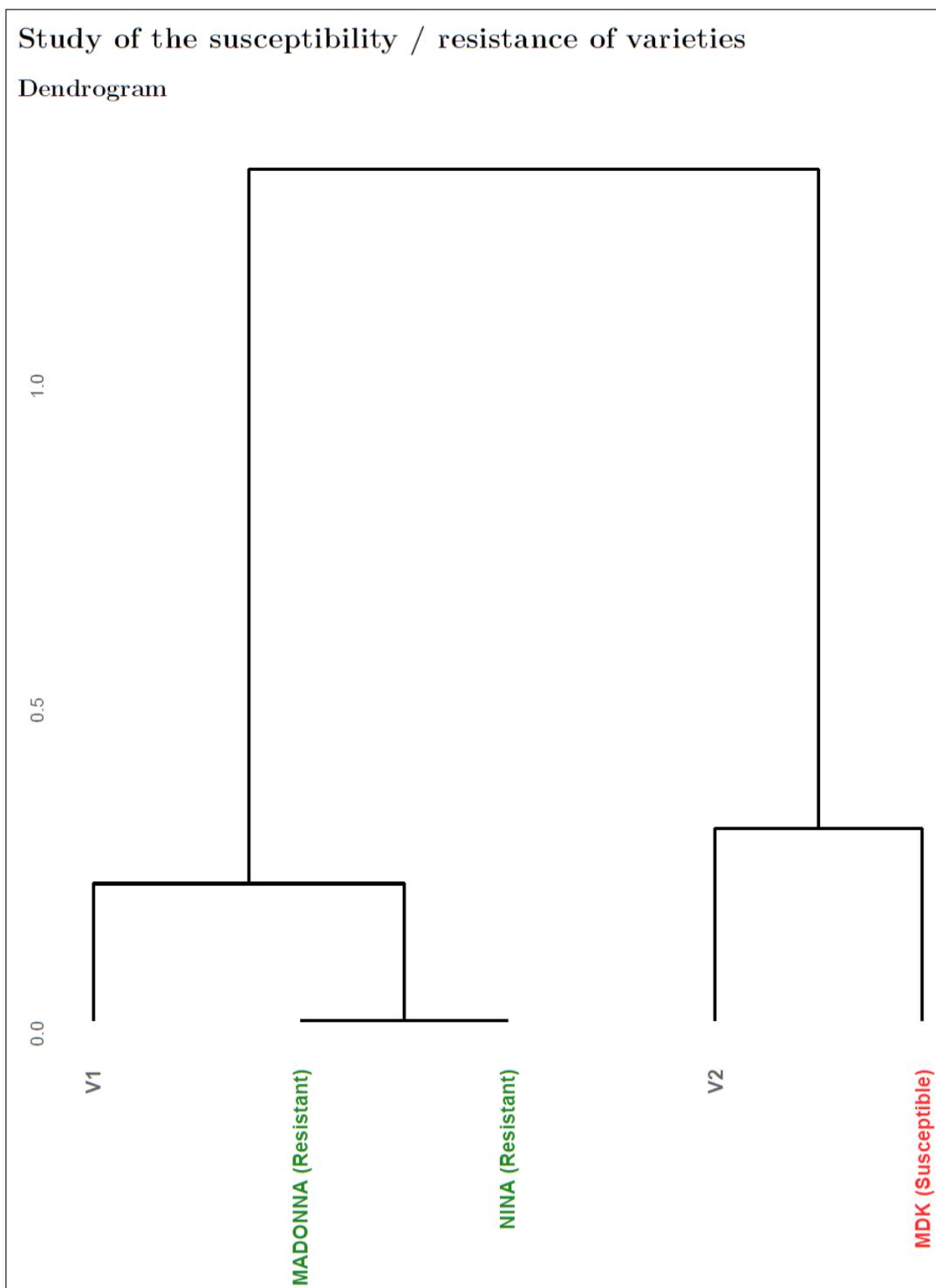
- Résultats of the test: Homogeneity of the distribution of symptom between repetitions

Homogeneity for each variety

Variety	P-value	Result
V1	1.03e-01	Homogeneity of repetitions
V2	3.03e-01	Homogeneity of repetitions
MDK (Susceptible)	1.00e+00	Homogeneity of repetitions
MADONNA (Resistant)	5.48e-03	Heterogeneity of repetitions
NINA (Resistant)	1.00e+00	Homogeneity of repetitions

The following page of the document is devoted to the dendrogram of the tab « Study of the susceptibility / resistance of varieties ».

