ISTA
Seed Quality Assurance

ISTA Secretariat
Richtiarkade 18, 8304 Wallisellen, Switzerland
Phone: +41 44 838 60 00 | Fax: +41 44 838 60 01
Email: ista.office@ista.ch
www.seedtest.org

Proficiency test (PT) report

ISTA PT23-SH 7-003

The ISTA Seed Health Proficiency Test for detection of *Botrytis cinerea* in sunflower (*Helianthus annuus*) seed samples

Date of publication: December 18th, 2024

N° of version: 1

Coordination	Full name	Position
Coordinator of ILC[1]	Thibaut DECOURCELLE	Coordinator of GEVES NRL[2] and ILC[1] activities
Validation of report's diffusion	Jaiana MALABARBA	Head of Pathology Laboratory and Responsible for GEVES NRL[2] for plant health
Coordinator and data validation	Justine FOUCHER	Methodology and Reference Manager - Project leader for GEVES NRL[2] for plant health
Person in charge of ILC[1] organization	Corinne SAHUGUEDE	Proficiency test organizer GEVES-SNES
Persons in charge of sample characterization tests	Céline LE GUISQUET Gaël CESBRON	Responsible for analysis in mycology GEVES-SNES

[1] : Inter-Laboratory Comparison (ILC)
[2] : National Reference Laboratory (NRL)

** Original report signed and archived*

Table of contents

Proficiency Test

1 PROFICIENCY TEST ORGANIZATION

The aim of this proficiency test was to verify the ability of laboratories to detect *Botrytis cinerea* on sunflower seeds. The schedule of this PT is presented in **Table 1**.

Table 1 : Schedule of the proficiency test

Sending of samples	8th of April 2024
Deadline to begin analysis	27th of May 2024
Deadline to send results	10th of June 2024
Sending by GEVES of report and individual sheet	18th of December 2024

Fourteen laboratories participated in this test and were randomly allocated a number, so that results remained anonymous.

1.1 Type of results

The laboratories indicated: qualitative and quantitative results for each sample and information about the method used.

1.2 Composition of the sample panel

11 samples of 400 seeds were sent to each laboratory. Three levels of contamination were represented in this panel, with a different number of samples per level as indicated in **Table 2**.

Table 2: Characteristics of samples in the panel

Level of contamination	Number of samples	Expected qualitative result	Expected percentage of infection (%)
Healthy	3	not detected	0
Medium	5	detected	5
Highly	3	detected	10

1.3 Statistical tools

Results of participants will be compared to the expected results defined by the results of homogeneity test and/or stability test.

The analysis of the results for a participating laboratory led to a declaration of conformity or non-conformity of the results in an individual sheet:
- “conform”: obtained results correspond to expected results.
- “not conform”: obtained results do not correspond to expected results.

1.3.1 Qualitative results

1.3.1.1 Diagnostic sensitivity, diagnostic specificity and accuracy

For the evaluation of diagnostic sensitivity and specificity, the analysis was done by addition of the results of the 2 lots with homogeneous samples (healthy and infected level) according to the Standard NF EN ISO 16140 for qualitative results. The expected values for each level were defined by GEVES independently of the results of the participating laboratories.

This norm gives assessment criteria on diagnostic sensitivity, diagnostic specificity and accuracy calculated as presented in **Table 3** and **Table 4**.

Table 3: Evaluation of criteria for diagnostic sensitivity, diagnostic specificity and accuracy

	Expected result + (infected sample)	Expected result – (healthy sample)
Obtained result +	Positive agreement +/+ (PA)	Positive deviation -/+ (PD)
Obtained result -	Negative deviation +/- (ND)	Negative agreement -/- (NA)

Table 4 : Conformity of results

Performance criteria	Level to obtain	Formula
Diagnostic sensitivity (SE)	100%: all infected samples are positive; no false negative results have been obtained	**$SE = (PA / N^{+}) * 100$**
Diagnostic specificity (SP)	100%: all healthy samples are negative; no false positive results have been obtained	**$SP = (NA/N^{-}) * 100$**
Accuracy (AC)	Synthesis of the two performance criteria. No false positive or negative results have been obtained	**$AC= (PA + NA) / N * 100$**

N= number of samples, N+ = number of infected samples, N- = number of healthy samples.

1.3.1.2 Rating system

On the qualitative data, the calculation of the rating is done with the Excel tool developed in collaboration with the Statistical committee of ISTA. It is based on an A, B, C and BMP rating. Numbers of samples identified as positive or negative are used for statistical analysis.

1.3.2 Quantitative results

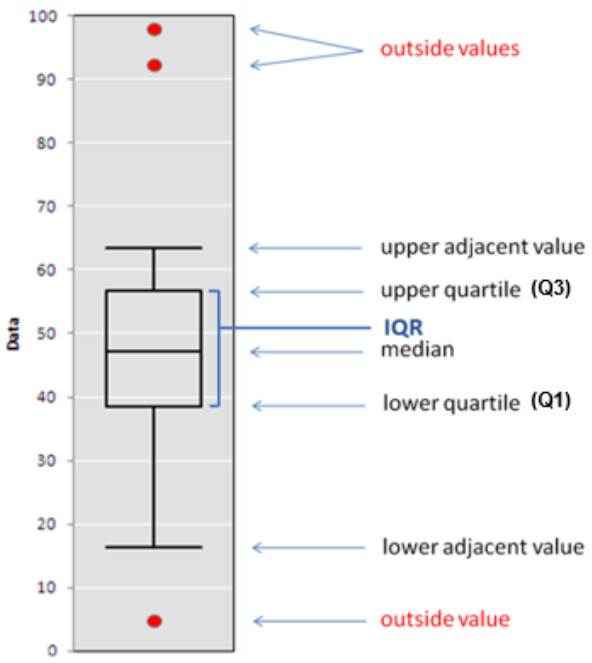
1.3.2.1 Hampel

Hampel's statistical test is used to detect outliers in a series of data. It is used for its robustness, which refers to its ability to provide accurate results even when the assumptions of normality are not respected.

