This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773139

Grant agreement N. 773139

DELIVERABLE N° 2.3

Title:

Scientific publication of the results of the validation of approach applied on five existing datasets

Validation of diagnostic tests to support plant health

Due date:	Month M39
Actual submission date	30-09-2021 (Month M41)
Start date of the project	01-05-2018
Deliverable lead contractor (organization name)	ULg
Participants (Partners short names)	ULg, EPPO, ANSES, CREA, FERA, NIB, UNITO, NVWA
Author(s) in alphabetical order	Brostaux Y., Lebas B., Massart S.
Contact for queries	sebastien.massart@uliege.be
Level of dissemination	Public
Type of deliverable	Report

Abstract

This deliverable summarizes the application of the statistical analyses on 12 datasets obtained from the test performance studies (TPS) conducted in the framework of VALITEST (as part of Work package 1). The application of the statistical approach on a large number of datasets throughout the project, e.g. including the datasets generated during the first round of TPS, and the iterative exchanges with TPS organizers from round 1 allowed to improve the analyses carried out. This document details the results of the application of the new and improved statistical approach on 5 datasets from TPS round 1 and round 2 in Annexes.

Partners involved: ULg, EPPO, ANSES, CREA, FERA, NIB, UNITO, NVWA

TABLE OF CONTENTS

1 Purpose

The proper validation of any diagnostic test, e.g. the determination of its performance characteristics, is a mandatory prerequisite before its application in plant pest diagnostics. Such evaluation can be carried out within the laboratory (intra-laboratory) or in the framework of test performance studies by two or more laboratories. The evaluation, carried out on a panel of reference samples, can include a single test or compare several tests simultaneously.

An appropriate statistical analysis increases the confidence in the conclusions drawn from the validation data. The use of statistics in data processing during intra- and inter-laboratories studies facilitates the interpretation and comparison between tests for a given performance criterion. The clear added value of statistical tools is also to provide confidence intervals and p-values associated with each estimate. For a given test, this allows a better interpretation of each calculated performance criterion for the intended use. This also allows the possibility to statistically compare tests for a given performance characteristic.

The purpose of this document is to demonstrate the applicability of the new and improved statistical approach described in the VALITEST deliverable report D2.1 (Guidelines for the revision of the EPPO Standards PM 7/98 and PM 7/122 for validation studies (including Test Performance Studies)) by analysing the datasets generated during the VALITEST test performance studies (as part of Work package 1).

2 Scope

This document describes the application of the statistical analyses of diagnostic tests validation data generated during intra-laboratory validation studies or test performance studies. Some statistical analyses outlined in this document can also be applicable to the analysis of data generated during proficiency tests programmes (e.g. diagnostic sensitivity, diagnostic specificity or repeatability).

The document is relevant to plant health diagnostic laboratories that perform validation studies before using a test routinely and organisers of test performance studies and proficiency testing studies. The scope of test includes the detection and/or identification of plant pests (e.g. arthropods, bacteria, fungi, nematodes, invasive plants, protozoan, viroids, viruses or weeds) from any types of matrices (e.g. pure microbial culture, plant tissue, soil, water).

3 Introduction

Diagnostic tests need to be appropriately validated before their use as a routine test in a plant health diagnostic laboratory. Test validation (or test verification if using a validated test) is mandatory to be accredited to ISO 17025 standard which will be a requirement for all national reference laboratories in the European Union by 29 April 2022 (EU Regulation 2017/625).

A new and improved statistical approach for the analysis of validation data has been proposed and described in VALITEST deliverable report 2.1. The repeatability (within a laboratory) and reproducibility (between laboratories) of a diagnostic test are estimated by calculating the accordance and concordance, respectively (Langton *et al.*, 2002). The analytical sensitivity is evaluated using a probability of detection model (Wehling *et al.*, 2011; Uhlig and Gowik, 2018). The diagnostic sensitivity and the diagnostic specificity for the validation of a new test by comparison with a validated test is determined using the calculation explained in EPPO standard PM 7/98 (2019). We also propose to broaden the EPPO approach by including the option for a newly developed test using

reference samples when, for example the new test is the only available test for the detection of a pest using reference material identified by another method, such as morphology, biochemical tests.

During the test performance studies conducted in the framework of VALITEST (as part of Work package 1), it appeared that more performance criteria for measuring the effectiveness of diagnostic tests may be valuable and were also proposed, as optional, in the data analyses. These additional performance criteria are diagnostic odd ratio, false positive and negative rates, rates of true positive and true negative, and positive and negative likelihood ratios. We believe that the plant health diagnostic laboratories would benefit from such analyses. For example, for the selection of the most appropriate test for the detection of a pest.

In addition, the confidence intervals were also proposed for most of the performance criteria. These confidence intervals are particularly useful when comparing the performance of several tests. The report also included information on how to establish the panel of samples, how to deal with inconclusive and missing results and how to identify and deal with outlier results.

The present document summarizes the results of the application of the new and improved statistical approach on 12 datasets obtained as part of test performance studies performed within the framework of VALITEST (under Work package 1) and provides details on 5 datasets from TPS round 1 and round 2 (Annexes 1 to 5).

4 Application of the new and improved statistical approach

4.1 Performance criteria analysis

The pests for the first round were *Bursaphelenchus xylophilus*, citrus tristeza virus, *Erwinia amylovora* (molecular and serological tests), *Fusarium circinatum*, *Pantoea stewartii* subsp. *stewartii* and plum pox virus. A total of 7 analyses (one method per pest + both methods for *E. amylovora*) have been carried out for the round 1.

The application of the new and improved validation approach on the TPS data was not always possible due to the lack of data for some statistical tools to perform correctly (e.g. no repeatability or reproducibility data, the number of dilution points).

The application of the new validation approach on data from TPS round 1 showed that:

- The determination of the repeatability and reproducibility should be done using the concordance and accordance of Langton *et al*. (2002).
- The determination of the analytical sensitivity using the probability of detection model performed well when there is a minimum of five dilution points.
- The determination of the diagnostic sensitivity and the diagnostic specificity as per EPPO PM7/98 (2019) for comparison of tests results with the status of reference material is better adapted in these TPS compared to the relative comparison between an existing validated test and a new test
- More criteria have been used by TPS organisers and could be useful to determine the performance of diagnostic tests. They have been summarized, integrated in the analyses and applied on the seven data sets from round 1.
- The identification of outliers can rely on several analyses and has been thoroughly described in the deliverable D2.1

The pests selected for the second round were *Cryphonectria parasitica*, plum pox virus (on-site tests), tomato brown rugose fruit virus, tomato spotted wilt virus, *Xylophilus ampelinus* and *Xanthomonas citri* pv. *citri*.

The statistical approach developed in the deliverable 2.1 have been carried out on 12 datasets obtained from the test performance studies conducted in the framework of VALITEST (as part of Work package 1) and most TPS organizers will include them in the corresponding publication. Therefore, in order not to interact with the planned publications (see chapter 5), this deliverable details the statistical analyses of TPS results for five pests from round 1 and from round 2, for which the agreement of the TPS organizer for publishing these analyses as an Annex has been obtained (see chapter 6).

4.2 Composition of sample panel

The composition of the sample panel plays a crucial role. The data generated during validation studies and test performance studies need to be of sufficient quantity for the statistical methods to perform correctly. Thus, a sample panel composed of 25 samples was proposed for the 2nd round of TPSs in the VALITEST deliverable report 2.1. taking into account that the composition of the sample panel is influenced by many constraints: statistical power of the tests, availability of financial resources and reference samples, number of participating laboratories, risk analysis determining critical performance criteria based on preliminary test characteristic evaluation by the organiser of an intra-laboratory validation study / test performance study, intended use of the test, etc. The proposed sample panel can therefore be adapted to these constraints.

Table 1 shows a summary of sample panels used during the second round of test performance studies in the framework of the VALITEST project. This table exemplifies the adaptations made by the organisers depending on the different constraints and the scope of their test performance studies.

Table 1. Information on the panel composition used during the second round of the VALITEST test performance study.

Pest	Minimum number of participants*	Maximum number of participants*	Number of samples in the panel	Use of biological replicates?	Number of samples infested with the target pest	Number of samples free from the target pest	Among which samples with close relative to the target pest	Dilution points	Replicates per dilution point	Scope of test performance study
Cryphonectria parasitica (Cp)	7	10	16	No	3	2	Yes	2	1 to 3	Detection of Cp in symptomatic and asymptomatic wood material using molecular tests
plum pox virus (PPV)	9	14	22	Yes	5	4	No	2	2	Detection of PPV in symptomatic and asymptomatic leaves of *prunus* spp. using serological and molecular on-site tests
tomato brown rugose fruit virus (ToBRFV)	26	34	22	No	2	2	No	5	2 for the least diluted 3 for other dilutions	Detection and identification of ToBRFV in symptomatic and asymptomatic leaves and fruit of tomato and pepper using serological and molecular tests
tomato spotted wilt virus (TSWV)	13	17	22	Yes	2	7 (depends on the method)	Yes	5	2 for the least diluted 3 for other dilutions (depends on the method)	Detection and identification of TSWV in symptomatic leaves of tomato using serological and molecular tests
Xylophilus ampelinus (Xa)	3	11	22	Yes	2	2	Yes	5	2 for the least diluted 3 for other dilutions	Detection of Xa in symptomatic and asymptomatic stem material using serological and molecular tests

*if several tests were included, the least number of participants and the highest number of participants for a test are provided

5 Publication of the validation approach

The statistical analyses proposed in the deliverable 2.1 will be published in the EPPO Bulletin with an anonymised dataset as example. Moreover, some TPS organizers will include the application of the statistical analyses on their datasets in publications.

Table 2. Planned publications with the application of the statistical analyses on the datasets of the TPS

TPS Round 1 and 2	Planned date of submission of the publications
TPS citrus tristeza virus (ANSES)	November 2021
TPS *Xanthomonas citri* pv. *citri* (ANSES)	January 2022
TPS *Xylophilus ampelinus* (FERA)	November 2021
TPS *Cryphonectria parasitica* (UNITO)	October 2021
TPS tomato brown rugose fruit virus (CREA)	December 2021

6 Detailed analyses carried out

The application of the statistical analyses on the 12 datasets from TPS round 1 and 2 of Valitest by WP2 has generated numerous amounts of information. This information has been sent to each of the TPS organizer to support results interpretation. The statistical analyses of TPS results for five pests from round 1 and from round 2, for which the agreement of the TPS organizer for publishing these analyses as an annex has been obtained are detailed in Annexes 1 to 5 (see Table 3).

Table 3. Statistical analyses on the TPS

Pest	Round	Method	Annex number
Bursaphelenchus xylophilus	1	Molecular	1
Fusarium circinatum	1	Molecular – Plating	2
Cryphonectria parasitica	2	Molecular and serological	3
plum pox virus (PPV)	1	Molecular	4
tomato spotted wilt virus (TSWV)	2	Molecular and serological	5

7 Conclusion

The application of the proposed framework of statistical analyses on TPS datasets allowed to iteratively improve the proposed approach, including the sample panel (after analysis of 1st round) and to provide an in-depth analysis of the validation results for all TPS (1st and 2nd rounds) which support the interpretation of the results for the TPS organizers. The use of the proposed approach will be integrated in several ongoing publications of the TPS.

References

European Plant Protection Organization (2014) PM7⁄122 (1) - Guidelines for the organization of interlaboratory comparisons by plant pest diagnostic laboratories. Bulletin OEPP/EPPO Bulletin, 44 (3): 390–399, https://onlinelibrary.wiley.com/doi/epdf/10.1111/epp.12162.

European Plant Protection Organization (2019). PM 7⁄98 (4) - Specific requirements for laboratories preparing accreditation for a plant pest diagnostic activity. *Bulletin OEPP/EPPO Bulletin*, 49 (3): 530-563, https://onlinelibrary.wiley.com/doi/epdf/10.1111/epp.12629.

ISO 17025 (2017) General requirements for the competence of testing and calibration laboratories. International organization for standardization, Geneva, Switzerland.

Langton S.D., Chevennement R., Nagelkerke N., Lombard B. (2002) Analysing collaborative trials for qualitative microbiological methods: accordance and concordance. *International Journal of Food Microbiology*, 79: 175-181.

Uhlig, S., Gowik P. (2018) Efficient estimation of the limit of detection and the relative limit of detection along with their reproducibility in the validation of qualitative microbiological methods by means of generalized linear mixed models. *Journal of consumer protection food safety*, 13: 79-87.

Wehling P., LaBudde R.A., Brunelle S.L., Nelson M.T. (2011) Probability of detection (POD) as a statistical model for the validation of qualitative methods. *Journal of AOAC International*, 94 (1): 335-347.

VALITEST TPS report

BX1 data

Yves Brostaux

26 juin 2020

Files reading and data preprocessing

BX1 data are stored in one csv files, with the following columns :

- sampid, sample ID
- method, method ID
- lab, lab ID
- test, result of the test (binary 0/1)
- ref, reference status of each sample (binary 0/1)
- dilution (dilution factor for diluted samples)
- link, ID of the original sample of the dilution serie

Those files have been read and combined into different data tables for the analysis.

- tidydata, 1 line per status result, same organization as the file
- dilu.dat, subset of tidydata including only diluted samples, with additional column
  - dilu, dilution exponent (negative base 10)
- spse.dat, subset of tidydata including only non diluted samples

```
## tidydata
```

```
##   sampid        method lab test ref dilution link
## 1      1 Burgermeister L01    0   0       NA   NA
## 2      2 Burgermeister L01    0   0       NA   NA
## 3      3 Burgermeister L01    0   0       NA   NA
## 4      4 Burgermeister L01    0   0       NA   NA
## 5      5 Burgermeister L01    0   0       NA   NA
## 6      6 Burgermeister L01    0   0       NA   NA
```

```
## dilu.dat
```

```
##    sampid        method lab test ref dilution link dilu
## 7       7 Burgermeister L01    1   1     1.00   NA    0
## 8       8 Burgermeister L01    1   1     0.10    7    1
## 9       9 Burgermeister L01    0   1     0.01    7    2
## 17      7 Burgermeister L02    1   1     1.00   NA    0
## 18      8 Burgermeister L02    1   1     0.10    7    1
## 19      9 Burgermeister L02    1   1     0.01    7    2
```

```
## spse.dat
```

```
##   sampid        method lab test ref dilution link
## 1      1 Burgermeister L01    0   0       NA   NA
## 2      2 Burgermeister L01    0   0       NA   NA
## 3      3 Burgermeister L01    0   0       NA   NA
## 4      4 Burgermeister L01    0   0       NA   NA
## 5      5 Burgermeister L01    0   0       NA   NA
## 6      6 Burgermeister L01    0   0       NA   NA
```

Note that the number of samples tested vary between methods and laboratories. This will impair the comparison of methods based on calculated parameters, as their basis differ.

```
##      method
## lab   Burgermeister ClearDetections François Kikushi Matsunaga
##   L01            10               20       20       0        10
##   L02            10               20       20       0        10
##   L03            10               20       20       0        10
##   L04            10               15       15      15        10
##   L05             0                0       20       0        10
##   L06            10                0       20       0        10
##   L08            10               20       20      20        10
##   L09            10               15       15       0        10
##   L10             0               20       20      20        10
##   L11            10               20       20      20        10
##   L12             0               20       20       0         0
##   L13            10                0        0       0        10
##   L15             0               20       20       0         0
##   L16             0               20       20       0         0
##   L17             0                0       20       0         0
##   L20            10                0        0       0        10
##   L21            10                0        0       0        10
##   L22            10                0        0       0        10
##   L23            10                0        0       0        10
##   L24            10                0        0       0        10
```