△ Just like stacked histogram of the previous page, size of the variety names on dendrogram have been reduced to keep a proportionality relationship between application and PDF report. Font size for variety names was reduces of 6 compared to application. In other words, if we have a text size of axis lower than 7 within application, that one will be null or negative inside the PDF document and the name of various varieties of the dendrogram will be not visible.

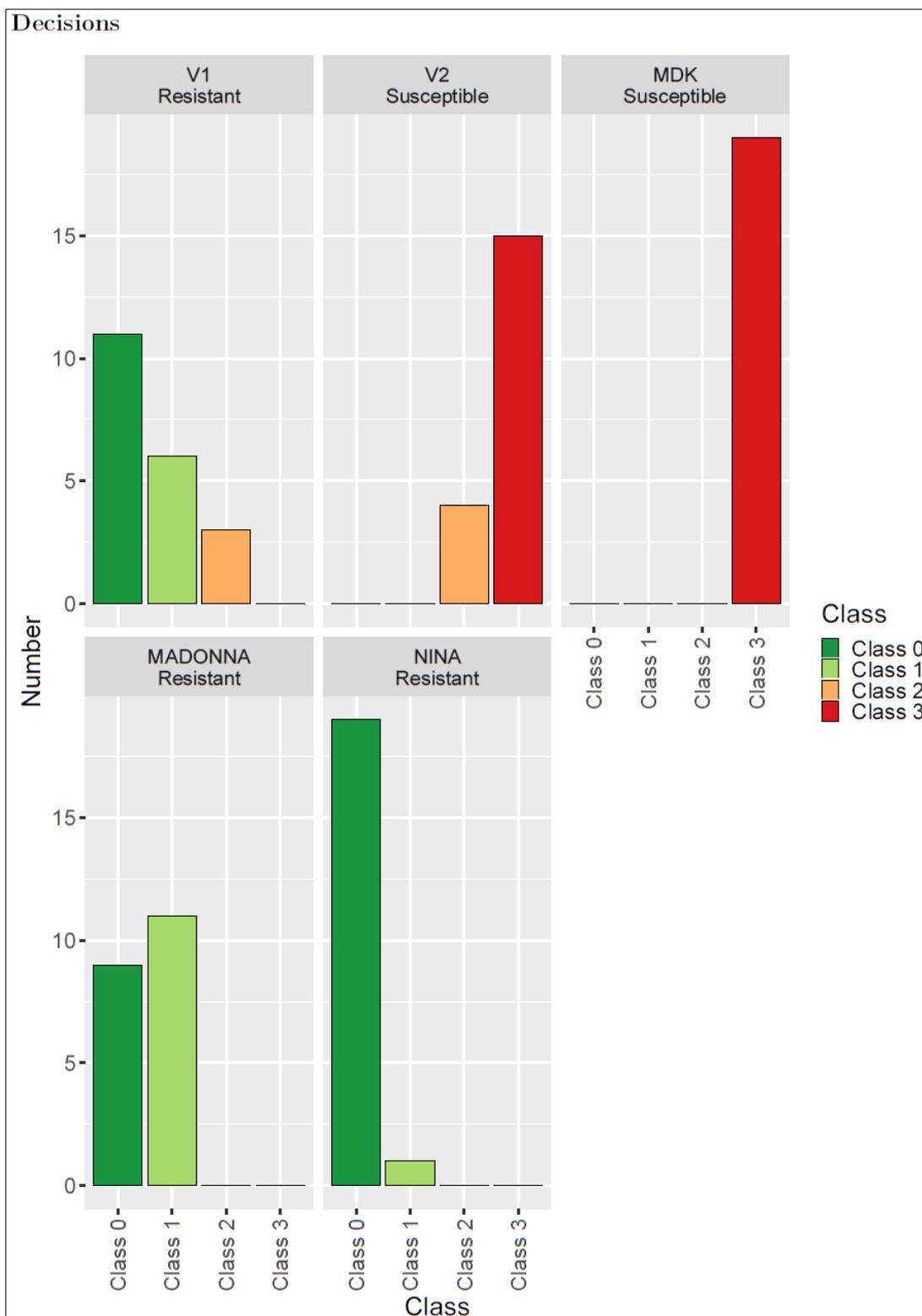


The last part of the PDF report is devoted to the decision. The first page is a copy of the output of the tab of the same name in application.

⚠ As for the previous graphic output, according to keep a similar proportionality relationship between the font sizes (labels, axis and legend) of application and PDF document, font size assigned within application is revised downwards when we compile the document :

- Labels size : it is reduced of 6 in the PDF document. In other words, if we assign a size of label lower than 7 within application, this one will be null or negative in the PDF document and labels won't be visible.

- Text size of axis : it is reduced of 5 if we want to display results in the form of histogram and of 8 if a pie chart is preferred. In other words, if we assign a size of axis lower than 6 for histogram and 9 for pie chart, within application, this one will be null or negative in PDF document and axis and their title won't be visible.
- Legend size : it is reduced of 5 in PDF document. In other words, if we assign a size of legend lower than 6 within application, this one will be null or negative in PDF document and legend won't be visible.



Still in the same part related to the Cochran-Armitage test for trend, straight after presentation of results, we find a table detailing the whole of results. This table summarize the whole of the various Cochran-Armitage test for trend realised for each variety.

Résultats détaillés				
Variété	P-value (TS)	P-value (TR1)	P-value (TR2)	Proposition
V1	8.77e-10	3.59e-02	3.59e-02	Sensible
V2	1.72e-02	1.33e-09	1.33e-09	Sensible

According to the kind of control variety and the CTPS decision rule, results and interpretation of the susceptibility / resistance of the variety are interpreted in different ways:

- Susceptible control (S): if p-value is higher than 0.05, we consider that the variety is not at a resistance level higher than control. If CTPS decision rule is « No control bound » or « Susceptible control higher bound » and that p-value in comparison to susceptible control(s) is higher than 0.05, then variety will be declared susceptible. If CTPS rules is « Susceptible control higher bound » and that p-value is lower than 0.05, then variety will be declared resistant.
- Intermediate resistant control (IR): if p-value is higher than 0.05, we consider that the variety is not at a susceptibility level higher than control. We only have intermediate resistant control for CTPS decision rule « IR control lower bound ». Here, if p-value is higher than 0.05 for intermediate resistant control, variety will be declared intermediate resistant (unless if variety is at the level of resistant control), otherwise variety will be found susceptible.
- Resistant control (R): if p-value is higher than 0.05, we consider that variety is not at a susceptibility level higher than control. No matter CTPS decision rule, if p-value is higher than 0.05, variety will be always declared resistant. For CTPS decision rule « No control bound », if p-value is lower than 0.05, then variety will be declared either susceptible, heterogeneous or heterogeneous-continuum according to comparison to susceptible control. For CTPS decision rule « IR control lower bound », if p-value is lower than 0.05, variety will be found either susceptible or intermediate according to comparison to resistant intermediate control(s). For CTPS decision rules « Resistant control lower bound » and « Class 1 mobile », as soon as p-value is lower than 0.05, variety will be always found susceptible.
- High intermediate resistant control (HIR): We only found high intermediate resistant control in the case of test with CTPS decision rule « IR control lower bound ». For this control, if p-value is higher than 0.05, we consider that the variety will be high intermediate resistant. Otherwise, variety will be susceptible or intermediate resistant according to the resistance level of intermediate resistant control.