The Hampel's statistical test works by comparing the value of each sample to the median value based on all samples. It is flagging the value that deviate by more than a specified number of median absolute deviations (MAD). To execute this test, an Excel tool has been developed by ISTA (Hampel's Test Calculations Example.xls) and is presented in the ISTA Guidelines, Organizing and analysing results of the Seed Health Proficiency tests. Hampel's statistical test could be used to evaluate the presence of outliers in the homogeneity test data, in the stability test data or in the participant's results.

1.3.2.2 Boxplot

The boxplot graphics can be used to visualize key statistical measures (**Figure 1**). The aim is to give a representation of the median, the variability of values and to identify aberrant values. To realize a boxplot, an Excel tool has been developed by ISTA (Box Plots.xls) and is available on the ISTA website.



Median: marks the mid-point of the data and is shown by the line that divides the box into two parts (50% of values will be lower, 50% of will be greater) sometimes known as the second quartile (Q2).

Upper quartile (Q3): 75% of the data are below the upper quartile value (also known as the third quartile). Thus, 25% of data are above this value.

Lower quartile (Q1): 25 % of the data are below the lower quartile value (also known as the first quartile). Thus, 75% of data are above this value.

IQR: inter Quartile Range (IQR) contained 50% of the data around the median. It is calculated by subtracting Q1 from Q3 (IQR = Q3 – Q1).

Upper adjacent value: greatest value excluding outliers. It corresponds to Q3 + 1.5*IQR.

Lower adjacent value: lowest value excluding outliers. It corresponds to Q1 - 1.5*IQR.

Outside value: data that are lower than the lower adjacent value, or that are greater that the upper adjacent value. It appears as separate red points.

Figure 1: a boxplot with key features labeled

1.3.2.3 Rating system

On the quantitative data, the calculation of the rating is done with the Excel file developed in collaboration with the Statistical committee of ISTA. It is based on an A, B, C and BMP rating.

The analysis is done using the Z scores, which correspond to the deviation from the average. It compares the mean value obtained by each laboratory (x) with the mean value of all laboratories (xi), considering the standard deviation of all laboratories (s): Z-score = (xi-x)/s. A Z score is calculated on each level of infection. For a level given, xi is estimated after removing values of laboratories that have outlier(s). To calculate Z-score, a tool was developed by ISTA and is presented in **Figure 2** below.

Process:
1. Sort the data on Lot, Lab and Rep, and indicate the value of each repetition.
2. For each Lot x Lab combination, compute the mean over the Rep (xi).
3. For each lot x Lab, perform Hampel's test for automatic detection of outliers. If no outlier is detected, the data kept corresponds to xi. If an outlier is detected the data kept is empty.
4. For each lot, compute mean and standard deviation (std-dev) of the data kept.
5. Z-score = (Data_kept - mean)/std-dev. For the healthy lot, the mean is set up to zero. If std-dev = 0, xi is reported in the column z-score.

1				2	3				4		5
					Hampel's test						
Lab	Lot	Rep	Value	x_i	Lot median (M)	$\lvert x_i - M \rvert$	MAD	Data kept	mean/Lot	std-dev/Lot	z-score
2	High	1	11.5								
2	High	2	9	10.83333333	7.38	3.458333333	2.21	10.83333333	7.21	3.26	1.11
2	High	3	12								
3	High	1	6								
3	High	2	6.5	6.5	7.38	0.875	2.21	6.5	7.21	3.26	-0.22
3	High	3	7								

Figure 2: Table used to calculate the Z score

1.4 Characterization of samples

1.4.1 Pre-test

Two seed lots were used in this PT and were characterized using the ISTA 7-003 method:

- Lot A: healthy lot
- Lot B: naturally infected lot (percentage of infection is around 30%)

The results are given in **Table 5**.

Table 5: Characterization of lots

Codification of lot	Level of infection expected	Number of tested samples	Level of infection obtained (%)	Comments *Botrytis cinerea*	Test execution date	Decision
A	Healthy	10	0		11/09/2023	Accepted
B	Infected by *Botrytis cinerea* (30%)	2	21.11 and 27.30	typical symptoms present, presence of a little saprophytic flora (*Rhizopus* sp.)	15/01/2024	Accepted

To illustrate the symptoms of *Botrytis cinerea* obtained for lot B see **Figure 3**. The picture presents the identification criteria for a naturally contaminated lot (i.e. rot from seedling to root and grey aerial mycelium). Examination by high magnification, the mycelium was sporulated and the spores were translucent and shiny.

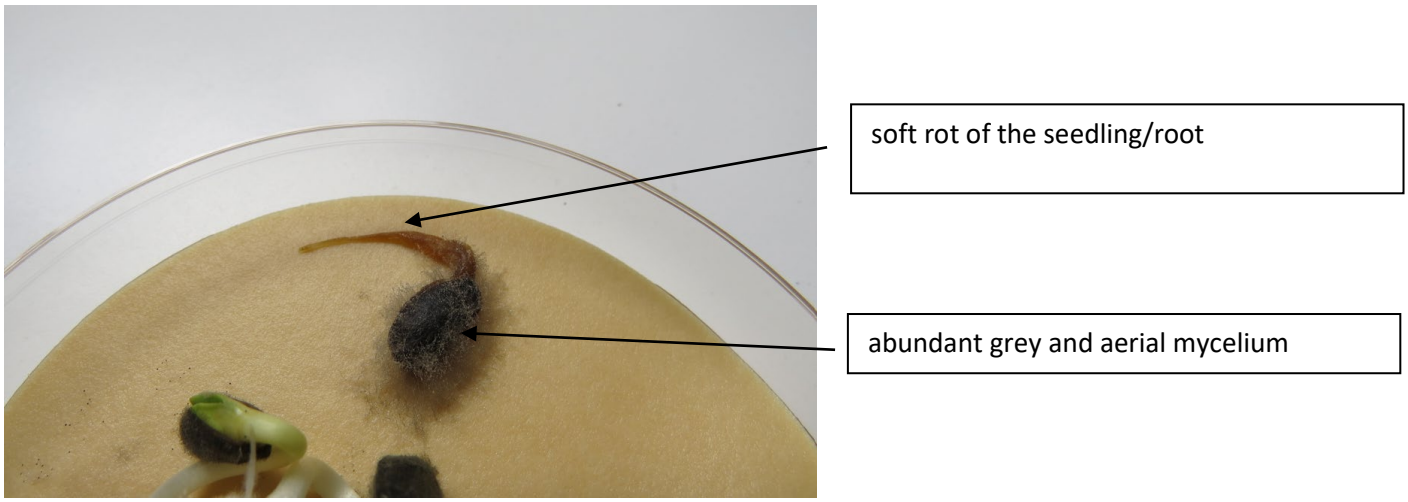


Figure 3: Picture of typical B. cinerea symptoms present in lot B. ©GEVES

The mixing of the 2 lots (Lot A + B) was tested to obtain 2 levels of infection: a medium infection at 5% and a highly infection at 10%.
The number of seeds from lot B to be added to lot A was calculated based on the level of infection (%) obtained during the characterization of lot B.
For each level of infection, 2 samples were analyzed to ensure the robustness of the result.
The results are given in **Table 6**.