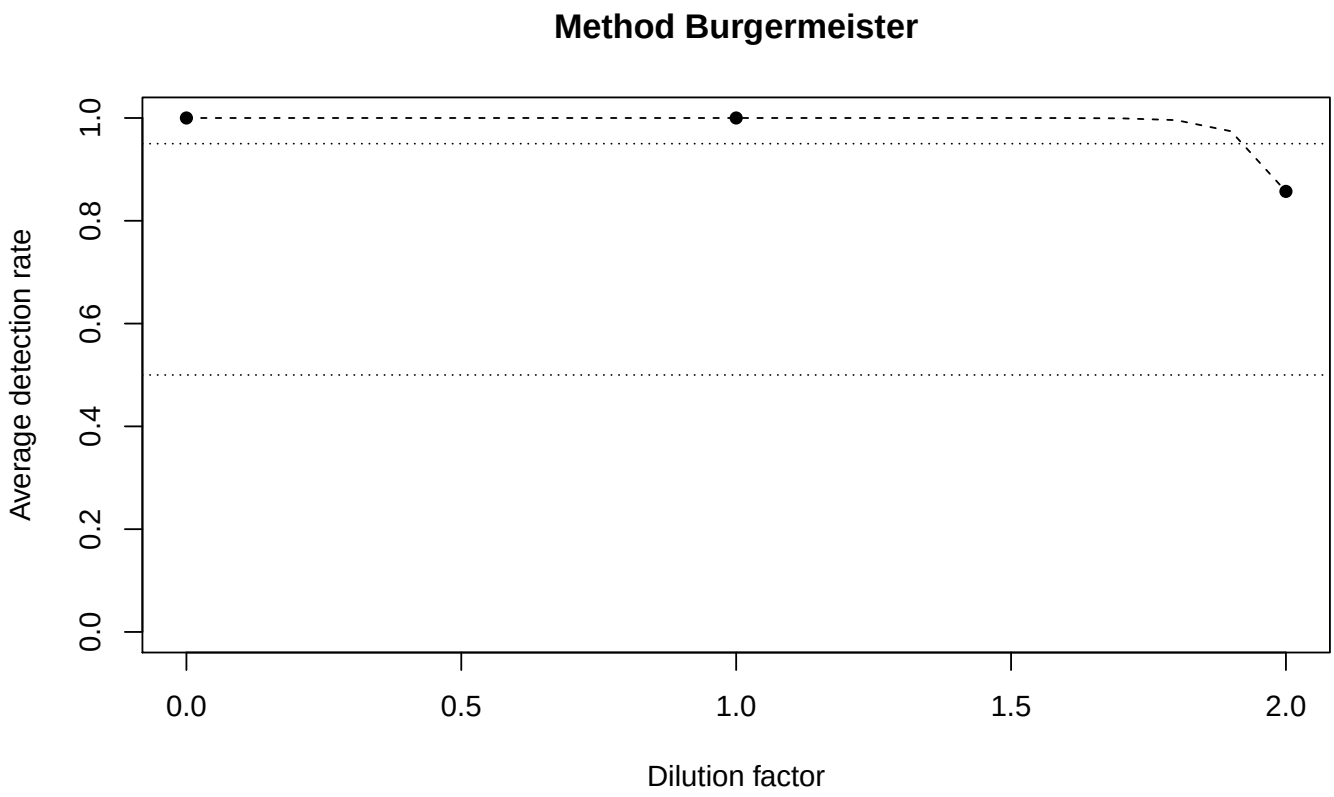
Repeatability and reproducibility

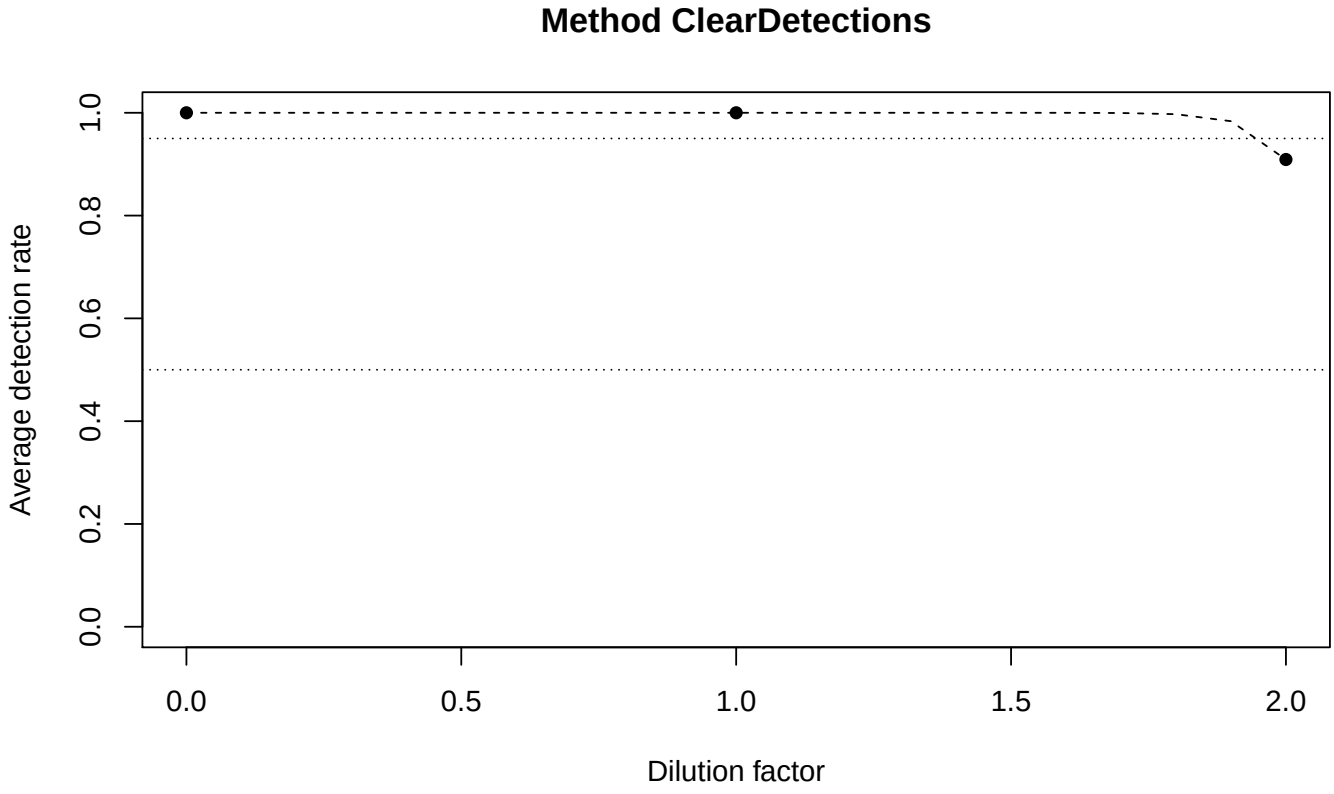
As none of the measure of this test was repeated, classical repetability and reproducibility of the methods are impossible to assess.

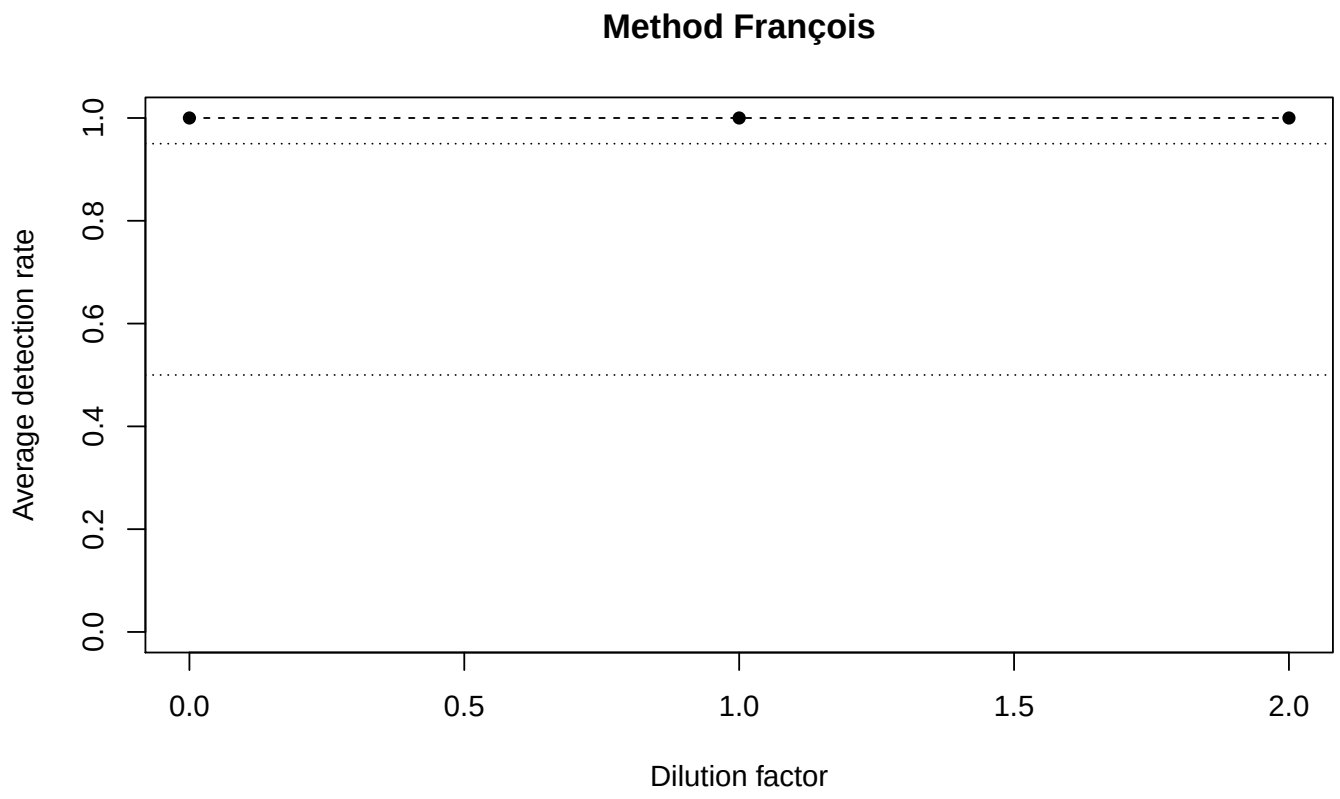
Analytical sensitivity

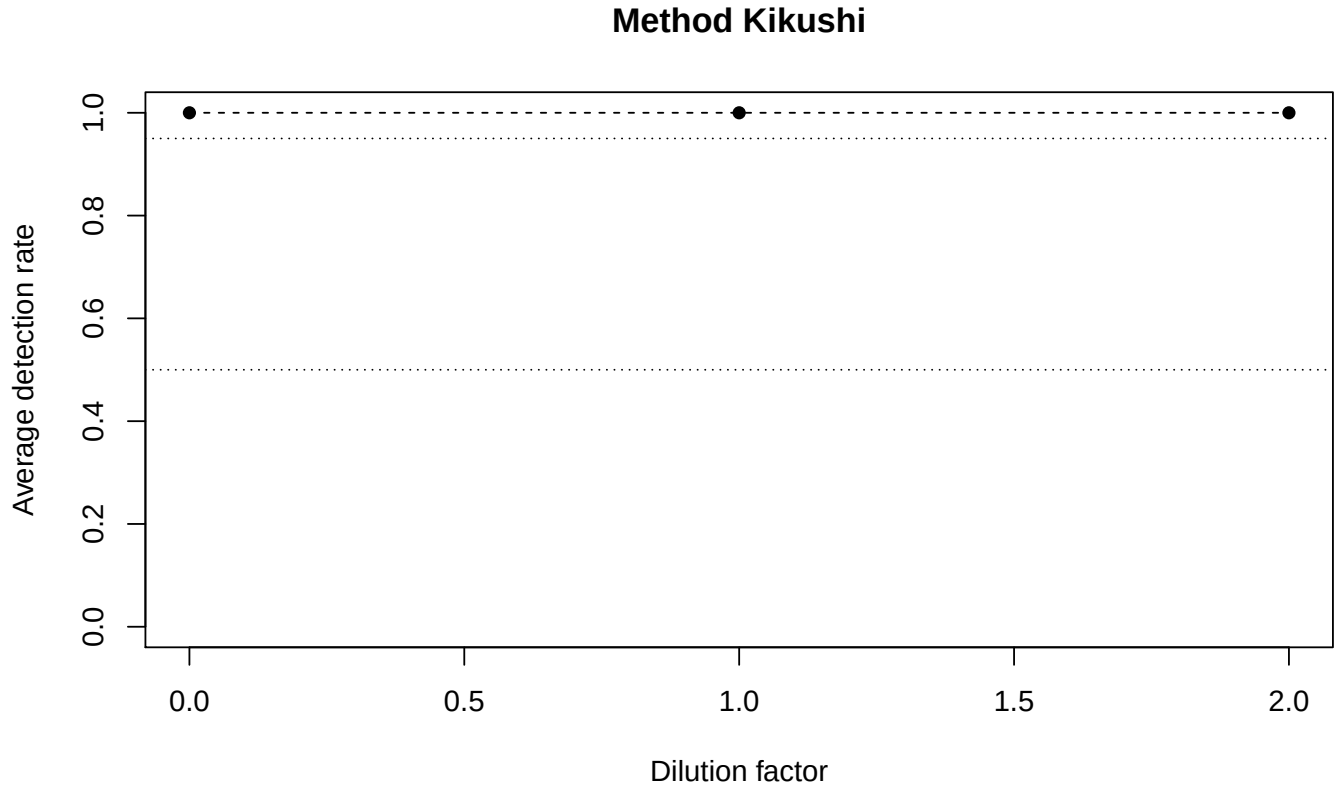
For each method , data of the diluted samples (7 to 9) were used to adjust binomial generalized linear models (bGLM) with logit link between the dilution (expressed by the base 10 negative exponent of the corresponding dilution) and the detection status. The number of dilution level being very limited, the adjustement of bGLM is not always possible as this method require at least 5 levels, and the laboratory effect has been neglected.

When possible, based on those models, dilutions corresponding to a 50% or 95% probability of detection have been calculated as an example of the possible LOD to report (LOD50 and LOD95).


Method Burgermeister
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Dilution factor


Method ClearDetections
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Dilution factor


Method François
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Dilution factor


Method Kikushi
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
Dilution factor

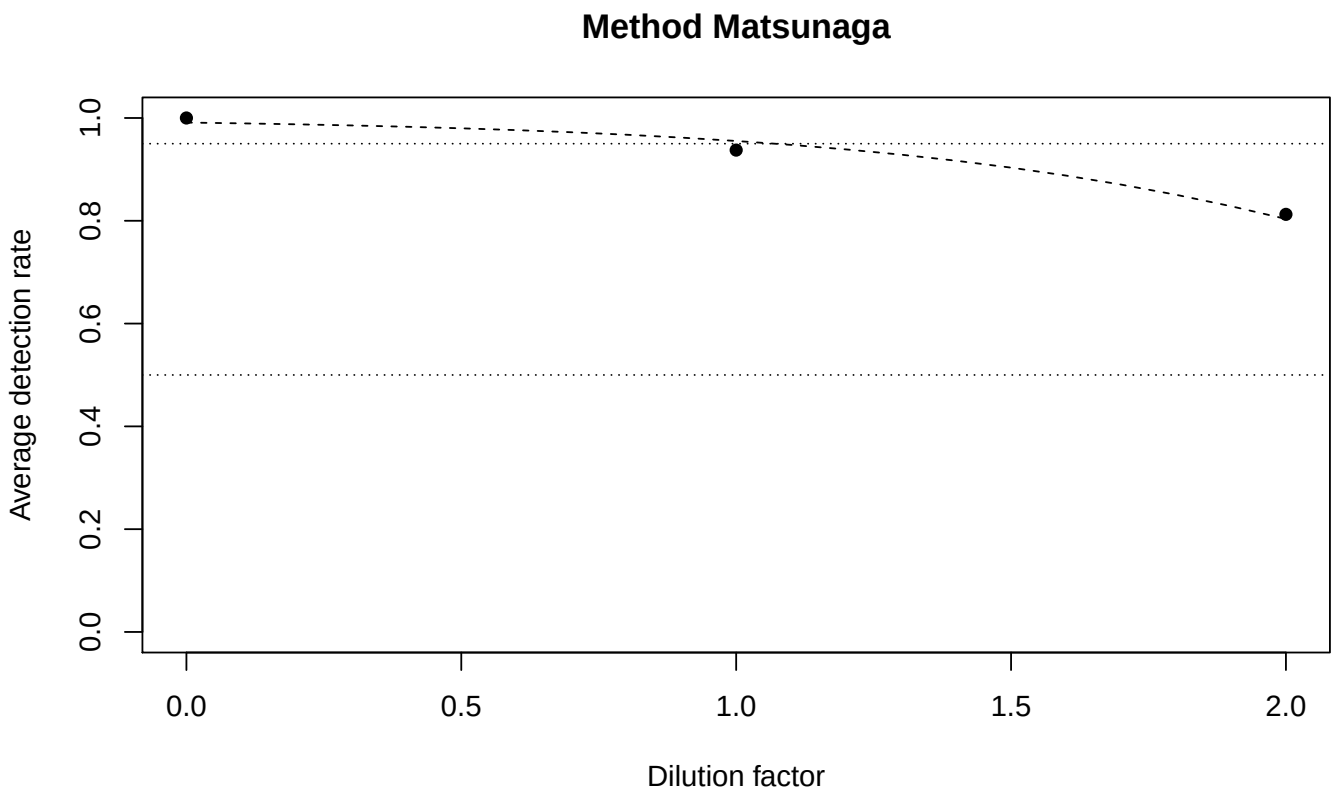


Table 1: Detection limits (log dilution factor) at 50% rate by methods

Burgermeister	ClearDetections	François	Kikushi	Matsunaga
NA	NA	NA	NA	NA

Table 2: Detection limits (log dilution factor) at 95% rate by methods

Burgermeister	ClearDetections	François	Kikushi	Matsunaga
1.9	2	NA	NA	1.1

Diagnostic sensitivity and specificity

Diagnostic sensitivity is estimated as the detection rate on samples with positive reference status, and the diagnostic specificity as the non detection rate on samples with negative reference status. Hence, those two parameters are heavily dependent on the choice of the reference positive and negative samples. Has the number of different samples vary from laboratories and methods, the comparison of the following estimates between methods has to be conducted with caution.