⚠ A variety can also be found heterogeneous according the heterogeneity rule. In other words, a variety with plants in both extreme classes, within the limits of the tolerated number of off-types, and with no plant in at least one of the intermediate classes will be found heterogeneous independently of the p-value in comparison with control varieties.

Report in Word format

As document in PDF format, Word report will present title followed by the today's date in *dd month yyyy* format and table of the resistance test with controls selected in application. According to option « Minimum threshold number » applied, varieties with a number lower than threshold could be ignored (« Don't apply any threshold »), could be indicated (« Inform about varieties for which a number < threshold ») [Arrival of column « Number ≥ threshold »], or could be clearly removed of resistance test (« Pull out varieties for which number < threshold »). Moreover, page layout and justification of table should be done manually, because we can't use these options when document is compiled in Word. In our example below, we have used option « Inform about varieties for which a number < threshold »:

POIS-APS						
16 December 2019						
Variety	Number ≥ threshold	Repetition	Class 0	Class 1	Class 2	Class 3
V1	☒	1	3	5	2	0
V1	☒	2	8	1	1	0
V2	X	1	0	0	1	9
V2	X	2	0	0	3	6
MDK (Susceptible)	X	1	0	0	0	10
MDK (Susceptible)	X	2	0	0	0	9
MADONNA (Resistant)	☒	1	8	2	0	0
MADONNA (Resistant)	☒	2	1	9	0	0
NINA (Resistant)	☒	1	10	0	0	0
NINA (Resistant)	☒	2	9	1	0	0

Directly below the table of the resistance test, we have a summarising table of the features of the resistance test. This table is informative as for pathosystem selected in the drop-down list, option of threshold chosen and value of the latter, various controls selected and CTPS decision rule for which we opted. Page layout and justification of the table should be done manually, because these options are not available when document is compiled in Word output.