Table 6: result of mixing the 2 lots A and B.

Level of contamination expected	Tested samples	Number of seeds		Level of infection obtained (%)	Decision
		Lot A = healthy	Lot B = infected		
Medium	1	305	95	6.33	Accepted
	2	305	95	6.77	
High	1	211	189	10.78	Accepted
	2	211	189	10.97	

Conclusion:
According to the results, three expected levels have been defined:

- Healthy level = Lot A
- Medium infected level = Lot M
- Highly infected level = Lot H

1.4.2 Homogeneity test

The homogeneity test was performed between 26th of February to 16th of March after packaging and just before shipping the seed samples to the participating laboratories. The method used was the ISTA 7-003 method. 10 samples of 400 seeds for each level of contamination were tested for the detection of *Botrytis cinerea* on sunflower seeds.

The homogeneity test results with the minimum, the maximum and the average values are given in **Table 7**. Expected results for each level are based on the pre-test results. The quantitative results were analyzed using Hampel's method to identify outliers. No outliers were found and the result is presented in Appendix 1.

Table 7: Results of homogeneity test

Codification of lot	Level of infection	Qualitative results			Quantitative results (%)		Deviation (%)	Conformity
		Expected	Obtained	Qualitative result	Expected results based on pre-test (%)	Obtained results (%)		
A	Healthy	Not detected	Not detected	$0^+/10$	0	0%	0	Conform
M	Medium infected	Detected	Detected	$10^+/10$	6.55	**4.51** ± 0.57	-2.04	Conform
H	Highly infected	Detected	Detected	$10^+/10$	10.87	**9.57** ± 1.99	-1.3	Conform

Conclusion of homogeneity test: the samples are homogeneous and there are no false positives at the healthy level. There is a slight decrease in the percentage of infection for both levels (medium and highly infected). For the medium infected lot, the percentage decreased to 2% and for the highly infected lot the percentage decreased to 1.3% compared to the pre-tests.

1.4.3 Stability Test

The stability test started on the 24th of May 2024 and ended on the 10th of June 2024. The method used was the ISTA 7-003 method. 25 samples were tested: 5 samples for the healthy level, 10 samples for the medium infected level and 10 samples for the highly infected level. The quantitative results were analyzed using Hampel's method to identify outliers. No outliers were found and the result is presented in Appendix 2. The results with the minimum, the maximum and the average values are given in **Table 8.**

Table 8: Results of stability test

Codification of lot	Level of infection	Qualitative results			Quantitative results (%)		Deviation (%)	Conformity
		Expected	Obtained	Qualitative result	Expected results based on homogeneity test (%)	Obtained results (%)		
A	Healthy	Not detected	Not detected	$0^+/5$	0	0%	0	Conform
M	Medium infected	Detected	Detected	$10^+/10$	4.51	**4.176** ± 0.76	-0.334	Conform
H	Highly infected	Detected	Detected	$10^+/10$	9.57	**7.375** ± 1.82	-2.195	Conform

Conclusion of stability test: The samples are homogeneous and there are no false positives results at the healthy level. There is a slight decrease in the percentage of infection for infected levels. For the medium infected level, the percentage decreased to 0.334% and for the highly infected level the percentage decreased to 2.195% compared to the homogeneity tests.

1.4.4 Overview of all tests carried out (pre-test, homogeneity and stability tests)

The comparison between homogeneity test results and stability test results are illustrated in **Figure 4**.

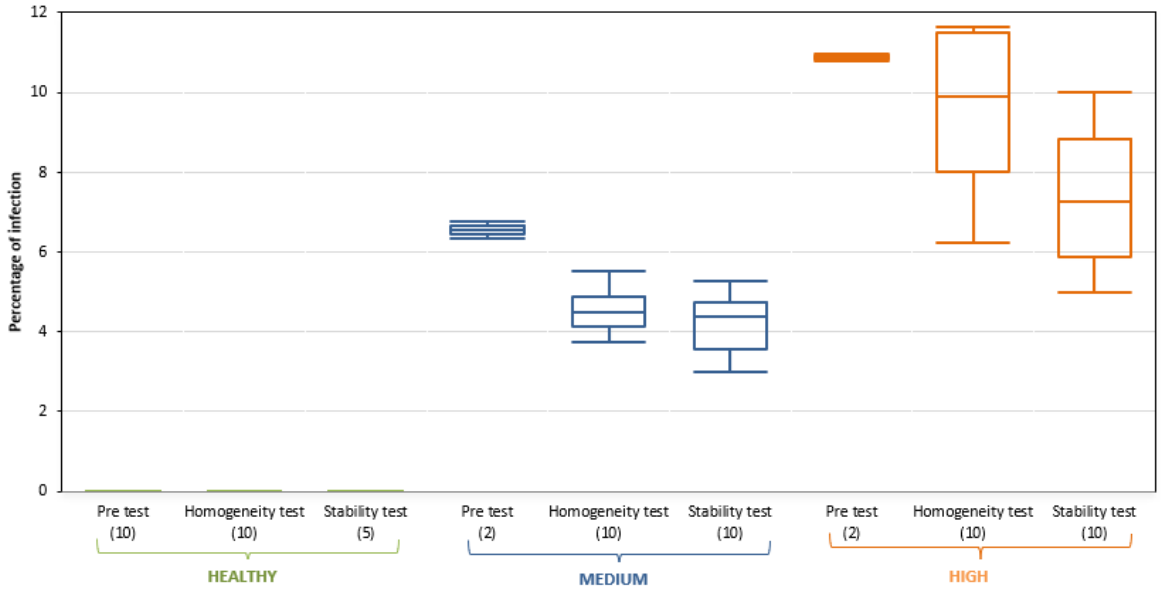


Figure 4: Boxplot representing the percentage of infection evaluated during the pre-test, the homogeneity and the stability tests. *Numbers in brackets correspond to the number of samples analyzed.*

Conclusion on the characterization of samples:

Results of healthy level

All samples were negative for the presence of *B. cinerea*. No false positive samples were observed during the process (pre-tests, homogeneity and stability tests). 25 samples of 400 seeds were tested, giving 0 infected seeds out of 10 000 seeds.