Table 3: Diagnostic sensitivity and specificity by methods

	Burgermeister	ClearDetections	François	Kikushi	Matsunaga
DSE	1.000	1.000	1.000	1.000	1.000
LCL	0.857	0.946	0.958	0.861	0.873
UCL	1.000	1.000	1.000	1.000	1.000
DSP	1.000	0.962	0.985	1.000	1.000

	Burgermeister	ClearDetections	François	Kikushi	Matsunaga
LCL	0.948	0.904	0.945	0.891	0.954
UCL	1.000	0.988	0.999	1.000	1.000

Likelihood ratios

Another drawback of a poor choice of reference samples is that 100% sensitivity or specificity estimates will impair the estimation of likelihood ratios, as they have either (1 - sensitivity) ou (1 - specificity) as denominator. Perfect specificity will then leads to infinite positive likelihood ratio, while perfect sensitivity leads to infinite negative likelihood ratio, obliterating the effect of the other parameter in the estimate. In this case, confidence limits can give some additional information by giving upper and lower bounds for this ratio, event in the case of an infinite estimation.

Table 4: Likelihood ratios by methods

	Burgermeister	ClearDetections	François	Kikushi	Matsunaga
LR+	Inf	26.500	68.000	Inf	Inf
LCL	22.867	10.749	19.218	10.892	25.991
UCL	Inf	67.701	247.391	Inf	Inf
LR-	Inf	Inf	Inf	Inf	Inf
LCL	8.289	21.501	28.173	8.549	9.330
UCL	Inf	Inf	Inf	Inf	Inf

Other diagnostic parameters

Table 5: Other diagnostic parameters by methods

	Burgermeister	ClearDetections	François	Kikushi	Matsunaga
Accuracy	1.000	0.979	0.992	1.000	1.000
LCL	0.960	0.945	0.968	0.935	0.965
UCL	1.000	0.994	1.000	1.000	1.000
Power	1.000	0.953	0.981	1.000	1.000
LCL	0.857	0.883	0.931	0.861	0.873
UCL	1.000	0.985	0.999	1.000	1.000
Rate True Positive	1.000	0.953	0.981	1.000	1.000
LCL	0.857	0.883	0.931	0.861	0.873
UCL	1.000	0.985	0.999	1.000	1.000
Rate True Negative	1.000	1.000	1.000	1.000	1.000
LCL	0.948	0.956	0.966	0.891	0.954
UCL	1.000	1.000	1.000	1.000	1.000

Outliers detection

Outliers can be detected by decomposing the calculation of previous parameters to the laboratory and/or the sample level, and looking for strong individual deviation among them. Those deviations can be investigated by expert, leading to the possible exclusion of the corresponding results from the general analysis if necessary.

Repetability & reproducibility

Impossible due to the absence of replicates

Analytical sensitivity

Impossible due to limited number of dilution levels

Diagnostic sensitivity and diagnostic specificity

Comparison impossible due to different number of samples between laboratories and methods

VALITEST TPS report

Fusarium data

Yves Brostaux

26 juin 2020

Files reading and data preprocessing

Fusarium data are stored in one csv files, with the following columns :

- sampid, sample ID
- method, method ID
- lab, lab ID
- test, result of the test (binary 0/1)
- ref, reference status of each sample (binary 0/1)
- dilution (dilution factor for diluted samples)
- link, ID of the original sample of the dilution serie

Following corrections were applied to the data prior to analysis : - ID of samples 2, 3, 4 and 5 have been converted to 15, 14, 13 and 12 - linked samples of samples 7, 10 & 11 have been set fo 12 (dilution serie)

Those files have been read and combined into different data tables for the analysis.

- tidydata, 1 line per status result, same organization as the file
- dilu.dat, subset of tidydata including only diluted samples, with additional column
  - dilu, dilution exponent (negative base 10)
  - reported results not equal to 0 ou 1 have been deleted
- spse.dat, subset of tidydata including only non diluted samples

```
## tidydata
```

```
##   sampid  method lab test ref dilution link
## 1      1 Plating   A    0   1            NA
## 2     15 Plating   A    0   0            NA
## 3     14 Plating   A    1   1            NA
## 4     13 Plating   A    1   1            NA
## 5     12 Plating   A    0   1            NA
## 6      6 Plating   A    0   0            NA
```

```
## dilu.dat
```

```
##    sampid method lab test ref dilution link dilu
## 7       7   PCR1   A    1   1    1:100   12    2
## 10     10   PCR1   A    0   1   1:1000   12    3
## 11     11   PCR1   A    1   1     1:10   12    1
## 12     12   PCR1   A    1   1            NA    0
## 17      7   PCR2   A    0   1    1:100   12    2
## 20     10   PCR2   A    0   1   1:1000   12    3
```

```
## spse.dat
```

```
##   sampid  method lab test ref dilution link
## 1      1 Plating   A    0   1            NA
## 2     15 Plating   A    0   0            NA
```

```
## 3     14 Plating   A    1   1           NA
## 4     13 Plating   A    1   1           NA
## 5     12 Plating   A    0   1           NA
## 6      6 Plating   A    0   0           NA
```

Note that the number of samples tested vary between methods and laboratories. This will impair the comparison of methods based on calculated parameters, as their basis differ.

```
##    method
## lab PCR1 PCR2 PCR3 PCR4 PCR5 Plating
##   A   10   10   10   10   10       6
##   B    0    5    5    5    5       0
##   C    0    0    0    0    0       6
##   D   10    0    0    0    0       6
##   E   10    0    0    0    0       6
##   F   10   10   10   10   10       0
##   G    0   10   10   10   10       6
##   H    1    0    0    0    0       6
##   I   10    0    0    0    0       6
##   J   10   10   10   10   10       6
##   K    0    0    0    0    0       6
##   L   10   10   10   10   10       0
##   M   10   10   10   10    0       6
##   N    5    5    5    5    5       0
##   O   10   10   10   10   10       6
##   P    5    5    5    5    5       0
##   Q    0   10   10   10   10       6
##   R    5    5    5    5    5       0
##   S   10   10   10   10   10       6
##   T   10   10   10   10   10       6
```

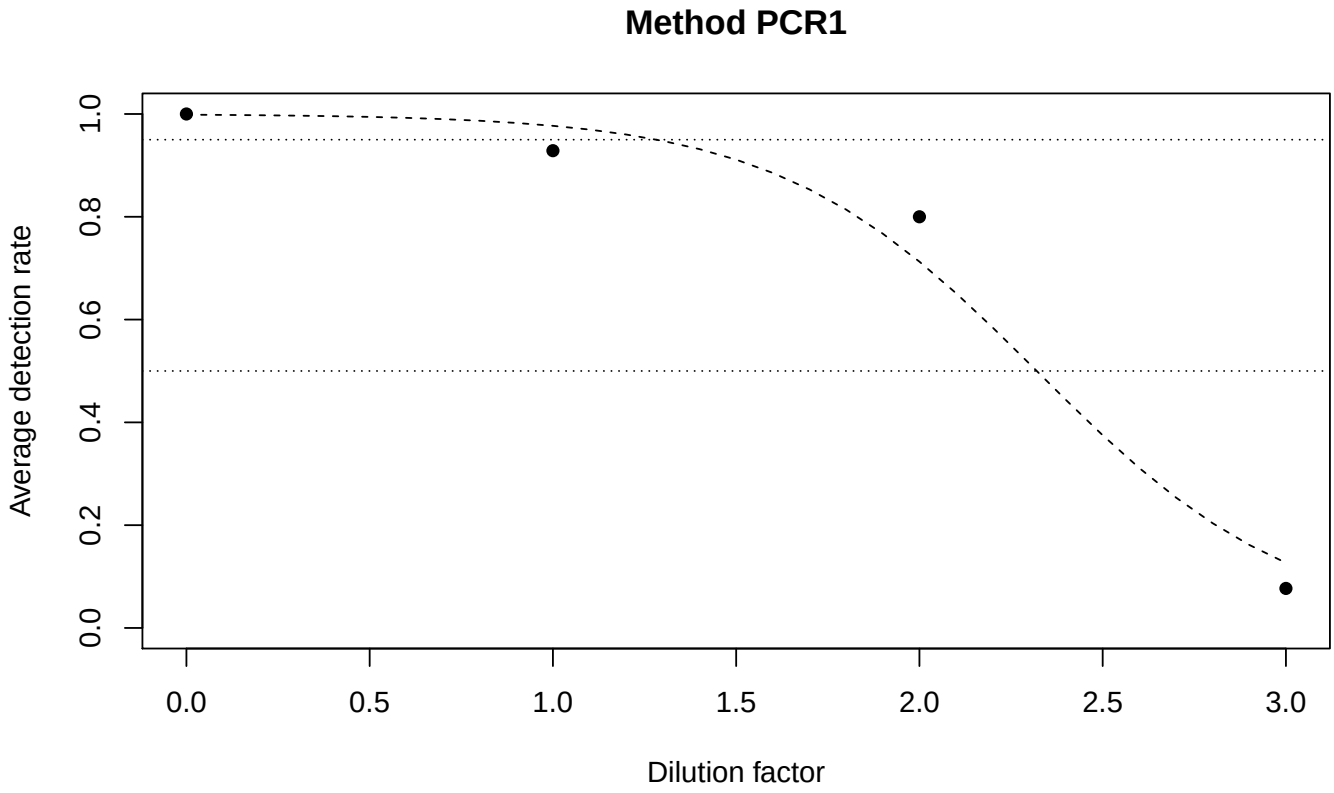
Repeatability and reproducibility

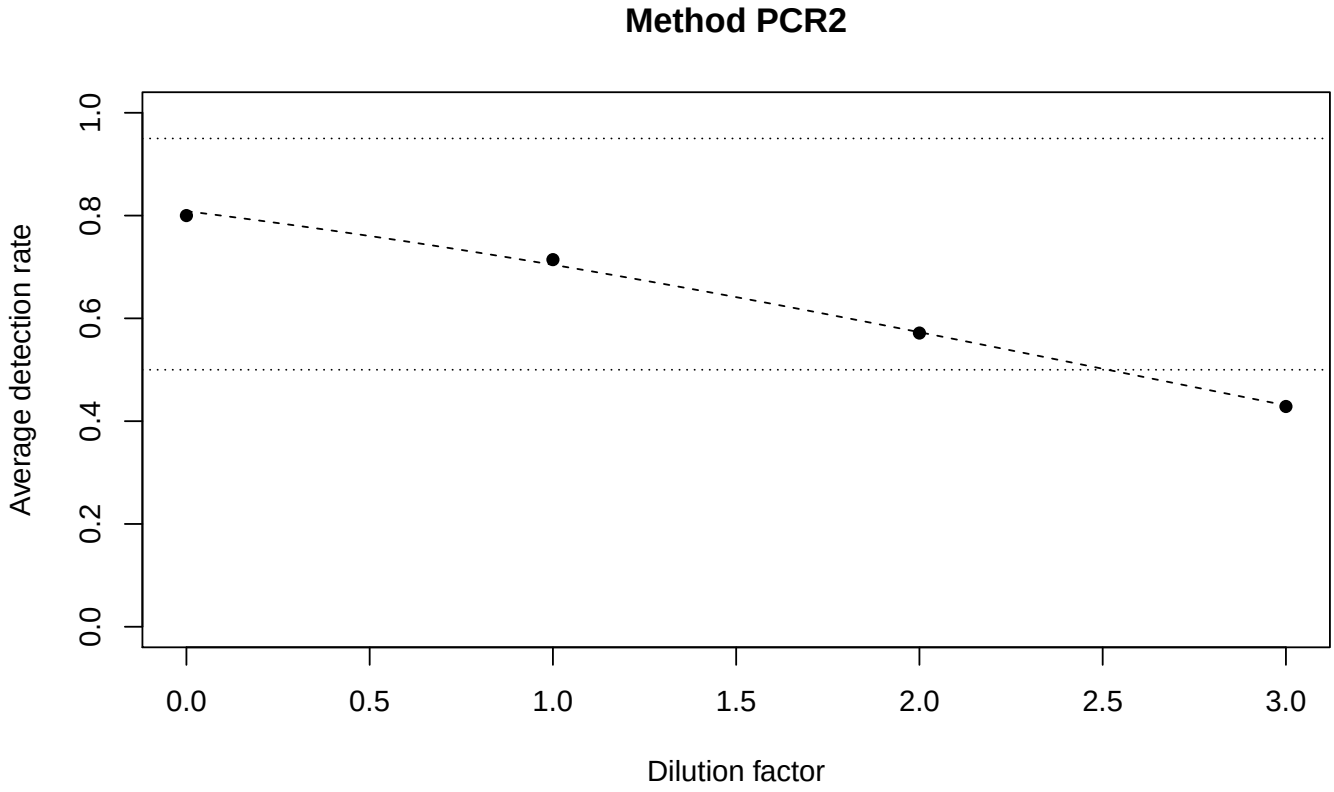
As none of the measure of this test was repeated, classical repetability and reproducibility of the methods are impossible to assess.

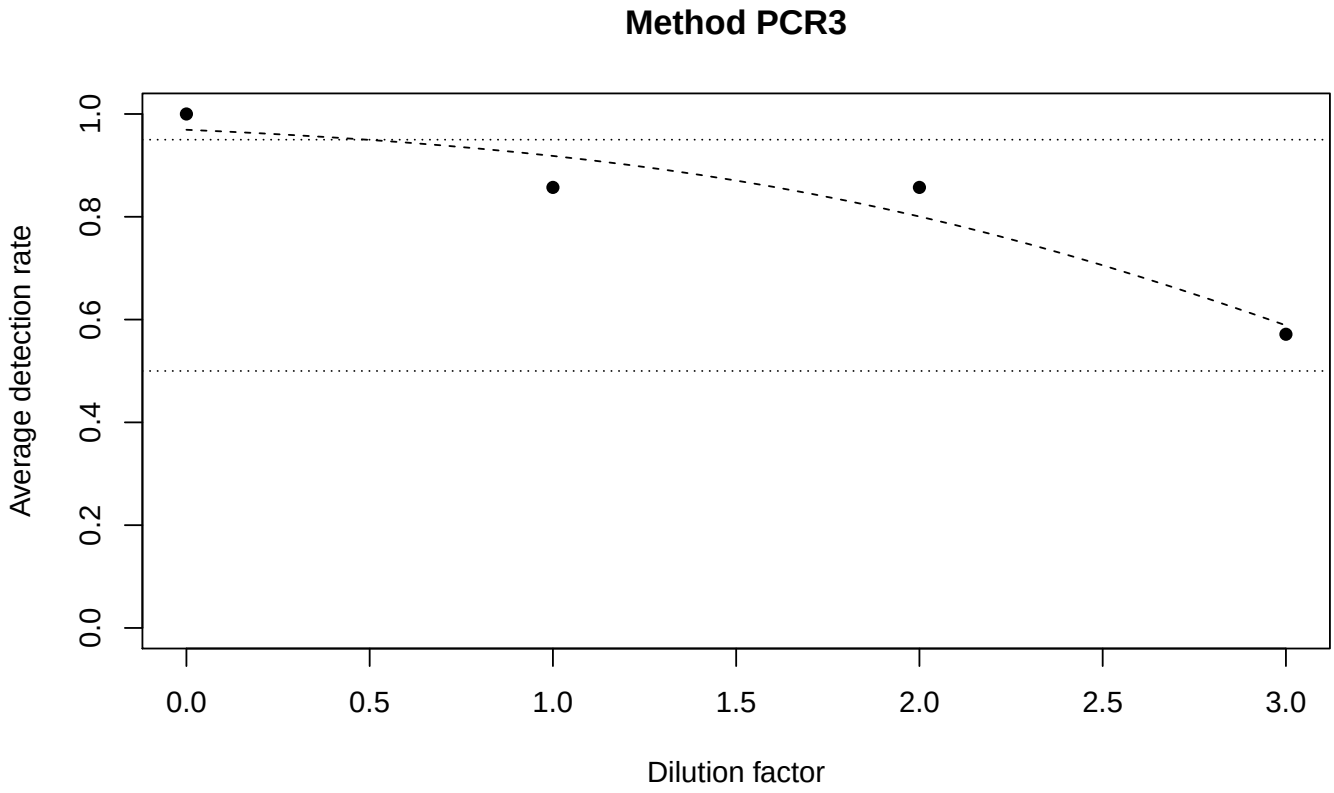
Analytical sensitivity

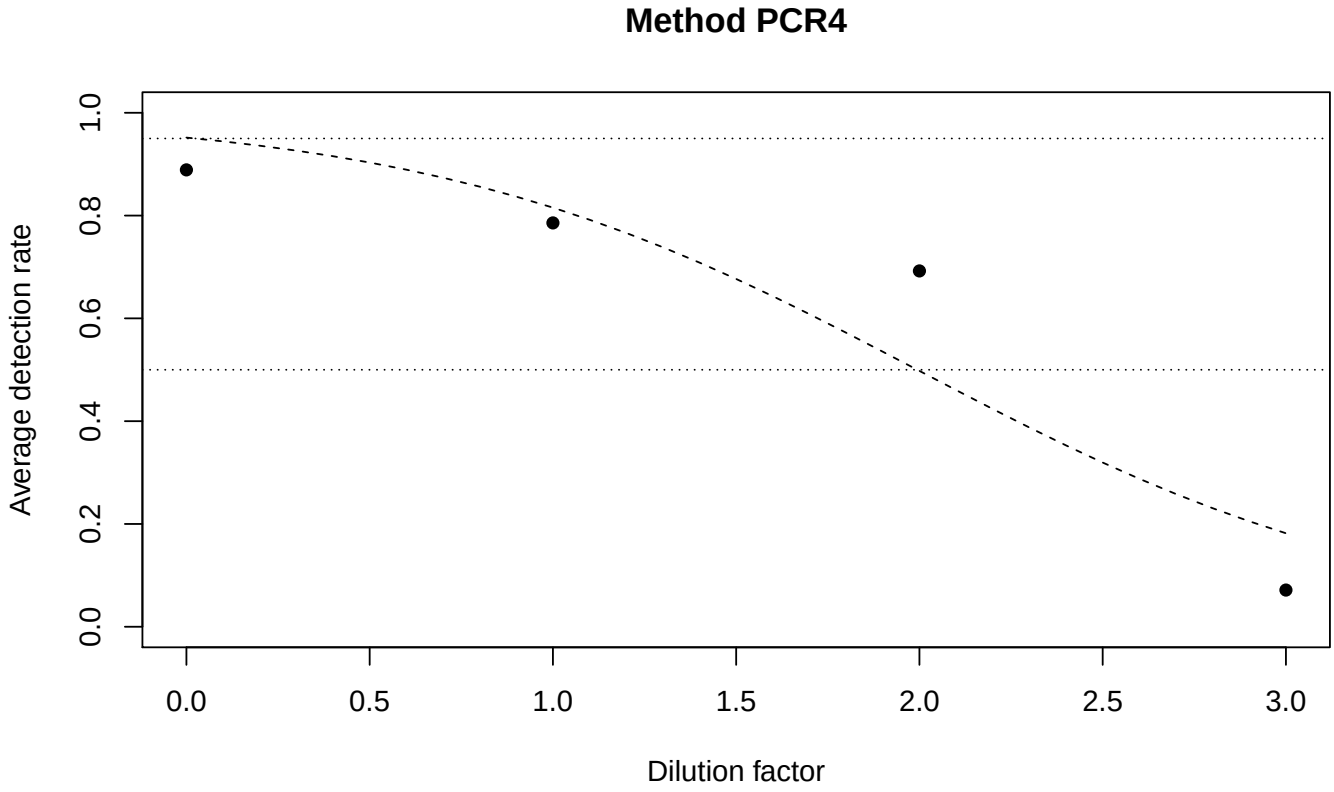
For each method , data of the diluted samples (7, 10, 11 and 12) were used to adjust binomial generalized linear models (bGLM) with logit link between the dilution (expressed by the base 10 negative exponent of the corresponding dilution) and the detection status. The number of dilution level being very limited, the adjustement of bGLM is not always possible as this method require at least 5 levels, and the laboratory effect has been neglected.