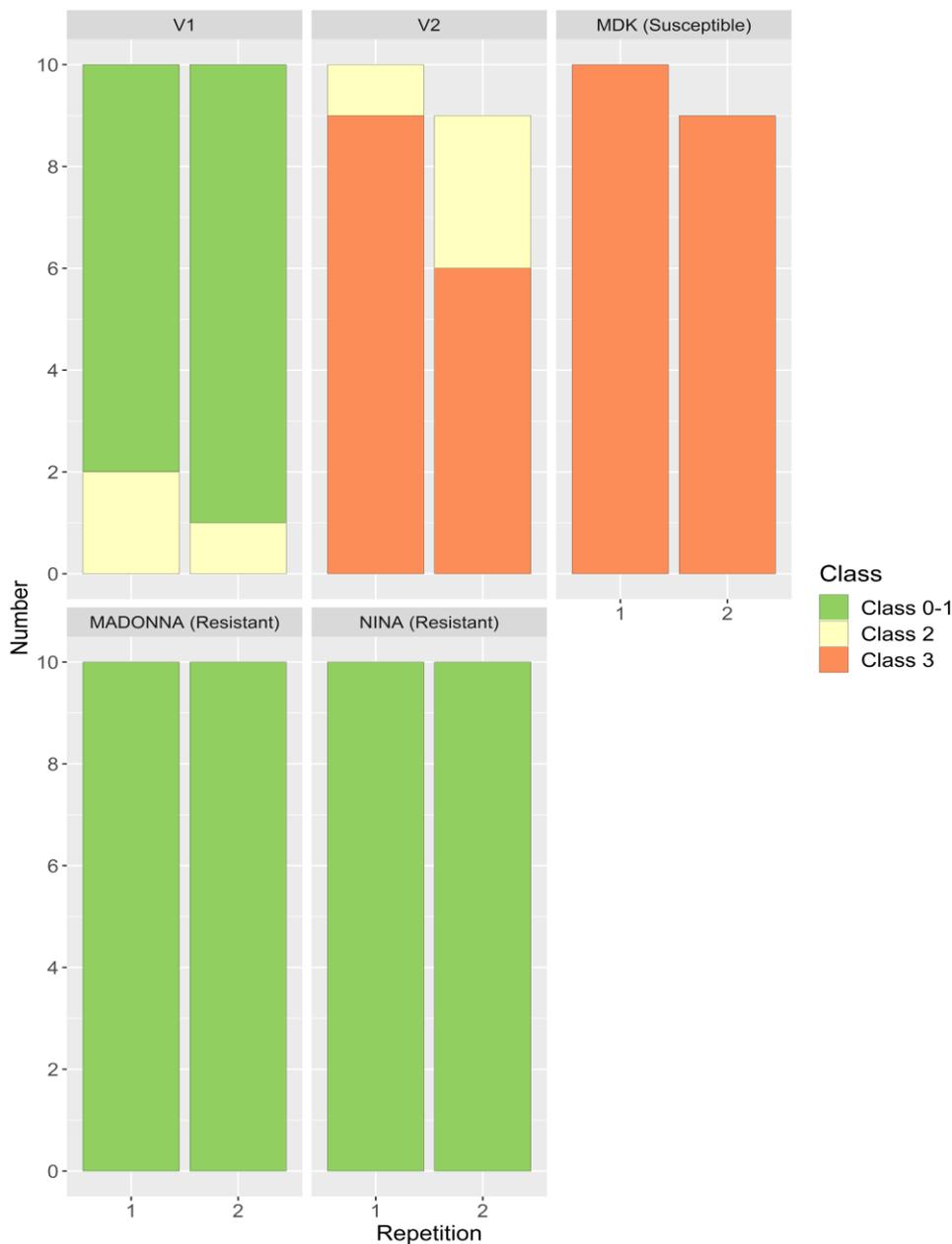
Resistance test features

Parameter	Value
Species	Pea
Pathogen	Ascochyta pisi
Use of threshold	Inform about varieties for which number < threshold
Number	20
Susceptible control	MDK
Intermediate control	
High intermediate control	
Resistant control	MADONNA, NINA
CTPS decision rule	Resistant control lower bound

⚠ To keep a similar proportionality relationship between font sizes (labels, axis and legend) of application and Word document, font size assigned within application is revised downwards when we compile the document:

- Labels size : it is reduced of 5 in the Word document. In other words, if we assign a size of label lower than 6 within application, this one will be null or negative in Word document and labels won't be visible.
- Text size of axis : it is reduced of 4 in Word document. In other words, if we assign a size of axis lower than 5 within application, this one will be null or negative in Word document and axis and their title won't be visible.
- Legend size : it is reduced of 3 in Word document. In other words, if we assign a size of legend lower than 4 within application, this one will be null or negative in Word document and legend won't be visible.

Comparison of distribution between repetitions



The next page is at the same time devoted to the test of homogeneity regardless of the variety and to the test of homogeneity for each variety. Results are presented according to the same architecture that for corresponding tabs of the application. Nevertheless, unlike PDF report, names of homogeneous varieties are not coloured in green, and names of heterogeneous varieties are not coloured in red purely and simply because it is not possible to add colour font when we compile our document in Word output. Moreover, tables are not centred in Word output. However, we always can modify page layout once document has been compiled.

Homogeneity regardless of the variety

- **Distribution of symptom per repetition:**

Repetition	Class 0	Class 1	Class 2	Class 3
1	21	7	3	19
2	18	11	4	15

- **Chi-squared test:**

Pearson's Chi-squared test: pooled data

Test statistic	df	P value
1.693	3	0.6385

- **Results of the test:** Homogeneity of the distribution of symptom between repetitions