Results of medium infected level

Samples are homogeneous for the presence of *B. cinerea*, there was no outlier in homogeneity and stability tests (no red point into the graph and no outliers detected by Hampel's statistical test). The difference between the percentage of infection obtained between the pre-test and the stability test highlights a 2.324% decrease in infection over time. This decrease represents a loss of 35% of the initial percentage of infection (6.55%) and is mainly observed between the pre-test and the homogeneity test (- 2.04%).

Results of highly infected level

The difference between the percentage of infection obtained between the pre-test and the stability test highlight a 3.5% decrease, that corresponds to a loss of 32% of the initial percentage of infection (10.88%). The decrease in the presence of *B. cinerea* is mainly observed between the homogeneity and the stability test (- 2.195%).

2 PROFICIENCY TEST RESULTS

2.1 Qualitative results

2.1.1 Diagnostic specificity, diagnostic sensitivity and accuracy

The results of participating laboratories were compared to the expected results determined by the homogeneity and stability tests results. The raw data of all laboratories are given in appendix 3 and the results of laboratories are given in **Table 9**.

Table 9: Overview of qualitative results obtained by each laboratory on the healthy, medium and highly infected B. cinerea samples

Lab number	Healthy	Medium	High
10	$0^+/3$	$5^+/5$	$3^+/3$
11	$0^+/3$	**$4^+/5$**	$3^+/3$
12	**$1^+/3$**	$5^+/5$	$3^+/3$
13	$0^+/3$	$5^+/5$	$3^+/3$
14	$0^+/3$	$5^+/5$	$3^+/3$
17	$0^+/3$	$5^+/5$	$3^+/3$
18	**$2^+/3$**	$5^+/5$	$3^+/3$
19	$0^+/3$	$5^+/5$	$3^+/3$
20	$0^+/3$	$5^+/5$	$3^+/3$
21	$0^+/3$	$5^+/5$	$3^+/3$
22	$0^+/3$	$5^+/5$	$3^+/3$
23	$0^+/3$	$5^+/5$	$3^+/3$
24	$0^+/3$	$5^+/5$	$3^+/3$
26	$0^+/3$	$5^+/5$	$3^+/3$

(Cells in yellow correspond to lab results different from expected ones)

The obtained results for healthy, medium and highly infected samples during the homogeneity and stability test allow to calculate the 3 criteria (diagnostic sensitivity, diagnostic specificity and accuracy).

The performance criteria are based:

- For the diagnostic specificity on 3 negative samples (healthy samples)
- For the diagnostic sensitivity on 8 positive samples (3 highly infected samples and 5 medium infected samples)
- For the accuracy on 3 negative samples (healthy) and 8 positive samples (3 highly infected samples and 5 medium infected samples)

Results of participating laboratories and the percentage of diagnostic sensitivity, diagnostic specificity and accuracy for each laboratory are presented in **Table 10**.

Table 10: Criteria of performance for each laboratory (percentage obtained for diagnostic specificity, diagnostic sensitivity and accuracy)

Lab number	Diagnostic specificity	Diagnostic sensitivity	Accuracy
10	100%	100%	100%
11	100%	88%	91%
12	67%	100%	91%
13	100%	100%	100%
14	100%	100%	100%
17	100%	100%	100%
18	33%	100%	82%
19	100%	100%	100%
20	100%	100%	100%
21	100%	100%	100%
22	100%	100%	100%
23	100%	100%	100%
24	100%	100%	100%
26	100%	100%	100%

(Cells in yellow correspond to lab results different from expected ones)

Eleven out of 14 laboratories obtained 100% diagnostic sensitivity (no false negative) and 100% diagnostic specificity (no false positive). Three laboratories did not find 100% of accuracy, it is due to:

False negative results for 1 laboratory (Lab 11) that obtained 4 positive samples out of 5 for the medium level (with a diagnostic sensitivity = 88%).

False positive result for 2 laboratories (Lab 12; Lab 18) that obtained:

Lab 12: 1 positive sample out of 3 for the healthy level (with a sensitivity = 67%).

Lab 18: 2 positive samples out of 3 for the healthy level (with a sensitivity = 33%).

2.1.2 Rating system

In this case:

- A corresponds to no false positive in healthy level and no false negative in medium and highly infected levels.
- B corresponds to 1 false positive in healthy that has a percentage of contamination strictly less than 1%, and no false negatives in medium and highly infected levels.
- C corresponds to 2 false positives in healthy level that have a percentage of infection strictly less than 1% and/or 1 false negative in medium infected level and no false negative in highly infected level.
- BMP (Below Minimum Performance) corresponds to other types of results.

The calculation of the rating for each laboratory is presented in **Table 11** and the distribution of the rating is presented in **Figure 5**.