When possible, based on those models, dilutions corresponding to a 50% or 95% probability of detection have been calculated as an example of the possible LOD to report (LOD50 and LOD95).


Method PCR1
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dilution factor


Method PCR2
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dilution factor


Method PCR3
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dilution factor


Method PCR4
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Dilution factor

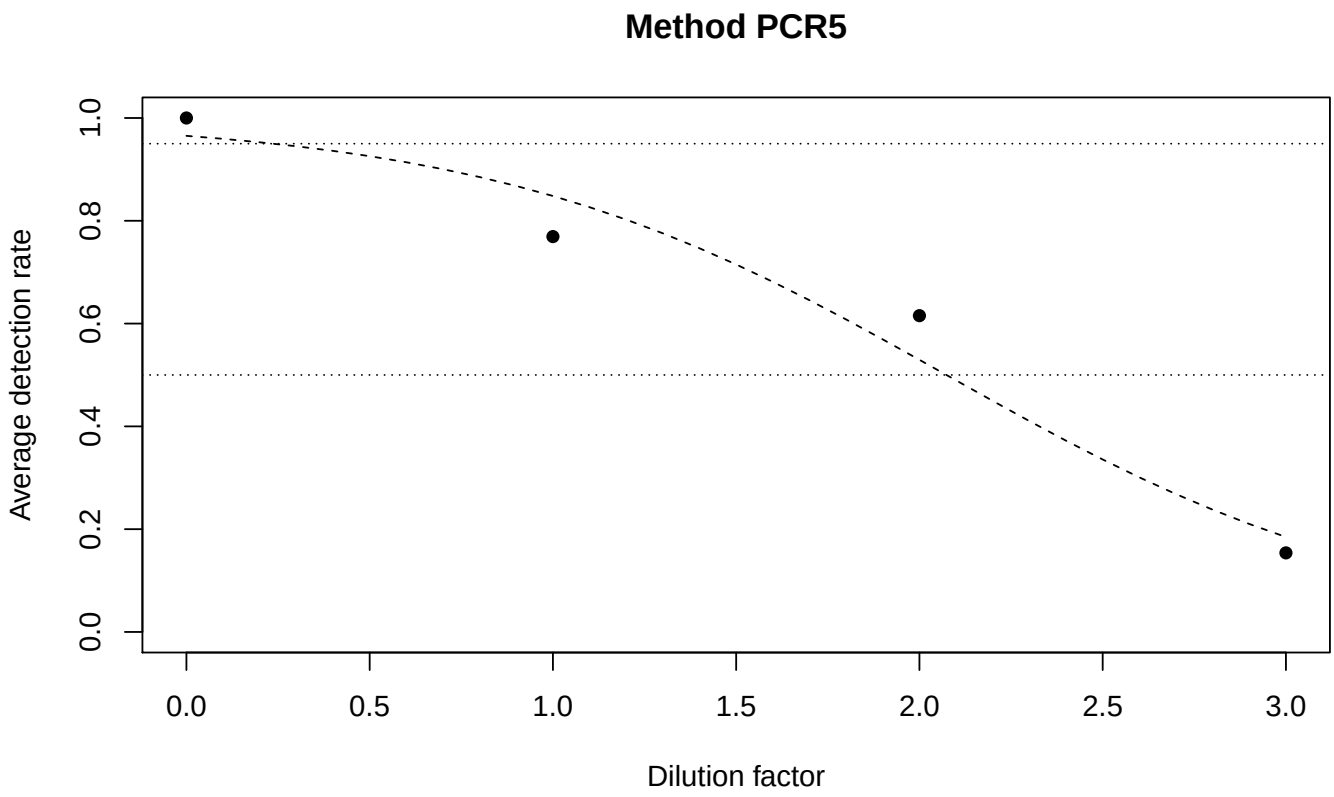


Table 1: Detection limits (log dilution factor) at 50% rate by methods

PCR1	PCR2	PCR3	PCR4	PCR5
2.3	2.5	NA	2	2.1

Table 2: Detection limits (log dilution factor) at 95% rate by methods

PCR1	PCR2	PCR3	PCR4	PCR5
1.3	NA	0.5	0	0.2

Diagnostic sensitivity and specificity

Diagnostic sensitivity is estimated as the detection rate on samples with positive reference status, and the diagnostic specificity as the non detection rate on samples with negative reference status. Hence, those two parameters are heavily dependent on the choice of the reference positive and negative samples. Has the number of different samples vary from laboratories and methods, the comparison of the following estimates between methods has to be conducted with caution.

Table 3: Diagnostic sensitivity and specificity by methods

	PCR1	PCR2	PCR3	PCR4	PCR5	Plating
DSE	1.000	0.667	1.000	0.867	0.963	0.679
LCL	0.876	0.487	0.865	0.697	0.802	0.548
UCL	1.000	0.809	1.000	0.953	1.000	0.787
DSP	0.972	0.971	0.971	0.941	0.903	0.964

	PCR1	PCR2	PCR3	PCR4	PCR5	Plating
LCL	0.846	0.838	0.838	0.799	0.743	0.808
UCL	1.000	1.000	1.000	0.993	0.974	1.000

Likelihood ratios

Another drawback of a poor choice of reference samples is that 100% sensitivity or specificity estimates will impair the estimation of likelihood ratios, as they have either (1 - sensitivity) ou (1 - specificity) as denominator. Perfect specificity will then leads to infinite positive likelihood ratio, while perfect sensitivity leads to infinite negative likelihood ratio, obliterating the effect of the other parameter in the estimate. In this case, confidence limits can give some additional information by giving upper and lower bounds for this ratio, event in the case of an infinite estimation.

Table 4: Likelihood ratios by methods

	PCR1	PCR2	PCR3	PCR4	PCR5	Plating
LR+	36.000	22.667	34.000	14.733	9.951	19.000
LCL	7.057	4.370	6.704	4.490	3.851	3.786
UCL	203.235	129.393	191.905	53.525	28.833	107.784
LR-	Inf	2.912	Inf	7.059	24.387	3.000
LCL	9.321	1.882	8.547	3.151	4.914	2.103
UCL	Inf	5.060	Inf	17.780	137.841	4.526

Other diagnostic parameters

Table 5: Other diagnostic parameters by methods

	PCR1	PCR2	PCR3	PCR4	PCR5	Plating
Accuracy	0.986	0.828	0.984	0.906	0.931	0.774
LCL	0.915	0.716	0.909	0.807	0.831	0.673
UCL	1.000	0.903	1.000	0.960	0.978	0.851
Power	0.971	0.488	0.968	0.722	0.839	0.507
LCL	0.838	0.343	0.824	0.559	0.669	0.396
UCL	1.000	0.635	1.000	0.843	0.934	0.617
Rate True Positive	0.971	0.952	0.968	0.929	0.897	0.974
LCL	0.838	0.756	0.824	0.763	0.728	0.856
UCL	1.000	1.000	1.000	0.991	0.972	1.000
Rate True Negative	1.000	0.767	1.000	0.889	0.966	0.600
LCL	0.882	0.621	0.876	0.741	0.814	0.454
UCL	1.000	0.870	1.000	0.962	1.000	0.730

Outliers detection

Outliers can be detected by decomposing the calculation of previous parameters to the laboratory and/or the sample level, and looking for strong individual deviation among them. Those deviations can be investigated by expert, leading to the possible exclusion of the corresponding results from the general analysis if necessary.

Repetability & reproducibility

Impossible due to the absence of replicates

Analytical sensitivity

Impossible due to limited number of dilution levels

Diagnostic sensitivity and diagnostic specificity

Comparison impossible due to different number of samples between laboratories and methods

VALITEST TPS2 report

Cp data

Yves Brostaux

23 octobre 2020

Files reading and data preprocessing

Cp data are stored in one csv files, with the following columns :

- Sample, Method, Laboratory Replicat
- Results and Status (binary)
- Concentration/quantity/dilution (renamed dilution)
- Linked sample, Reference method and Sample info

Dilution series have been wrongly reported. Column *linked samples* should contain the sample ID of the original sample of the dilution serie, and dilution information is not present for QPCR1 and QPCR2.

The following hypothesis has been infered from data interpretation and the data have been corrected/completed accordingly : 3 dilution series (CP10, CP23 and CP24) with 3 dilution levels (1, 0.25 and 0.05), associated respectively with samples 1 to 3 (1), 3 to 5 (0.05) and 12 to 14 (0.25).

Results columns contains some missing data. Without any further information, those values have been deleted from the analysis.

After correction , the date were combined into different data tables for the analysis.

- tidydata, 1 line per status result,
  - sampid, sample ID
  - method, method ID
  - lab, lab ID
  - test, result of the test (binary 0/1)
  - ref, reference status of each sample (binary 0/1)
  - dilution (dilution factor for diluted samples)
  - link, ID of the original sample of the dilution serie
- dilu.dat, subset of tidydata including only diluted samples, with additional column dilu (target concentration (log base))
- spse.dat, subset of tidydata including only non diluted samples

```
## tidydata
```

```
##   sampid method lab replic test ref dilution link refmeth sinfo dilu
## 1    NAC   PCR1   A      1    0   0       NA          Yes   DNA   NA
## 2    NAC   PCR1   A      2    0   0       NA          Yes   DNA   NA
## 3    NIC   PCR1   A      1    0   0       NA          Yes  wood   NA
## 4    NIC   PCR1   A      2    0   0       NA          Yes  wood   NA
## 5    PAC   PCR1   A      1    1   1       NA          Yes   DNA   NA
## 6    PAC   PCR1   A      2    1   1       NA          Yes   DNA   NA
```

```
## dilu.dat
```

```
##    sampid method lab replic test ref dilution                        link refmeth sinfo    dilu
## 9       4   PCR1   A      1    1   1     0.05 Cryphonectria parasitica CP10     Yes   DNA 1.30103
## 10      4   PCR1   A      2    1   1     0.05 Cryphonectria parasitica CP10     Yes   DNA 1.30103
## 11     12   PCR1   A      1    0   1     0.25 Cryphonectria parasitica CP10     Yes  wood 0.60206
```

```
## 12     12   PCR1   A      2    0   1     0.25 Cryphonectria parasitica CP10     Yes  wood 0.60206
## 13      1   PCR1   A      1    1   1     1.00 Cryphonectria parasitica CP10     Yes   DNA 0.00000
## 14      1   PCR1   A      2    1   1     1.00 Cryphonectria parasitica CP10     Yes   DNA 0.00000
```

```
## spse.dat
```

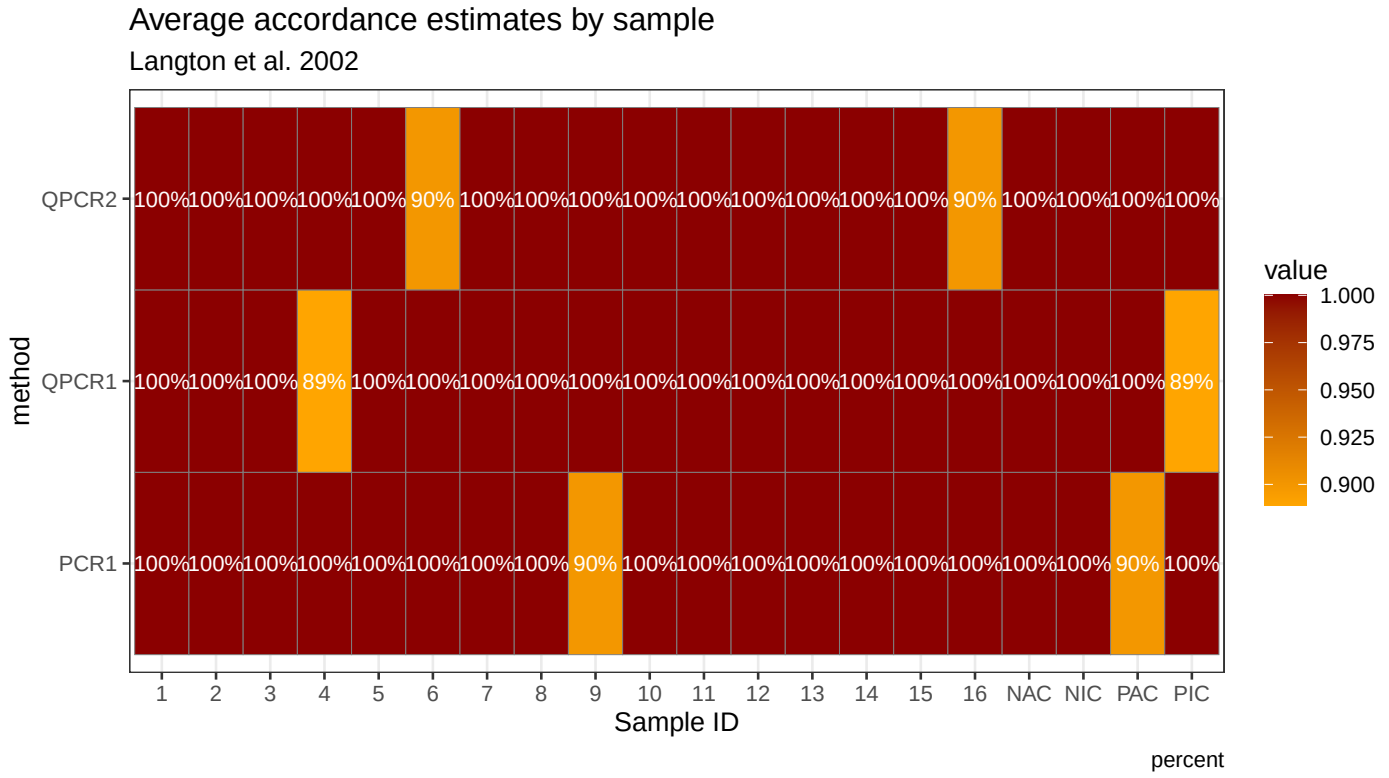
```
##   sampid method lab replic test ref dilution link refmeth sinfo dilu
## 1    NAC   PCR1   A      1    0   0       NA          Yes   DNA   NA
## 2    NAC   PCR1   A      2    0   0       NA          Yes   DNA   NA
## 3    NIC   PCR1   A      1    0   0       NA          Yes  wood   NA
## 4    NIC   PCR1   A      2    0   0       NA          Yes  wood   NA
## 5    PAC   PCR1   A      1    1   1       NA          Yes   DNA   NA
## 6    PAC   PCR1   A      2    1   1       NA          Yes   DNA   NA
```

```
##    method
## lab PCR1 QPCR1 QPCR2
##   A   40    40    20
##   B   40    40    20
##   C   40    40    20
##   D   40    40    20
##   E   40    40    20
##   F   20    20    20
##   G   40    40    20
##   H   40    40    20
##   I   40    40    20
##   J   38    38    20
```