Homogeneity for each variety

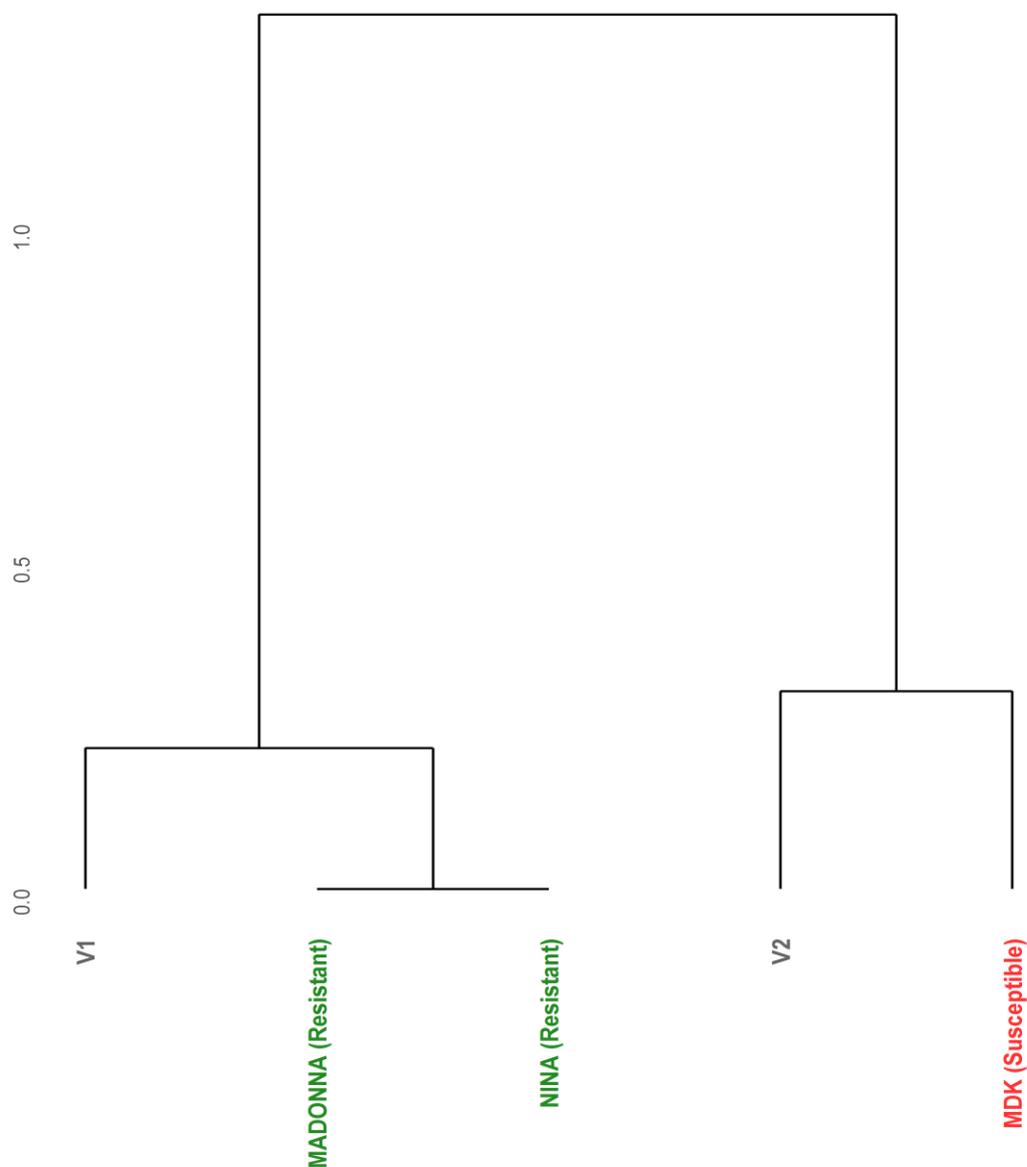
Variety	P-value	Result
V1	1.03e-01	Homogeneity of repetitions
V2	3.03e-01	Homogeneity of repetitions
MDK (Susceptible)	1.00e+00	Homogeneity of repetitions
MADONNA (Resistant)	5.48e-03	Heterogeneity of repetitions
NINA (Resistant)	1.00e+00	Homogeneity of repetitions

The following page of the document is devoted to the dendrogram of the tab « Study of the susceptibility / resistance of varieties ».