Table 11: Computation of ratings for each laboratory. *TH and TS correspond to homogeneity and stability tests respectively.*

Rating for qualitative SH PTs — Change any value in a yellow cel

Minimum requirements for A rating :

Max # of pos reps (Healthy lot)		Min # of pos reps (High level lot)		Min prob for observing k (Medium level lot)
0	and	3	and	5

Minimum requirements for B rating :

Max # of pos reps (Healthy lot)		Min # of pos reps (High level lot)		Min prob for observing k (Medium level lot)
1	and	3	and	5

Minimum requirements for C rating :

Max # of pos reps (Healthy lot)		Min # of pos reps (High level lot)		Min prob for observing k (Medium level lot)
2	and	3	and	4

Rating	Lab	Healthy lot # of pos reps	High level lot # of pos reps	# of seeds/rep: True cont. rate: # of reps tested: Medium level lot # of pos reps (k)
A	10	0	3	5
C	11	0	3	**4**
B	12	**1**	3	5
A	13	0	3	5
A	14	0	3	5
A	17	0	3	5
C	18	**2**	3	5
A	19	0	3	5
A	20	0	3	5
A	21	0	3	5
A	22	0	3	5
A	23	0	3	5
A	24	0	3	5
A	26	0	3	5
A	TH	0	3	5
A	TS	0	3	5

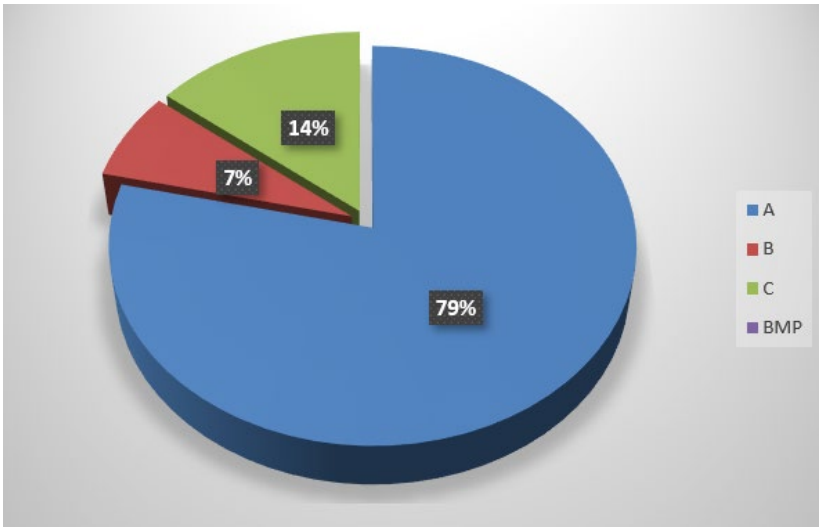


Figure 5: Percentage of laboratories according to the qualitative rating

The distribution of the rating evaluated on qualitative data demonstrates that note A represents 79% of the laboratories.

The B rating (Lab 12) is due to a false positive result in the healthy level with a percentage of 0.25%.

The C rating is due to:

1 false negative result in medium infected level (Lab 11)

2 false positive results in the healthy level (Lab 18).

2.2 Quantitative results

Raw data of all laboratories are given in Appendix 3. The **Figure 6** presents the obtained results by the laboratories and during the homogeneity and stability tests for each level.

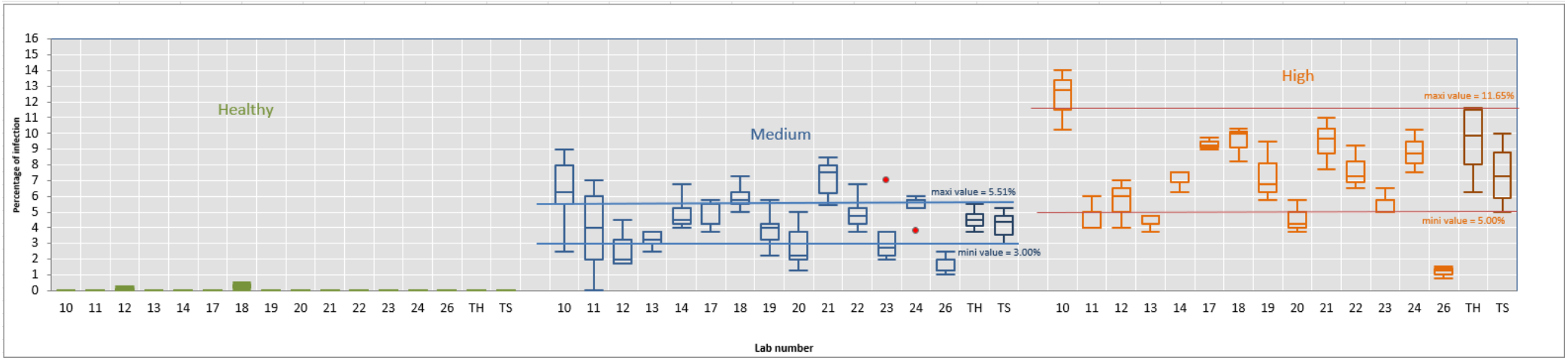


Figure 6: Boxplot presenting the percentage of infection obtained by each laboratory for each level of infection.

On the graph, red dots indicate outliers, only the medium level has outliers.

For **healthy level**: 12 out of 14 laboratories obtained a percentage of infection egal to 0% for the 3 replicates. One laboratory (Lab 12) obtained 0.25% infection for 1 replicate out of 3. Another laboratory (Lab 18) obtained two positive replicates out of 3 with 0.25% and 0.50% infection respectively.

No contamination was observed in the pre-test, homogeneity and stability test, but a low level of contamination of the healthy lot cannot be ruled out. Considering the number of seeds sown by fourteen participants and including all the tests carried out during the characterization, a total of 26 800 seeds have been sown. Three contaminated seeds will represent a potential contamination rate of 0.011%.

For **infected levels** :

- **Medium infected level:**

The limits were determined by the maximum and minimum values obtained in the homogeneity (max value = 5.51%) and stability tests (min value = 3.00%) indicated by blue lines in **Figure 6.**

The graph indicated two outliers (red points):

Lab 23: 1 outlier value (7.00%)

Lab 24: 1 outlier value (3.75%)

The results obtained can be divided into 3 groups:

Group 1: with 6 laboratories (Lab 11; Lab 13; Lab 14; Lab 17; Lab 19; Lab 22) whose results are in the expected interval. The result for Lab 11 shows a wide dispersion range (minimum value - maximum value) of 7, with 2 higher values (6% and 7%) and 2 lower values, one of which is 0%.