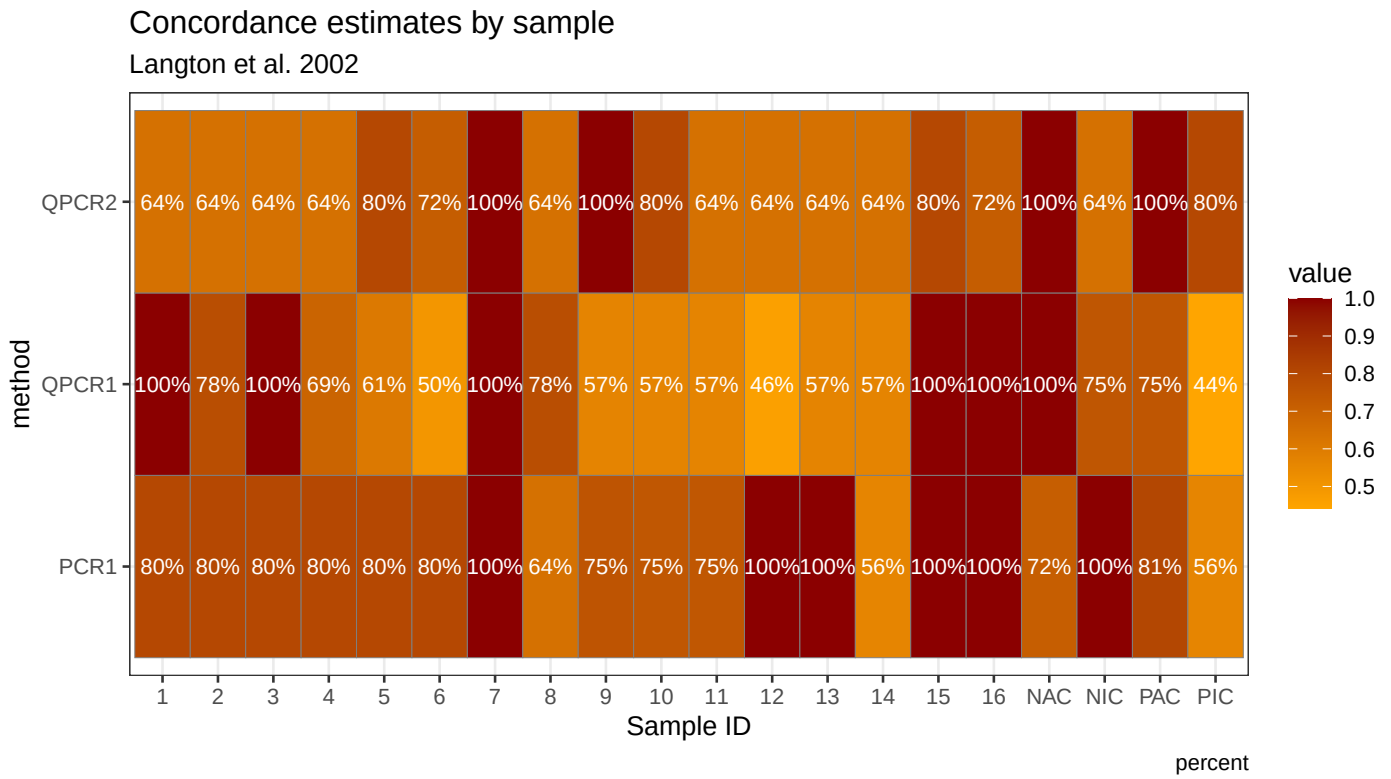
Repeatability and reproducibility

Accordance and concordance coefficients are used to estimate repeatability and reproducibility through laboratories, methods and samples. Unconclusive results are excluded from this analysis.

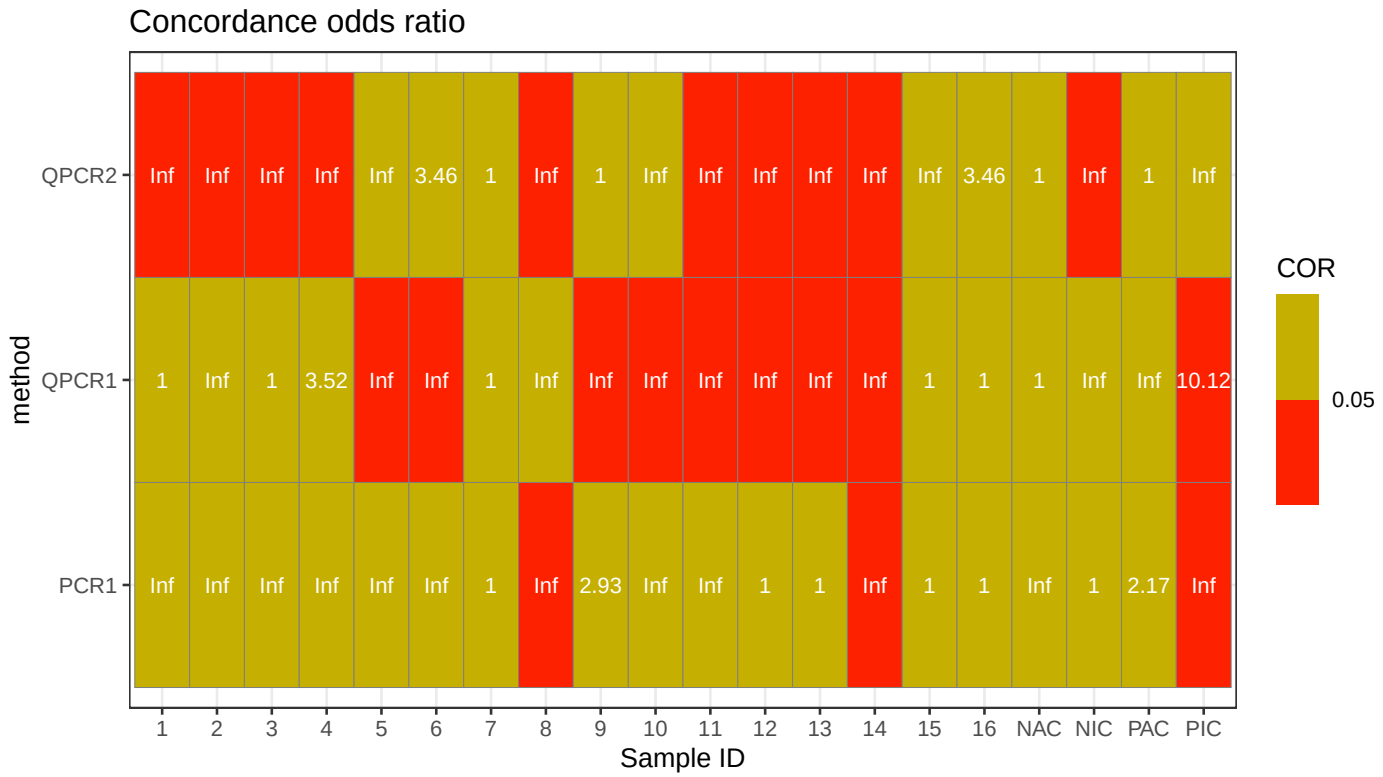
Accordance estimated values averaged by samples can give some insight about the high dependency of those value to the chosen sample.



Concordance estimates are equalily useful in this special case, because it highlights the samples which leads to high discrepancies between laboratories.



Concordance odds ratio (COR) can also assess the degree of interlaboratory variation ratio. COR removes the bias related to the accuracy of the results (i.e. numbers of true positive/negative and of false positive/negative) which are used to calculate the two parameters (i.e. concordance and accordance) taken separately. But because of the numerous 100% values, concordance odds ratios are of little help to discriminate all but the least efficient methods, as most of the estimates are either 1 or infinite values. This can be completed at the sample level by a Fisher's test, which test the hypothesis that there is a significant variation of the results between laboratories for a particular sample (the COR is significantly greater than one).



We also can characterize globally the repetability and reproducibility by calculating those parameters by method.

```
kable(RRtab(acc.tab, conc.tab), digits=3, caption="Repeatabilty & Reproducibility by Method")
```

Table 1: Repeatabilty & Reproducibility by Method

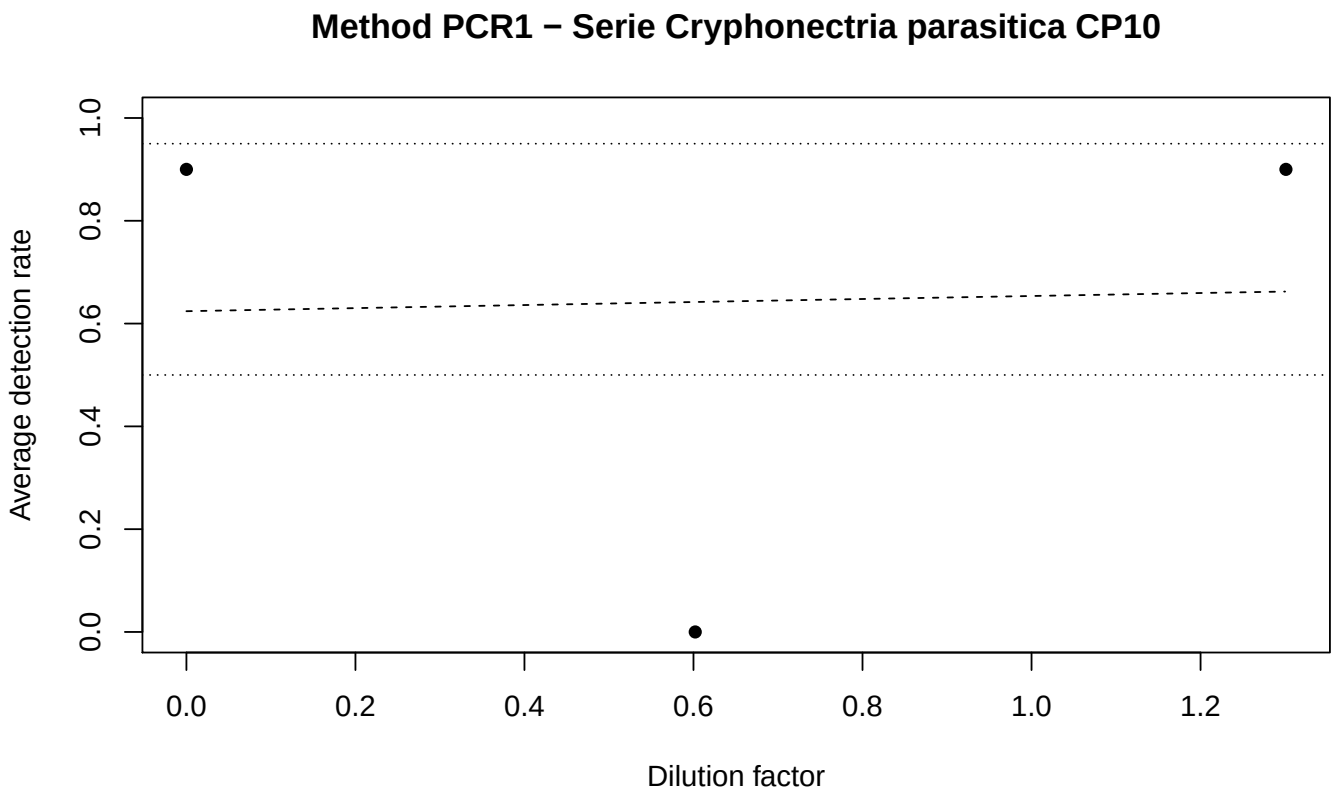
	PCR1	QPCR1	QPCR2
acc	0.990	0.989	0.990
conc	0.817	0.730	0.754
cor	22.180	32.957	32.222

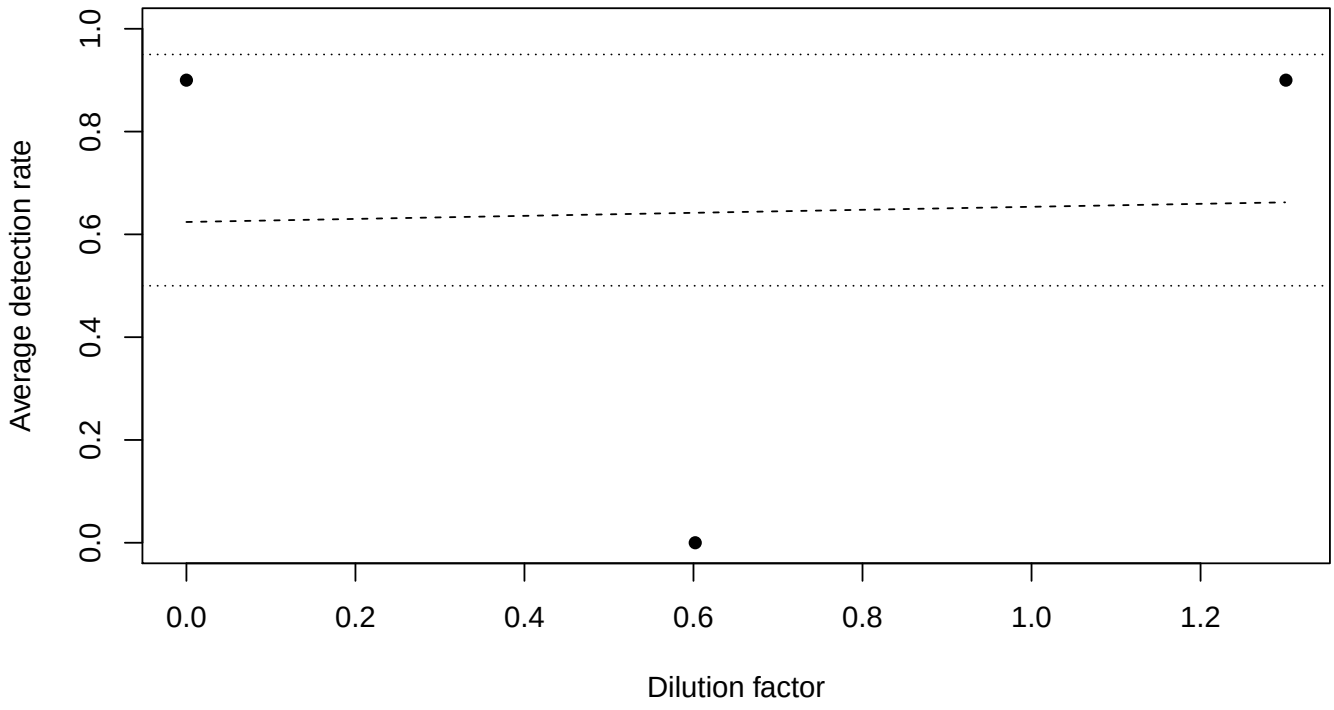
Analytical sensitivity

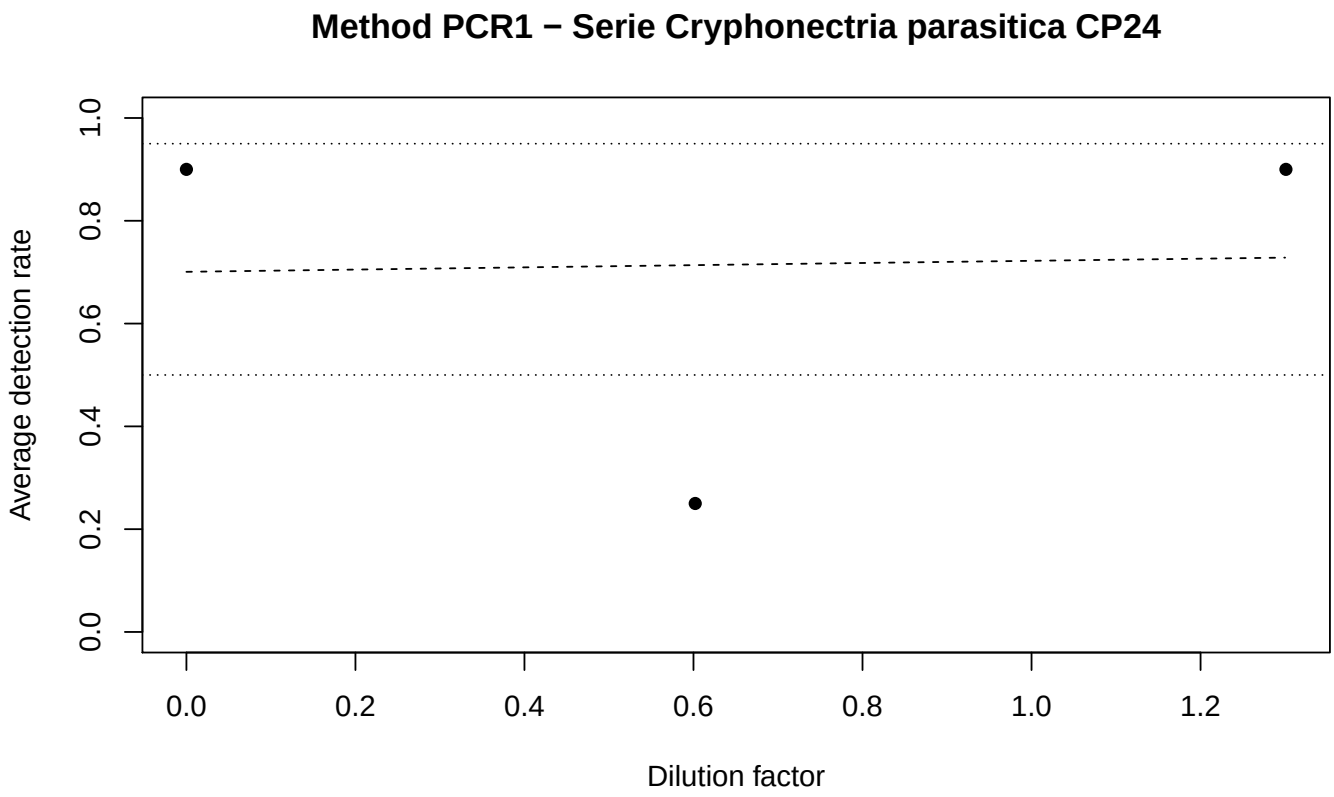
For each method , data of the diluted samples were used to adjust binomial generalized linear models (bGLM) with logit link between the dilution (expressed by the base 10 negative exponent of the corresponding dilution) and the detection status. The number of dilution level being very limited, the adjustement of bGLM is not always possible as this method require at least 5 levels, and the laboratory effect has been neglected.