⚠ Just like stacked histogram of the previous page, size of the variety names on dendrogram have been reduced to keep a proportionality relationship between application and Word report. Font size for variety names is reduced of 6 compared to application. In other words, if we have a text size of axis lower than 7 within application, that one will be null or negative inside the Word document and the name of various varieties of the dendrogram will be not visible.

Study of the susceptibility / resistance of varieties

Dendrogram



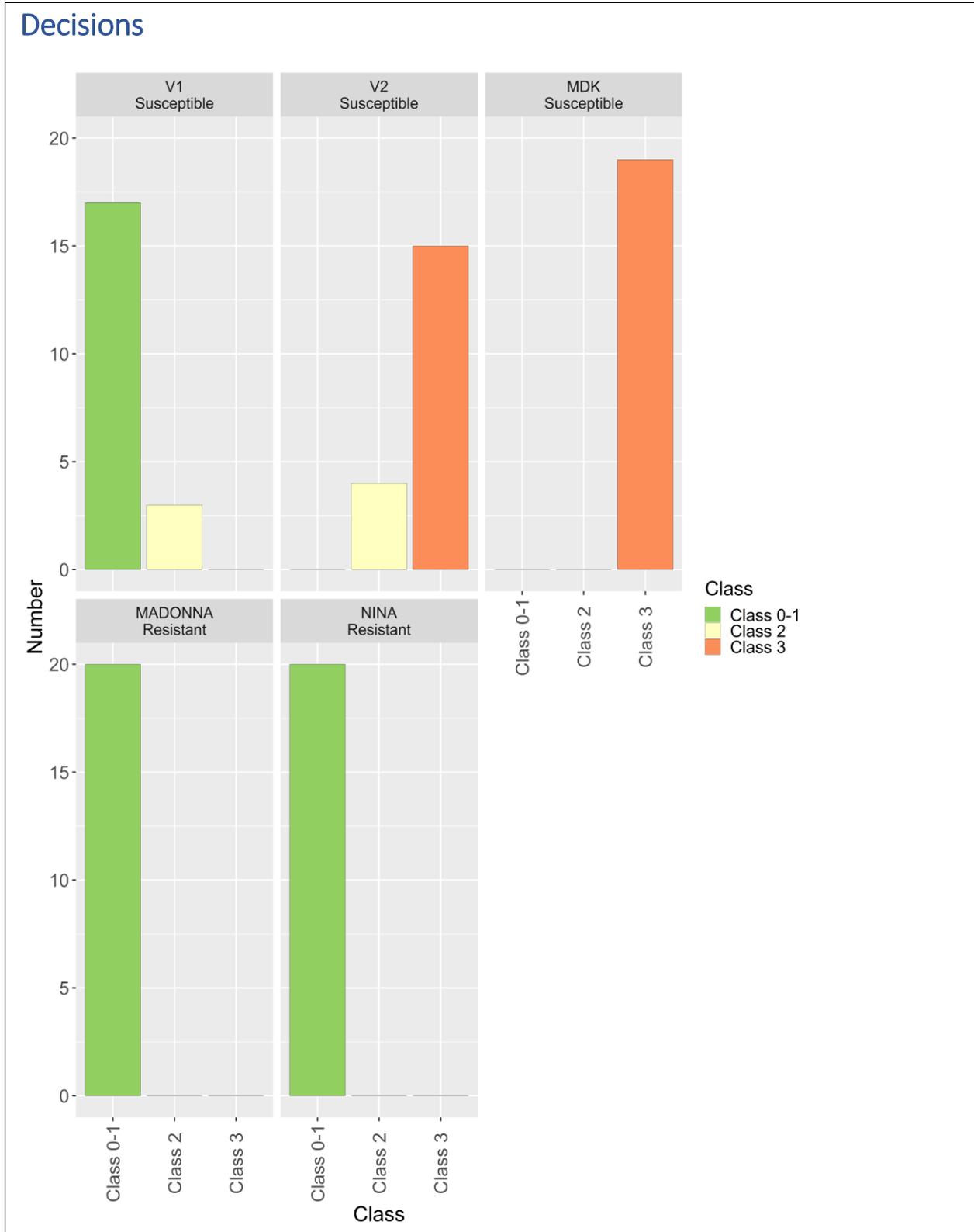
The last part of the Word report is devoted to the decisions related to Cochran-Armitage test for trend. The first page is a copy of the output of the tab of the same name in application.