Group 2: with 4 laboratories (Lab 12; Lab 20; Lab 23; Lab 26) whose results are close but significantly lower than those of groups 1 and 3. The obtained results are above the minimum expected value:

- for 3 values out of 5 for 3 laboratories (Lab 12; Lab 20; Lab 23)

- for 5 values out of 5 for 1 laboratory (Lab 26)

Group 3: with 4 laboratory (Lab 10; Lab 18; Lab 21; Lab 24) whose obtained results over than the groups 1 and 2. The obtained results are above the maximum expected value:

- 3 values out of 5 for 3 laboratories (Lab 10; Lab 18; Lab 24)

- 4 values out of 5 for 1 laboratory (Lab 21).

- **Highly infected level**:

The limits were determined by the maximum and minimum values obtained in the homogeneity (max value = 11.65%) and stability tests (min value = 5.00%) indicated by red lines in **Figure 6**.

The results obtained is divided in 3 groups:

Group 4: with 9 laboratories (Lab 12; Lab 14; Lab 17; Lab 18; Lab 19; Lab 21; Lab22; Lab 23; Lab 24) that are included in the expected interval. For one laboratory (lab 12) one value is above the min value of 5%.

Group 5: with 4 laboratories (Lab 11; Lab 13; Lab 20; Lab 26) whose results are close but significantly lower than those of group 4. This is the group with the lowest results.

Group 6: with 1 laboratory (Lab 10) whose obtained results over than the groups 4 and 5. The obtained result shows that 2 value out of 3 are over the maximum expected value.

2.2.1 Rating system

Rule of decision

For the infected levels: Limits of acceptable z-scores for A, B, C and BMP will be dependent on limits given by quantile of a normal distribution (**Table 12**) and the calculation of the rating for each laboratory for their quantitative results are presented (**Table 13**).

Table 12: Decision rule for each note

Rating	Healthy level		Infected levels
A	0 infected seeds out of 3 samples (0%)	**and**	Z-score < 0.67 which corresponds to 0.75 quantile of a normal distribution. It means that the Z-score obtained by the laboratory is within 75% of the possible values following a normal distribution
B	1 false positive seed in 1 sample of the healthy level (0.25% of infection). Z-score corresponding to 1 infected seed in a sample = 0.08	**and**	Z-score between 0.67 and 1.5 which means between A and C. It means that the z-score obtained by the laboratory is within 87% of the possible values following a normal distribution
C	Z-score between 0.08 and 1.00 which corresponds to false positive samples with a % ≤ 1% of infected seed in each sample.	**and**	Z-score between 1.5 and 2.33 which corresponds to the 0.99 quantile of a normal distribution. It means that the z-score obtained by the laboratory is within 99% of the possible values following a normal distribution.
BMP	1 to 3 false positive samples with a % > 1% infected seed in each sample.	**and**	Z-score > 2.33

The calculation of the rating for each laboratory is presented in **Table 13** and the distribution of the rating is presented in **Figure 7**.

Table 13: Computation of ratings for each laboratory

Rating for quantitative SH PTs

Change any value in a yellow cell

	Max abs(z-score) for the Healthy lot		Max abs(z-score) for the Medium lot		Max abs(z-score) for the High lot
Minimum requirements for A rating :	0	and	0.67	and	0.67
Minimum requirements for B rating :	0.08	and	1.5	and	1.5
Minimum requirements for C rating :	1.00	and	2.33	and	2.33

Lab	abs(z-score) for the Healthy lot	abs(z-score) for the Medium lot	abs(z-score) for the High lot	Rating
10	0.00	1.33	1.89	C
11	0.00	0.38	0.91	B
12	0.08	1.18	0.54	B
13	0.00	0.76	1.00	B
14	0.00	0.43	0.03	A
17	0.00	0.25	0.79	B
18	0.25	1.12	0.86	C
19	0.00	0.31	0.06	A
20	0.00	1.04	0.94	B
21	0.00	1.95	0.84	C
22	0.00	0.42	0.19	A
23	0.00	0.56	0.60	A
24	0.00	0.67	0.61	A
26	0.00	1.92	2.19	C
TH	0.00	0.11	0.88	B
TS	0.00	0.12	0.08	A

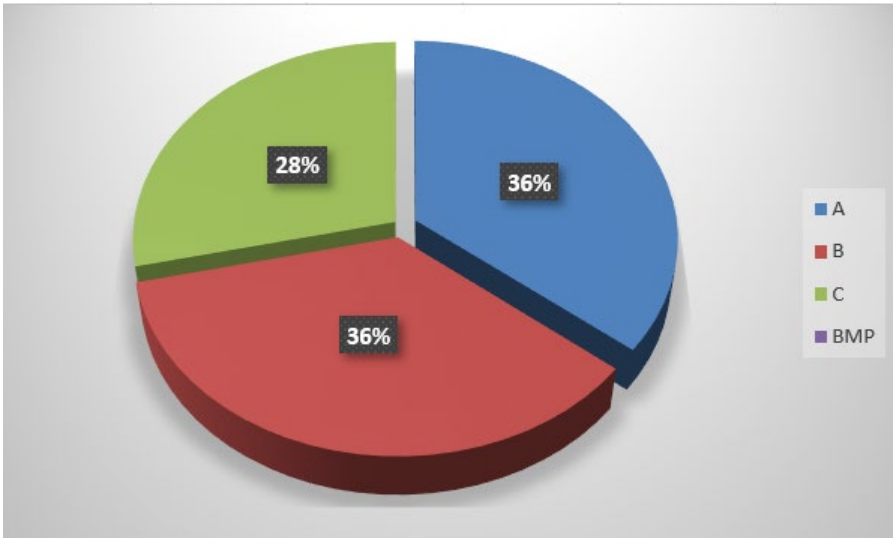


Figure 7: Percentage of laboratories according to the quantitative rating

3 CONCLUSIONS

Table 14 summarizes the different scores obtained for the different parts of this proficiency test. The final score corresponds to the lowest score obtained between qualitative and quantitative ratings.

Table 14: Summary of obtained ratings

Lab number	Final rating		
	Qualitative	Quantitative	Final
10	A	C	C
11	C	B	C
12	B	B	B
13	A	B	B
14	A	A	A
17	A	B	B
18	C	C	C
19	A	A	A
20	A	B	B
21	A	C	C
22	A	A	A
23	A	A	A
24	A	A	A
26	A	C	C

For this proficiency test, the A and B ratings represent 64% of the laboratories and no laboratory obtained a BMP rating. The percentage of laboratory for each rating is presented in **Figure 8**.