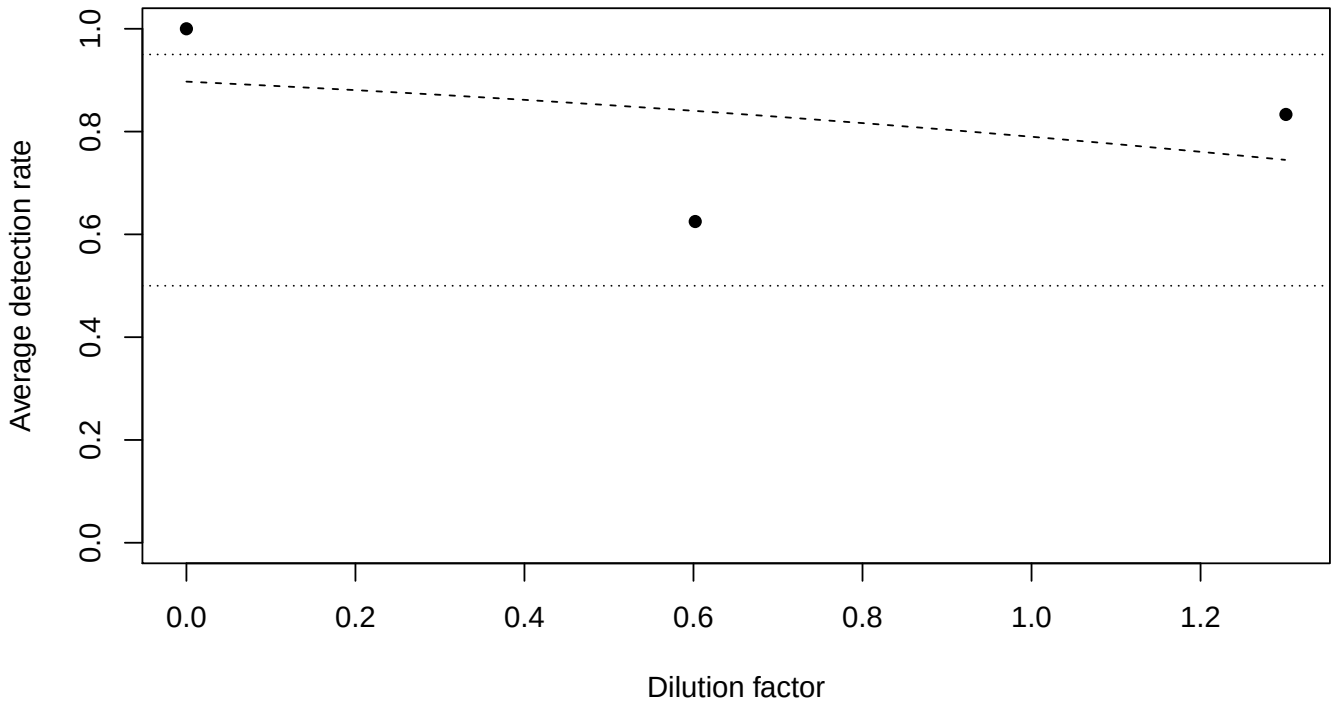
When possible, based on those models, dilutions corresponding to a 50% or 95% probability of detection have been calculated as an example of the possible LOD to report (LOD50 and LOD95).

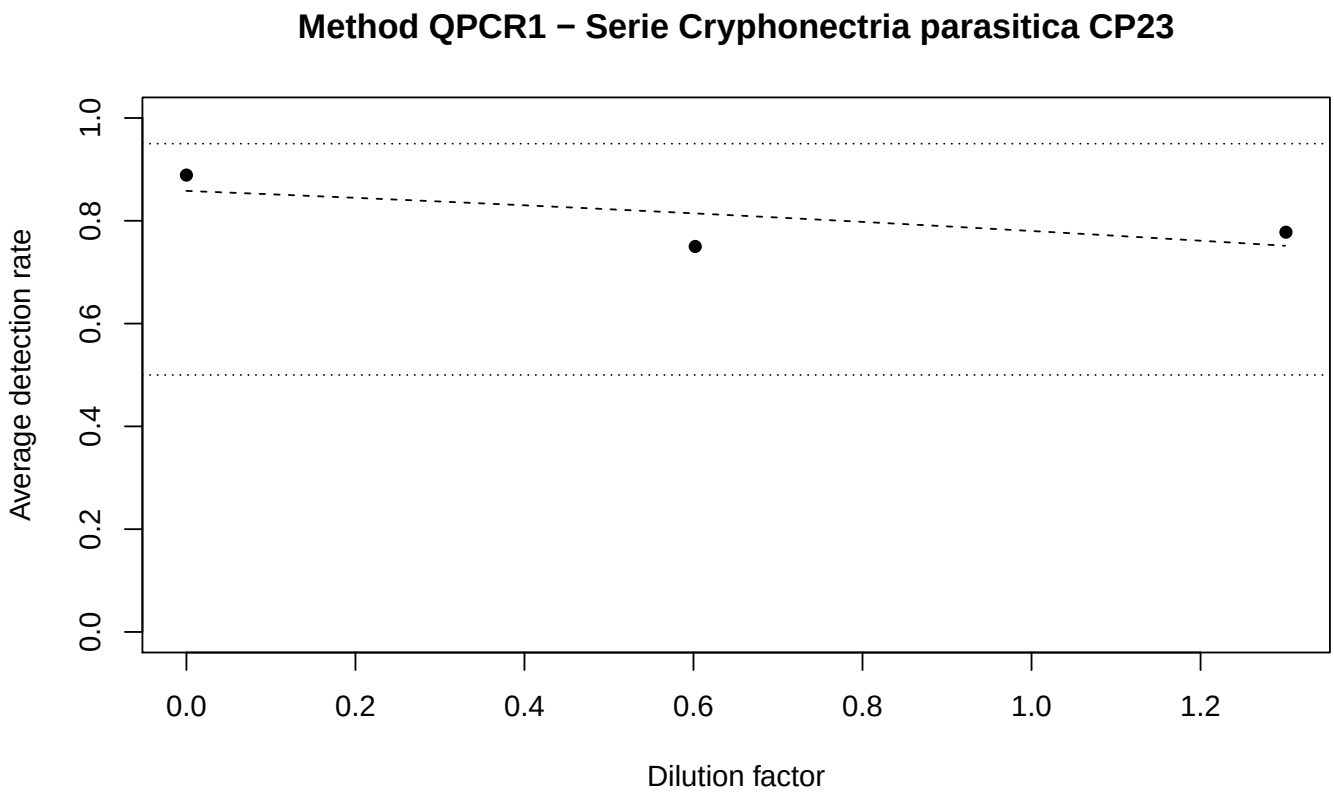
Results show that samples with dilution 1:20/0.05 perform on average better than samples diluted at 1:4/0.25, which is not the expected behavior. It may come from a wrong interpretation of the dilution information and need to be checked.

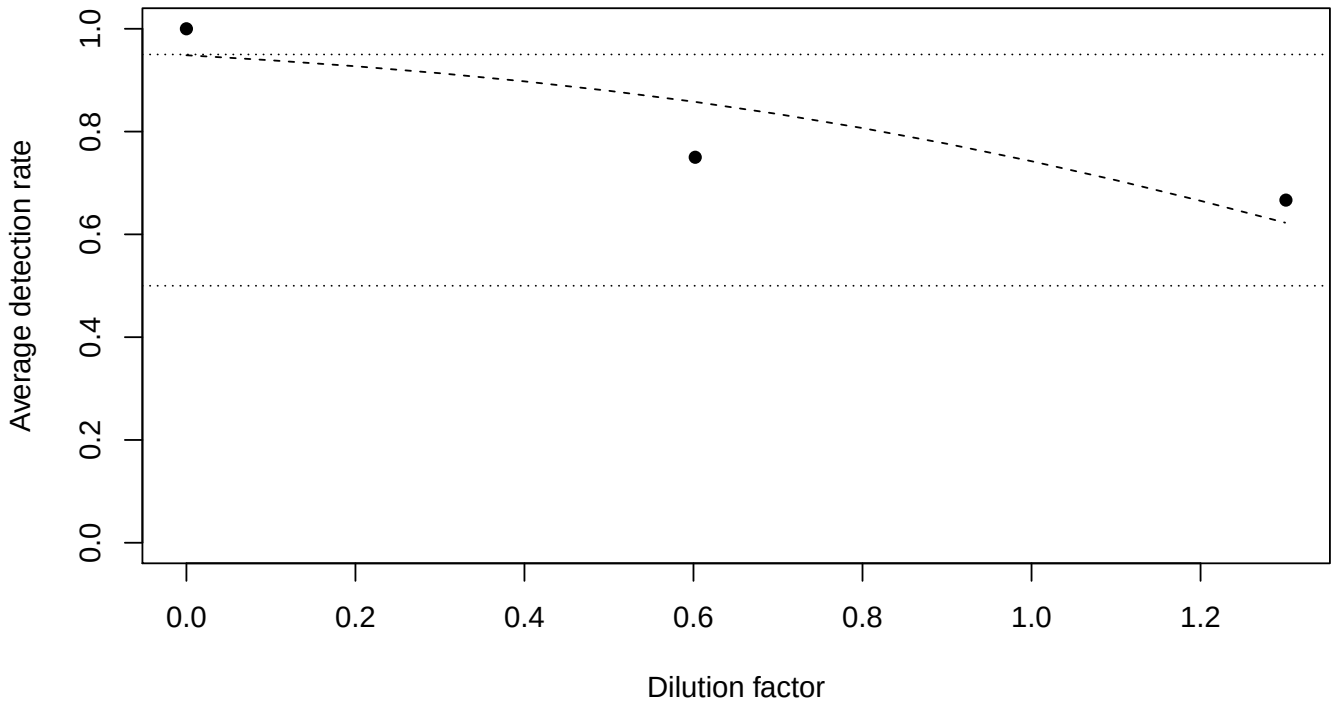

Method PCR1 − Serie Cryphonectria parasitica CP10
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Dilution factor

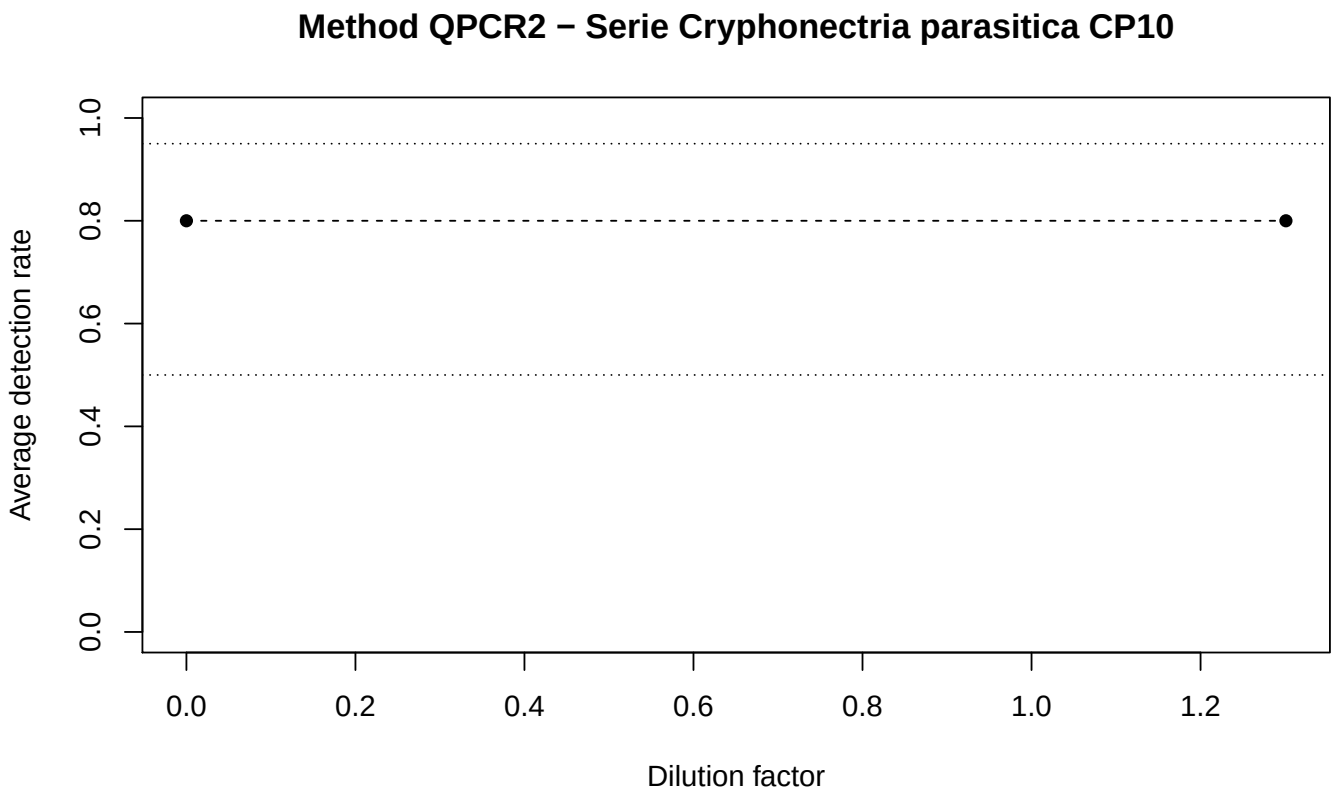

Method PCR1 − Serie Cryphonectria parasitica CP23
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Dilution factor

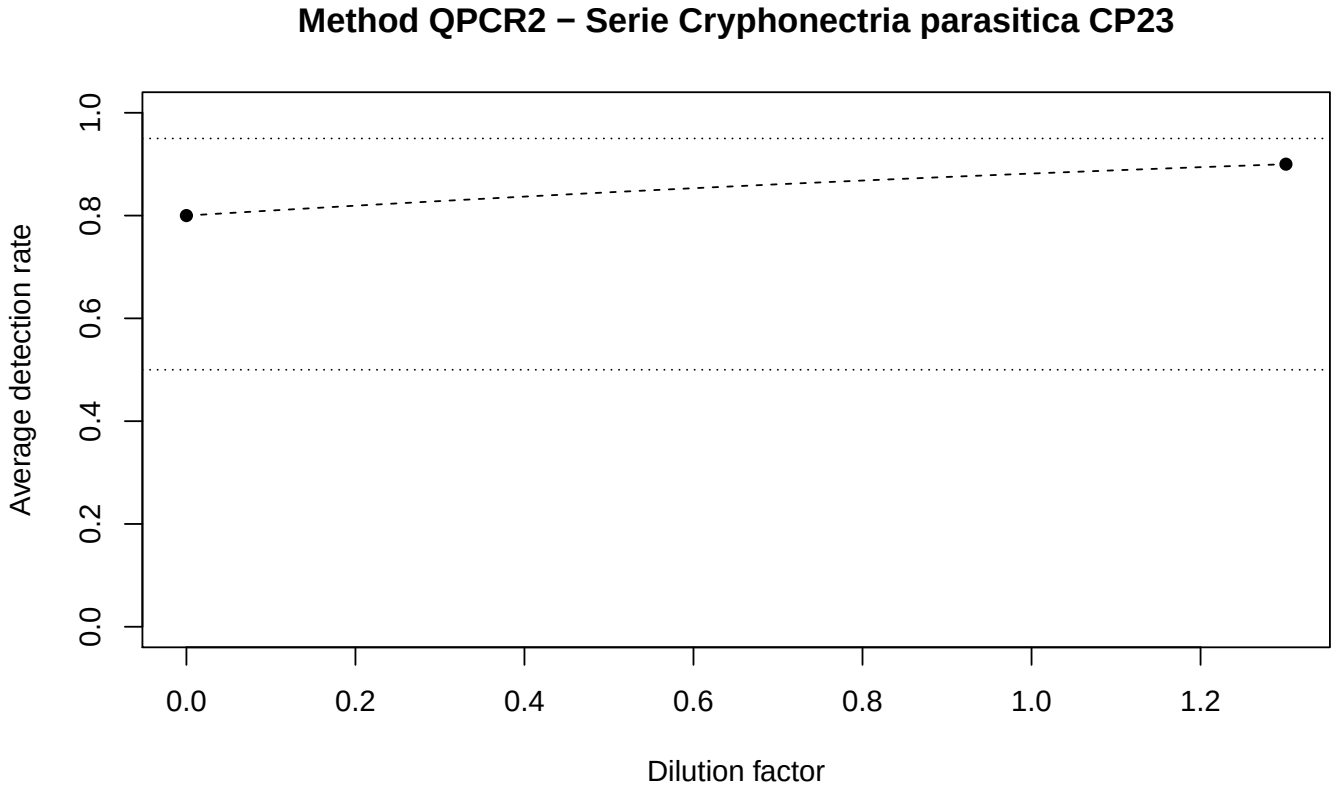

Method PCR1 − Serie Cryphonectria parasitica CP24
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Dilution factor