△ As for the previous graphic output, according to keep a similar proportionality relationship between the font sizes (labels, axis and legend) of application and Word document, font size assigned within application is revised downwards when we compile the document :

- Labels size : it is reduced of 5 in the PDF document. In other words, if we assign a size of label lower than 6 within application, this one will be null or negative in the Word document and labels won't be visible.
- Text size of axis : it is reduced of 3 if we want to display results in the form of histogram and of 5 if a pie chart is preferred. In other words, if we assign a size of axis lower than 4 for

histogram and 6 for pie chart, within application, this one will be null or negative in Word document and axis and their title won't be visible.

- Legend size : it is reduced of 4 in Word document. In other words, if we assign a size of legend lower than 5 within application, this one will be null or negative in Word document and legend won't be visible.



Still in the same part related to the Cochran-Armitage test for trend, straight after presentation of results, we find a table detailing the whole of results. This table summarize the whole of the various Cochran-Armitage test for trend realised for each variety.

Detailed results

Variety	P-value (S)	P-value (R1)	P-value (R2)	Proposal
V1	8.77e-10	3.59e-02	3.59e-02	Susceptible
V2	1.72e-02	1.33e-09	1.33e-09	Susceptible

According to the kind of control variety and the CTPS decision rule, results and interpretation of the susceptibility / resistance of the variety are interpreted in different ways:

- Susceptible control (S): if p-value is higher than 0.05, we consider that the variety is not at a resistance level higher than control. If CTPS decision rule is « No control bound », « Susceptible control higher bound » and that p-value in comparison to susceptible control(s) is higher than 0.05, then variety will be declared susceptible. If CTPS rules is « Susceptible control higher bound » and that p-value is lower than 0.05, then variety will be declared resistant.
- Intermediate resistant control (IR): if p-value is higher than 0.05, we consider that the variety is not at a susceptibility level higher than control. We only have intermediate resistant control for CTPS decision rule « IR control lower bound ». Here, if p-value is higher than 0.05 for intermediate resistant control, variety will be declared intermediate resistant (unless if variety is at the level of resistant control), otherwise variety will be found susceptible.
- Resistant control (R): if p-value is higher than 0.05, we consider that variety is not at a susceptibility level higher than control. No matter CTPS decision rule, if p-value is higher than 0.05, variety will be always declared resistant. For CTPS decision rule « No control bound », if p-value is lower than 0.05, then variety will be declared either susceptible, heterogeneous or heterogeneous-continuum according to comparison to susceptible control. For CTPS decision rule « IR control lower bound », if p-value is lower than 0.05, variety will be found either susceptible or intermediate according to comparison to resistant intermediate control(s). For CTPS decision rule « Resistant control lower bound » and « Class 1 mobile », as soon as p-value is lower than 0.05, variety will be always found susceptible.
- High intermediate resistant control (HIR): We only found high intermediate resistant control in the case of test with CTPS decision rule « IR control lower bound ». For this control, if p-value is higher than 0.05, we consider that the variety will be high intermediate resistant. Otherwise, variety will be susceptible or intermediate resistant according to the resistance level of intermediate resistant control.

△ A variety can also be found heterogeneous according the heterogeneity rule. In other words, a variety with plants in both extreme classes, within the limits of the tolerated number of off-types, and with no plant in at least one of the intermediate classes will be found heterogeneous independently of the p-value in comparison with control varieties.