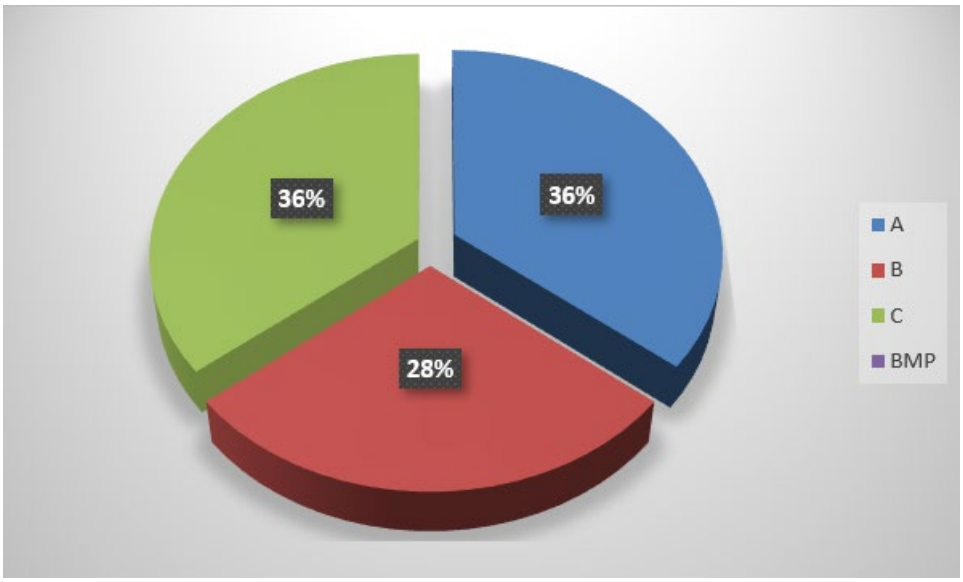


Figure 8: Percentage of laboratories with each rating

Figure 9 summarizes the rating obtained and the nature of the differences observed. The deviations observed allow to identify the laboratories whose results are in conformity at all levels or which are not in conformity at one or more levels:

Conform at all levels: 4 laboratories (Lab 14; Lab 17; Lab 19; Lab 22) obtained results within the expected limits for both infected levels.

For both levels (medium and highly infected), Lab 10 shows an overestimation while Labs 20 and 26 show an underestimation.

2 laboratories (Lab11 and Lab 13) show an underestimation of the highly infected levels, and 2 laboratories (Lab 12, Lab 23) show an underestimation of the medium infected levels.

3 laboratories (Lab 18; Lab 21; Lab 24) show an overestimation of the medium infected levels, these laboratories obtained percentages of infection between 5.3 and 7.1%.

Two laboratories obtained false positive results in healthy level (Lab 12 and Lab 18). The percentages obtained are minimal (between 0.25 % and 0.50 %). A cross contamination or a confusion in the identification of *Botrytis cinerea* could be responsible for this false positive.

Lab number	Rating	Impact on which level	Due to
10	C	Medium / High	over estimation
11	B	Medium	1 false negative
		High	underestimation
12	B	Healthy	1 false positive
		Medium	underestimation
13	B	High	underestimation
14	A	Healthy / Medium / High	in expected values
17	B	Healthy / Medium / High	in expected values
18	C	Healthy	2 false positives
		Medium	over estimation
19	A	Healthy / Medium / High	in expected values
20	B	Medium / High	underestimation
21	C	Medium	over estimation
22	A	Healthy / Medium / High	in expected values
23	A	Medium	underestimation and 1 outlier value
24	A	Healthy / Medium / High	in expected values
26	C	Medium / High	underestimation

Figure 9: Summary of ratings and deviations.

Appendix:

1) Identification of outliers on the homogeneity test data using Hampel's statistical test

Medium	Highly infected

MS Excel Hampels Outlier Test Example

Lab	Lab Values (Xi)	| Xi - M |	Status
1	4.52	0.020	OK
2	4.50	0.000	OK
3	5.51	1.010	OK
4	4.50	0.000	OK
5	3.75	0.750	OK
6	4.00	0.500	OK
7	3.77	0.730	OK
8	4.50	0.000	OK
9	5.01	0.510	OK
10	5.01	0.510	OK

Median (M):	4.500
MAD:	0.505
5.2 X MAD	2.626

MS Excel Hampels Outlier Test Example

Lab	Lab Values (Xi)	| Xi - M |	Status
1	11.59	1.705	OK
2	9.52	0.365	OK
3	6.24	3.645	OK
4	11.65	1.765	OK
5	11.25	1.365	OK
6	11.59	1.705	OK
7	10.25	0.365	OK
8	8.04	1.845	OK
9	7.58	2.305	OK
10	8.02	1.865	OK

Median (M):	9.885
MAD:	1.735
5.2 X MAD	9.022

2) Identification of outliers on the stability test data using Hampel's statistical test

Medium	High

MS Excel Hampels Outlier Test Example

Lab	Lab Values (Xi)	| Xi - M |	Status
1	3.00	1.375	OK
2	5.26	0.885	OK
3	4.25	0.125	OK
4	3.25	1.125	OK
5	3.75	0.625	OK
6	4.75	0.375	OK
7	4.75	0.375	OK
8	4.50	0.125	OK
9	3.50	0.875	OK
10	4.75	0.375	OK

Median (M):	4.375
MAD:	0.500
5.2 X MAD	2.600

MS Excel Hampels Outlier Test Example

Lab	Lab Values (Xi)	| Xi - M |	Status
1	7.50	0.250	OK
2	9.75	2.500	OK
3	7.25	0.000	OK
4	5.50	1.750	OK
5	9.25	2.000	OK
6	5.25	2.000	OK
7	10.00	2.750	OK
8	5.00	2.250	OK
9	7.00	0.250	OK
10	7.25	0.000	OK