Method QPCR1 − Serie Cryphonectria parasitica CP10
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Dilution factor


Method QPCR1 − Serie Cryphonectria parasitica CP23
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Dilution factor


Method QPCR1 − Serie Cryphonectria parasitica CP24
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Dilution factor


Method QPCR2 − Serie Cryphonectria parasitica CP10
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Dilution factor


Method QPCR2 − Serie Cryphonectria parasitica CP23
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Dilution factor

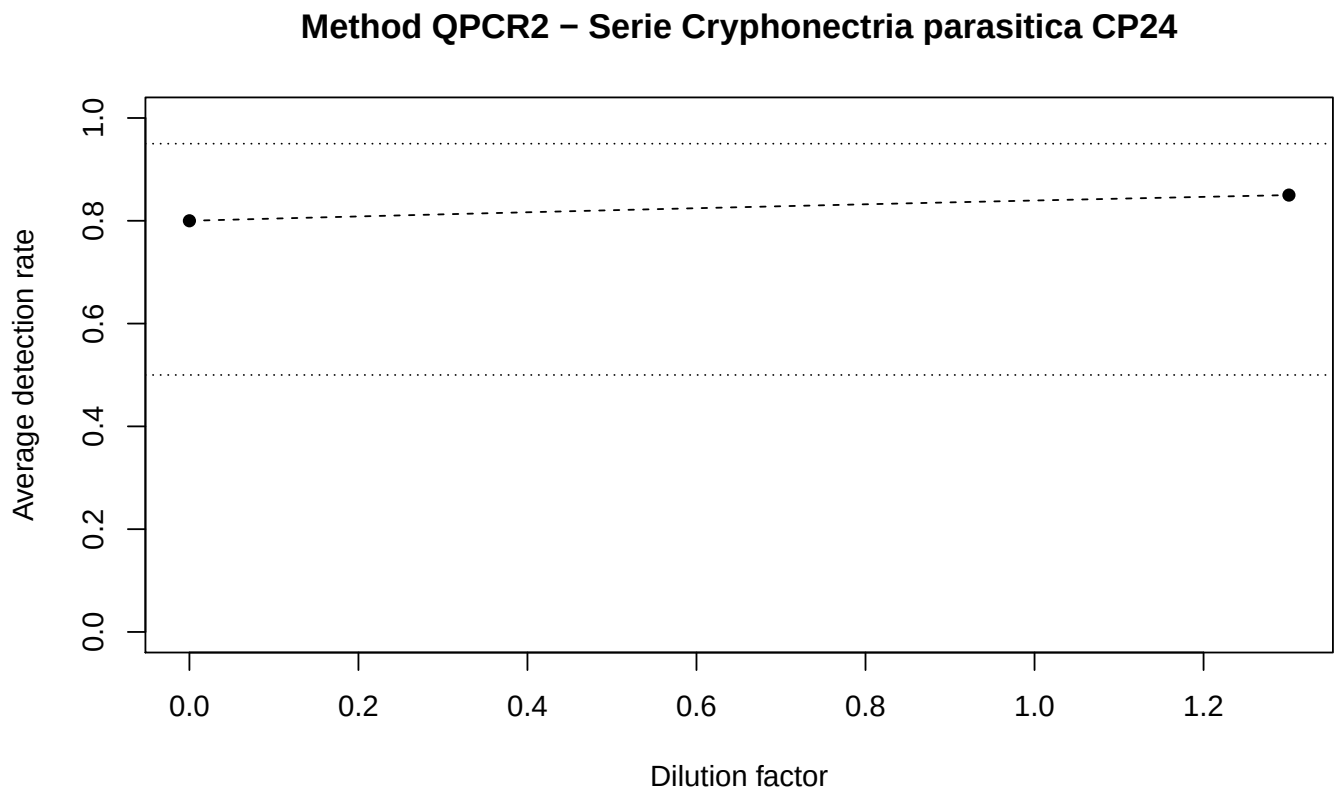


Table 2: Detection limits (log dilution factor) at 50% rate by methods

	PCR1	QPCR1	QPCR2
Cryphonectria parasitica CP10	NA	NA	NA
Cryphonectria parasitica CP23	NA	NA	NA
Cryphonectria parasitica CP24	NA	NA	NA

Table 3: Detection limits (log dilution factor) at 95% rate by methods

	PCR1	QPCR1	QPCR2
Cryphonectria parasitica CP10	NA	NA	NA
Cryphonectria parasitica CP23	NA	NA	NA
Cryphonectria parasitica CP24	NA	NA	NA

Diagnostic sensitivity and specificity

Diagnostic sensitivity is estimated as the detection rate on samples with positive reference status, and the diagnostic specificity as the non detection rate on samples with negative reference status. Hence, those two parameters are heavily dependent on the choice of the reference positive and negative samples. Has the number of different samples vary from laboratories and methods, the comparison of the following estimates between methods has to be conducted with caution.

Table 4: Diagnostic sensitivity by methods

	PCR1	QPCR1	QPCR2
DSE	0.547	0.740	0.825
LCL	0.467	0.664	0.726
UCL	0.624	0.804	0.894

Table 5: Diagnostic specificity by methods

	PCR1	QPCR1	QPCR2
DSP	0.904	0.860	0.933
LCL	0.834	0.783	0.836
UCL	0.947	0.913	0.978

Likelihood ratios

Another drawback of a poor choice of reference samples is that 100% sensitivity or specificity estimates will impair the estimation of likelihood ratios, as they have either (1 - sensitivity) ou (1 - specificity) as denominator. Perfect specificity will then leads to infinite positive likelihood ratio, while perfect sensitivity leads to infinite negative likelihood ratio, obliterating the effect of the other parameter in the estimate. In this case, confidence limits can give some additional information by giving upper and lower bounds for this ratio, event in the case of an infinite estimation.

Table 6: Likelihood ratios by methods

	PCR1	QPCR1	QPCR2
LR+	5.665	5.272	12.375
LCL	3.252	3.388	5.147
UCL	10.156	8.457	31.598
LR-	1.993	3.306	5.333
LCL	1.670	2.529	3.401
UCL	2.423	4.414	8.741

Other diagnostic parameters

Table 7: Other diagnostic parameters by methods

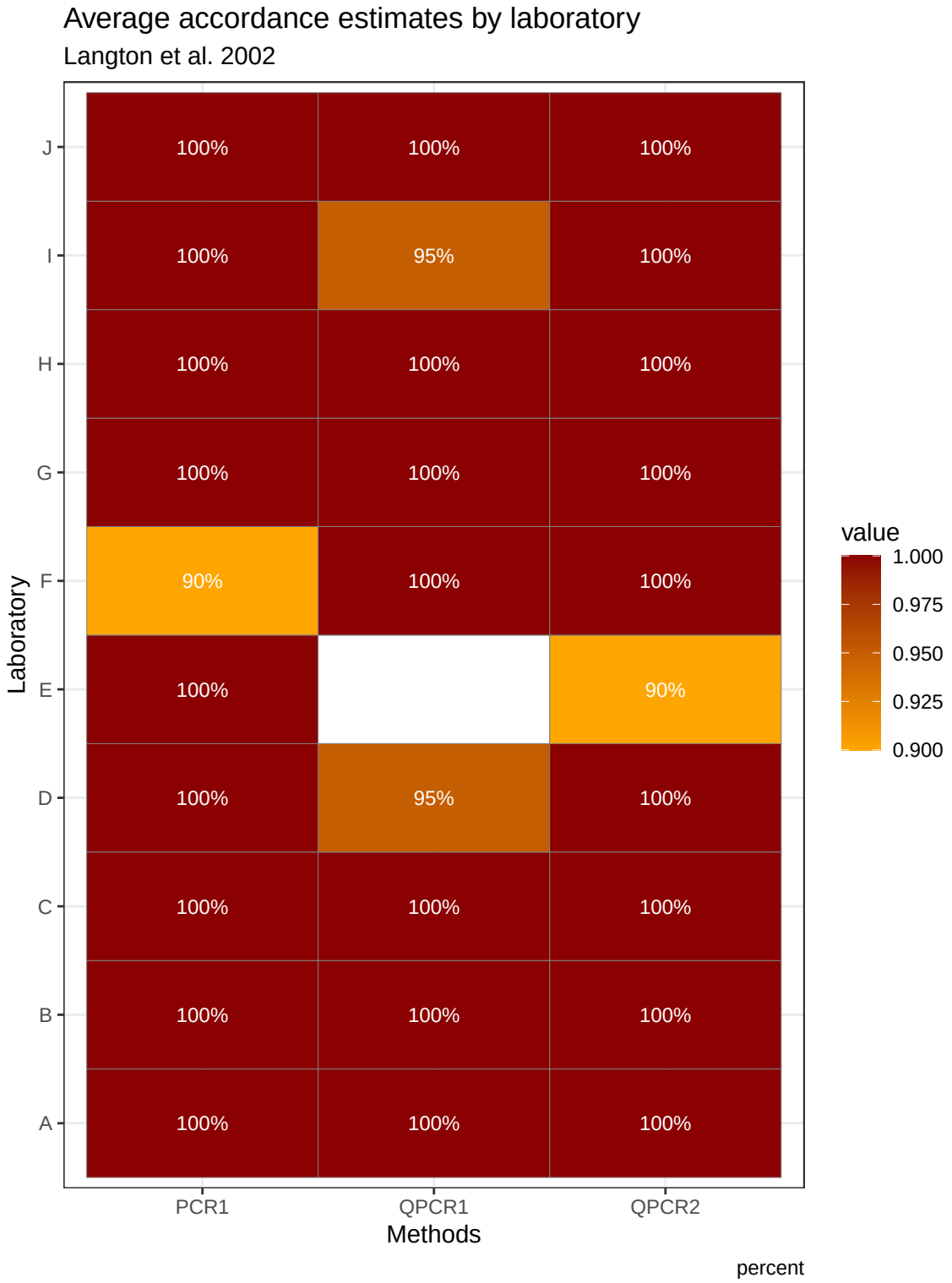
	PCR1	QPCR1	QPCR2
Accuracy	0.701	0.792	0.871
LCL	0.643	0.738	0.805
UCL	0.753	0.836	0.918
Power	0.358	0.541	0.673
LCL	0.299	0.473	0.575
UCL	0.422	0.608	0.758
Rate True Positive	0.882	0.874	0.943
LCL	0.799	0.804	0.858
UCL	0.934	0.922	0.982
Rate True Negative	0.602	0.715	0.800
LCL	0.527	0.634	0.691
UCL	0.673	0.784	0.878

Outliers detection

Outliers can be detected by decomposing the calculation of previous parameters to the laboratory and/or the sample level, and looking for strong individual deviation among them. Those deviations can be investigated by expert, leading to the possible exclusion of the corresponding results from the general analysis if necessary.

Repetability & reproducibility

Detailed accordance splitted by laboratories can help detecting labs which had some trouble with their repetability.



Analytical sensitivity

For each method and each laboratory, data of the diluted samples were used to adjust binomial generalized linear models (bGLM) with logit link between the target concentration (expressed in log scale) and the detection status.

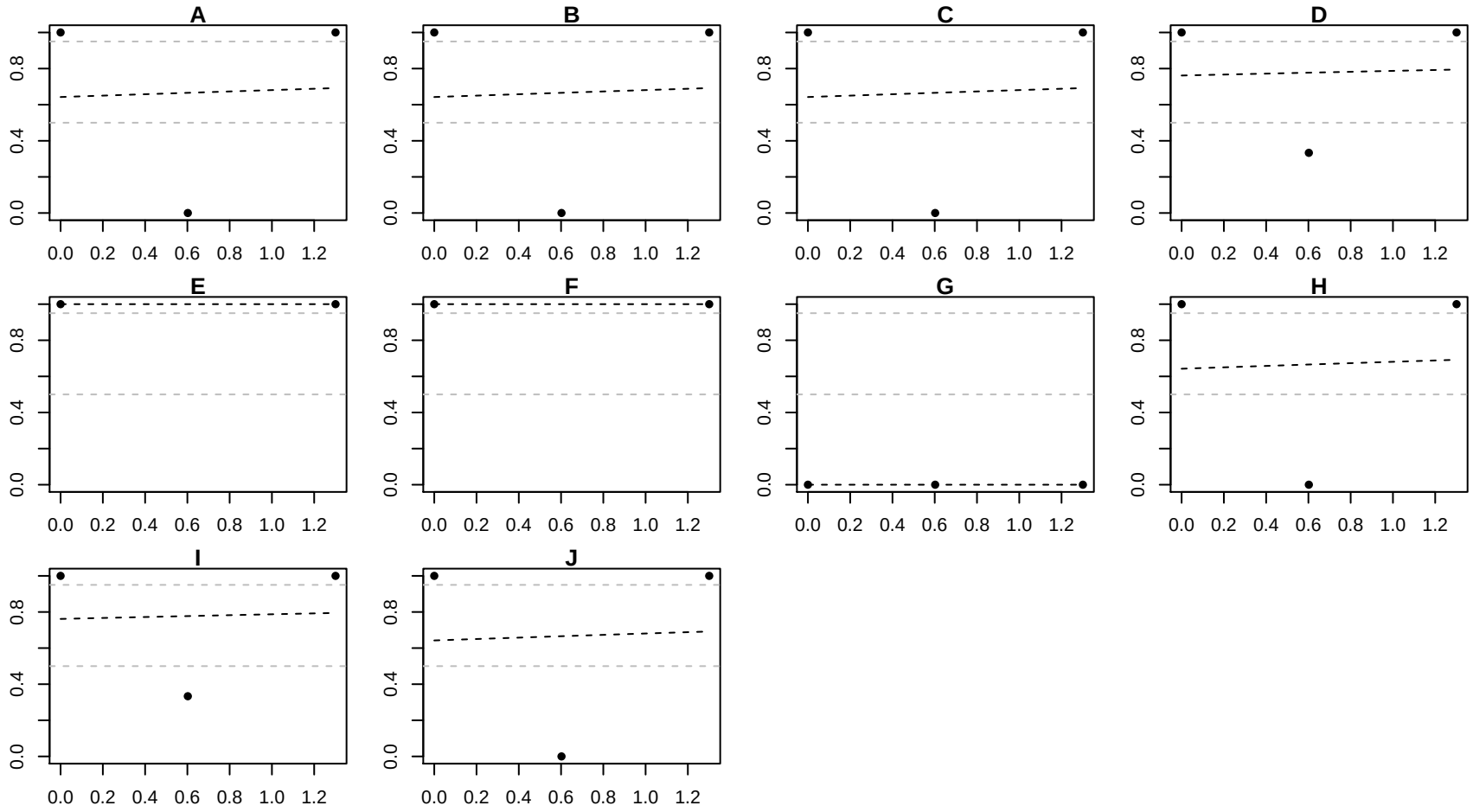
From those results, some warnings must be addressed when using bGLM to calculate LOD. Those models are based on the acceptable assumption that the probability of detection is decreasing when the dilution level increases. At the laboratory level, this hypothesis can be violated in different cases :

- when all the samples show the same status whatever the dilution level;
- when the observed detection rate shows non monotonous behaviour (e.g. decrease then increase again).