Median (M):	7.250
MAD:	1.875
5.2 X MAD	9.750

3) Raw results of participants

Lab number	Level of contamination	Number of samples	Qualitative results			Quantitative results			
			Expected	Obtained	Finald result	Obtained results	Average	Median	Expected results
Lab 10	Healthy	138	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		173				0.00			
		291				0.00			
	Medium	41	$5^+/5$	Detected	$5^+/5$	9.00	6.25	6.25	5.00
		26				2.50			
		87				6.25			
		341				5.50			
		311				8.00			
	High	283	$3^+/3$	Detected	$3^+/3$	14.00	12.33	12.75	10.00
		110				12.75			
		40				10.25			
Lab 11	Healthy	184	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		366				0.00			
		99				0.00			
	Medium	164	$5^+/5$	Detected	$4^+/5$	7.00	3.80	4.00	5.00
		350				2.00			
		65				6.00			
		123				4.00			
		120		not detected		0.00			
	High	156	$3^+/3$	Detected	$3^+/3$	4.00	4.67	4.00	10.00
		327				4.00			
		32				6.00			
Lab 12	Healthy	145	$0^+/3$	not detected	$1^+/3$	0.00	0.08	0.0000	0.00
		63				0.00			
		219		Detected		0.25			
	Medium	71	$5^+/5$	Detected	$5^+/5$	1.75	2.65	2.00	5.00
		43				3.25			
		161				1.75			
		47				4.50			
		15				2.00			
	High	153	$3^+/3$	Detected	$3^+/3$	6.00	5.67	6.00	10.00
		266				4.00			
		357				7.00			
Lab 13	Healthy	368	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		343				0.00			
		241				0.00			
	Medium	93	$5^+/5$	Detected	$5^+/5$	3.75	3.25	3.25	5.00
		279				3.25			
		13				3.00			
		220				3.75			
		150				2.50			
	High	317	$3^+/3$	Detected	$3^+/3$	4.75	4.42	4.75	10.00
		133				4.75			
		200				3.75			
Lab 14	Healthy	203	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		260				0.00			
		141				0.00			
	Medium	261	$5^+/5$	Detected	$5^+/5$	4.00	4.96	4.50	5.00
		361				4.50			
		230				6.77			
		302				4.25			
		344				5.28			
	High	174	$3^+/3$	Detected	$3^+/3$	7.50	7.08	7.50	10.00
		20				7.50			
		19				6.25			
Lab 17	Healthy	148	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		347				0.00			
		224				0.00			
	Medium	215	$5^+/5$	Detected	$5^+/5$	5.50	4.70	4.25	5.00
		256				3.75			
		326				4.25			
		151				4.25			
		216				5.75			
	High	35	$3^+/3$	Detected	$3^+/3$	9.25	9.33	9.25	10.00
		27				9.75			
		277				9.00			

Lab number	Level of contamination	Number of samples	Qualitative results			Quantitative results			
			Expected	Obtained	Finald result	Obtained results	Average	Median	Expected results
Lab 18	Healthy	358	$0^+/3$	Detected	$2^+/3$	0.50	0.25	0.25	0.00
		365		not detected		0.00			
		374		Detected		0.25			
	Medium	243	$5^+/5$	Detected	$5^+/5$	6.25	5.95	5.75	5.00
		81				7.25			
		182				5.00			
		265				5.75			
		177				5.50			
	High	113	$3^+/3$	Detected	$3^+/3$	10.00	9.51	10.00	10.00
		240				8.25			
		159				10.28			
Lab 19	Healthy	83	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		73				0.00			
		152				0.00			
	Medium	292	$5^+/5$	Detected	$5^+/5$	5.75	3.90	4.00	5.00
		275				4.25			
		363				2.25			
		360				3.25			
		232				4.00			
	High	280	$3^+/3$	Detected	$3^+/3$	9.50	7.33	6.75	10.00
		301				6.75			
		106				5.75			
Lab 20	Healthy	70	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		115				0.00			
		98				0.00			
	Medium	273	$5^+/5$	Detected	$5^+/5$	1.25	2.85	2.25	5.00
		30				2.00			
		310				2.25			
		270				3.75			
		103				5.00			
	High	320	$3^+/3$	Detected	$3^+/3$	4.25	4.58	4.25	10.00
		204				3.75			
		294				5.75			
Lab 21	Healthy	287	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		143				0.00			
		167				0.00			
	Medium	205	$5^+/5$	Detected	$5^+/5$	6.23	7.14	7.50	5.00
		129				5.47			
		74				8.50			
		140				7.98			
		371				7.50			
	High	39	$3^+/3$	Detected	$3^+/3$	7.71	9.46	9.70	10.00
		121				10.97			
		229				9.70			
Lab 22	Healthy	296	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		329				0.00			
		58				0.00			
	Medium	262	$5^+/5$	Detected	$5^+/5$	5.25	4.95	4.75	5.00
		118				4.75			
		308				4.25			
		199				6.75			
		163				3.75			
	High	290	$3^+/3$	Detected	$3^+/3$	7.25	7.67	7.25	10.00
		330				9.25			
		276				6.50			

Lab number	Level of contamination	Number of samples	Qualitative results			Quantitative results			
			Expected	Obtained	Finald result	Obtained results	Average	Median	Expected results
Lab 23	Healthy	162	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		18				0.00			
		146				0.00			
	Medium	251	$5^+/5$	Detected	$5^+/5$	2.75	3.55	2.75	5.00
		221				2.25			
		5				2.00			
		197				3.75			
		183				7.00			
	High	96	$3^+/3$	Detected	$3^+/3$	6.50	5.50	5.00	10.00
		246				5.00			
		352				5.00			
Lab 24	Healthy	132	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		69				0.00			
		346				0.00			
	Medium	91	$5^+/5$	Detected	$5^+/5$	3.75	5.30	5.75	5.00
		155				5.75			
		79				5.25			
		255				6.00			
		104				5.75			
	High	195	$3^+/3$	Detected	$3^+/3$	10.25	8.83	8.75	10.00
		172				8.75			
		53				7.50			
Lab 26	Healthy	253	$0^+/3$	not detected	$0^+/3$	0.00	0.00	0.00	0.00
		57				0.00			
		16				0.00			
	Medium	298	$5^+/5$	Detected	$5^+/5$	1.25	1.60	1.25	5.00
		127				1.00			
		322				2.50			
		77				2.00			
		55				1.25			
	High	38	$3^+/3$	Detected	$3^+/3$	1.50	1.17	1.25	10.00
		351				1.25			
		90				0.75			