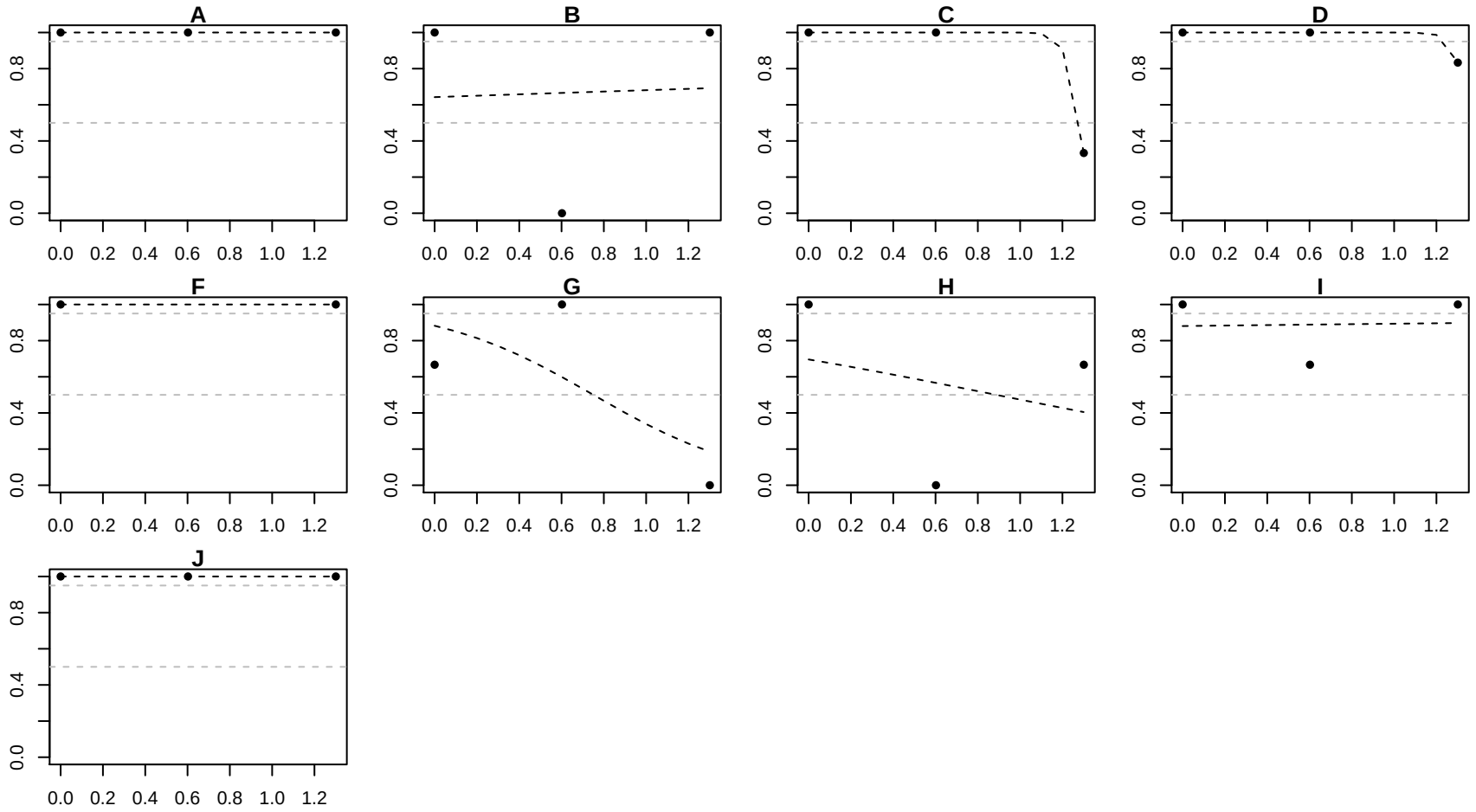
When this happens, the calculated LODxx calculated from the model cannot be trusted. Hence, expert verification of the fit of the model is mandatory before interpreating the calculated LOD value.

On the following figures, observed status and adjusted bGLM are shown for each method and laboratories. The problematic cases are very easy to identify on such graphical representation.

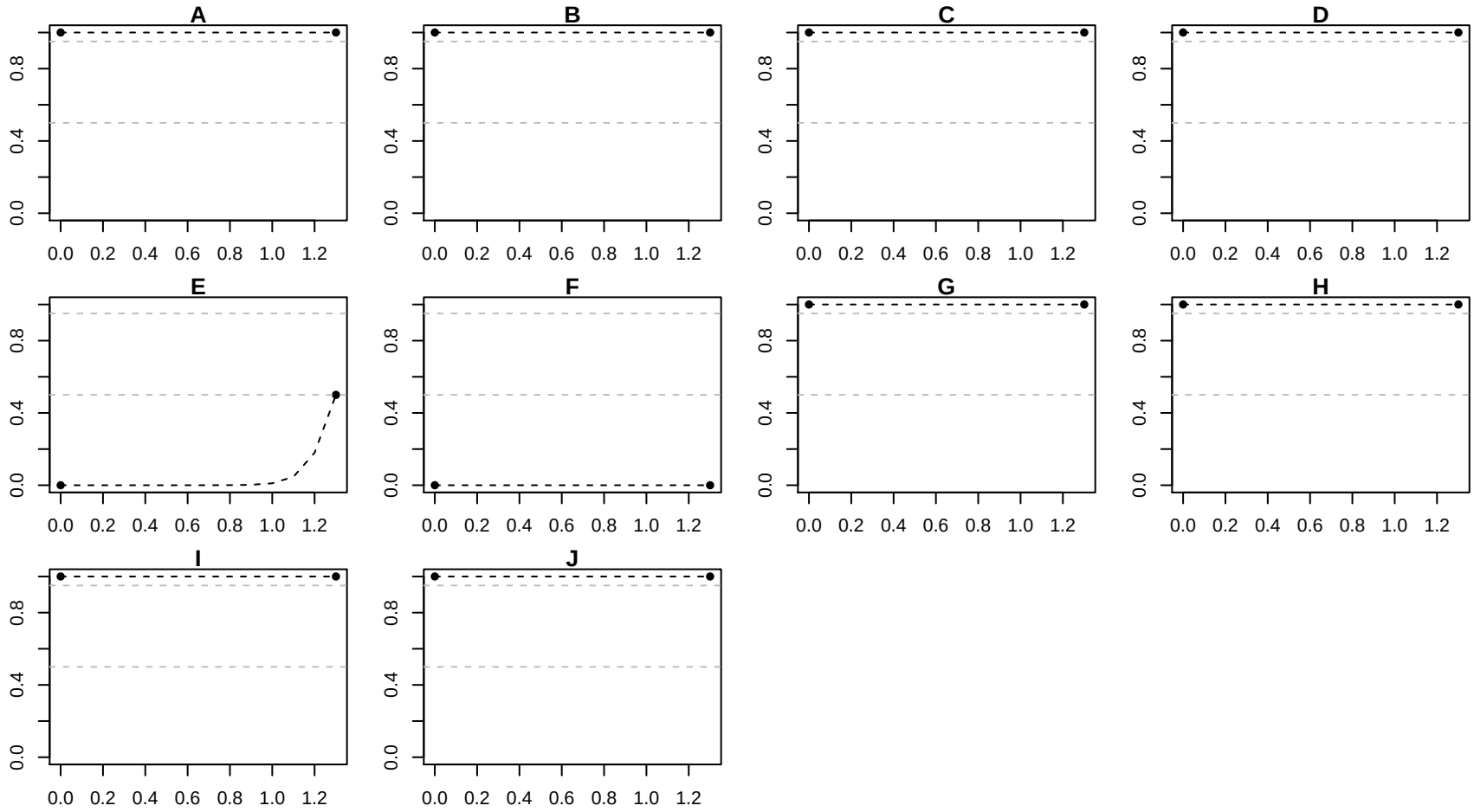
Method PCR1



Method QPCR1

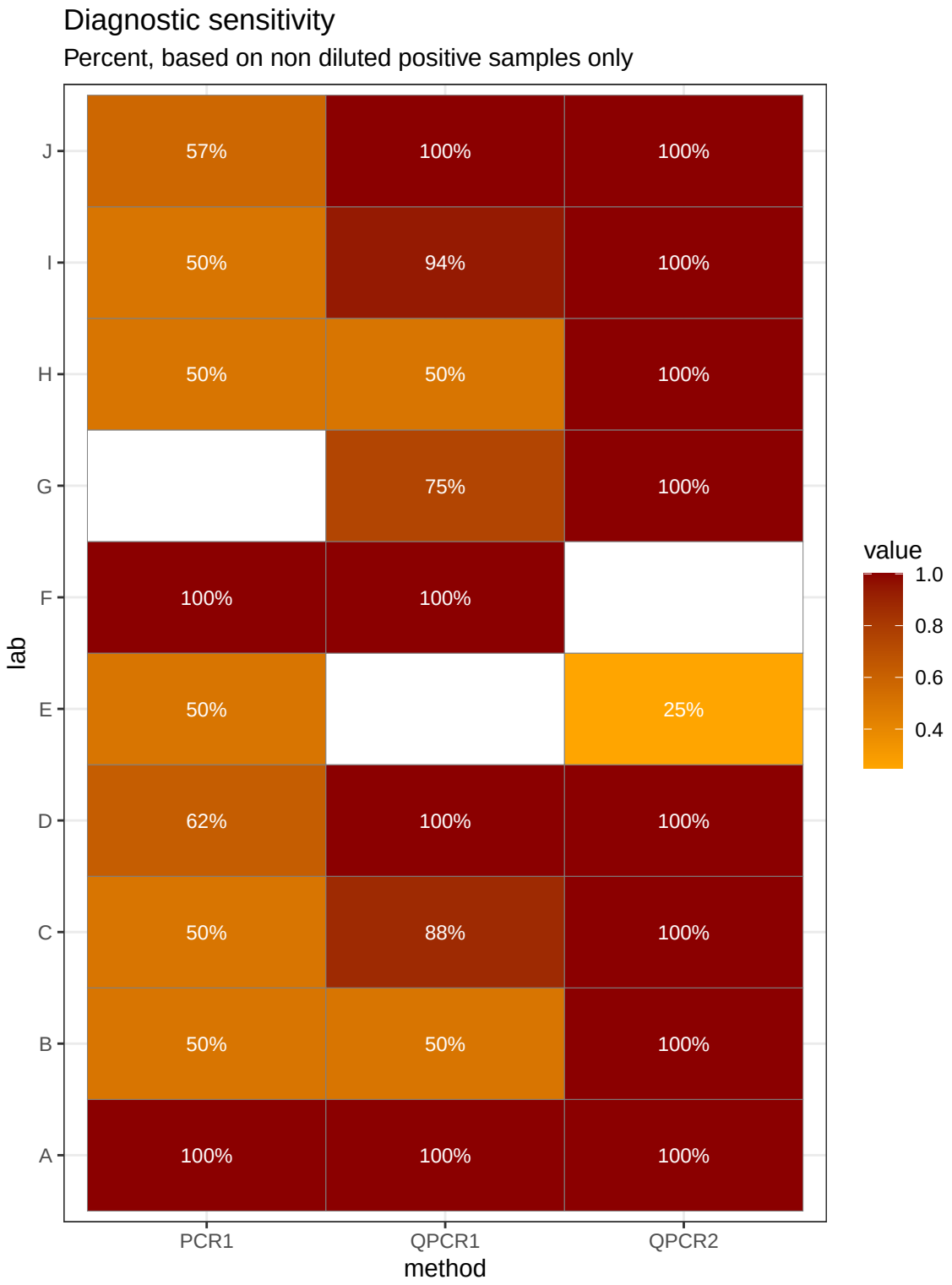


Method QPCR2



Diagnostic sensitivity and diagnostic specificity

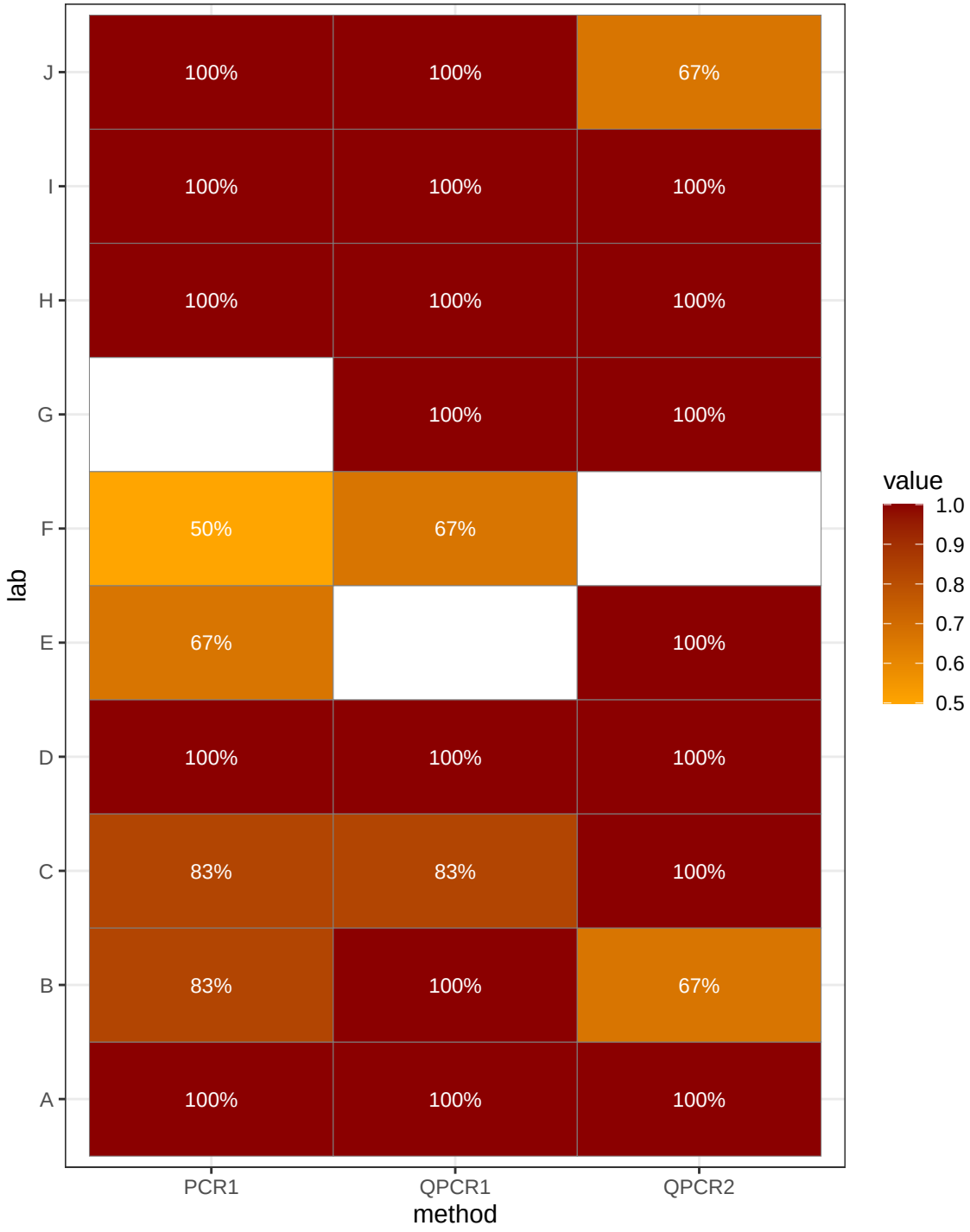
Even in the case of very high diagnostic sensitivity or specificity, those parameters can still be usefull to detect outliers, which show below than average performance event in those situations.



Negative samples with very low confusing sequences will lead to the same results for specificity (very frequent 100% estimation), but can also detect some poor performance of labs, methods or a combination of the two.

Diagnostic specificity

Percent, based on non diluted samples only



VALITEST TPS report

PPV RT-PCR data

Yves Brostaux

28 septembre 2021

Files reading and data preprocessing

Fusarium data are stored in one csv files, with the following columns :

- sampid, sample ID
- method, method ID
- lab, lab ID
- test, result of the test (binary 0/1)
- ref, reference status of each sample (binary 0/1)
- dilution (dilution factor for diluted samples)
- link, ID of the original sample of the dilution serie

Following corrections were applied to the data prior to analysis : - reported results not equal to 0 ou 1 have been set to 2 (undetermined)

Those files have been read and combined into different data tables for the analysis.

- tidydata, 1 line per status result, same organization as the file
- dilu.dat1, dilu.dat2, dilu.dat2, subsets of tidydata including only diluted samples, with additional column dilu (dilution exponent (negative base 10)), corresponding to 3 dilutions series of samples 97-009, T+PPV#23, 17-0125-05 respectively
- spse.dat, subset of tidydata including only non diluted samples

```
## tidydata
```

```
##   sampid        method lab test ref dilution    link
## 1      A Levy & Hadidi L01    0   0        0     NIC
## 2      B Levy & Hadidi L01    1   1        1     PIC
## 3      C Levy & Hadidi L01    1   1        1    SK59
## 4      D Levy & Hadidi L01    1   1        1 00-0034
## 5      E Levy & Hadidi L01    1   1        1 94-0025
## 6      F Levy & Hadidi L01    1   1       10  97-009
```

```
## dilu.dat1
```

```
##    sampid        method lab test ref dilution   link dilu
## 6       F Levy & Hadidi L01    1   1       10 97-009    1
## 7       G Levy & Hadidi L01    1   1      100 97-009    2
## 8       H Levy & Hadidi L01    1   1     1000 97-009    3
## 9       I Levy & Hadidi L01    1   1    10000 97-009    4
## 28      F Levy & Hadidi L03    1   1       10 97-009    1
## 29      G Levy & Hadidi L03    1   1      100 97-009    2
```

```
## dilu.dat2
```

```
##    sampid        method lab test ref dilution     link    dilu
## 12      L Levy & Hadidi L01    1   1       50 T+PPV#23 1.69897
## 13      M Levy & Hadidi L01    1   1      250 T+PPV#23 2.39794
## 14      N Levy & Hadidi L01    1   1     1250 T+PPV#23 3.09691
```

```
## 34      L Levy & Hadidi L03    1   1       50 T+PPV#23 1.69897
## 35      M Levy & Hadidi L03    1   1      250 T+PPV#23 2.39794
## 36      N Levy & Hadidi L03    1   1     1250 T+PPV#23 3.09691
```

```
## dilu.dat3
```

```
##    sampid        method lab test ref dilution       link dilu
## 15      O Levy & Hadidi L01    1   1       10 17-0125-05    1
## 16      Q Levy & Hadidi L01    1   1      100 17-0125-05    2
## 17      R Levy & Hadidi L01    1   1     1000 17-0125-05    3
## 18      S Levy & Hadidi L01    1   1    10000 17-0125-05    4
## 37      O Levy & Hadidi L03    1   1       10 17-0125-05    1
## 38      Q Levy & Hadidi L03    1   1      100 17-0125-05    2
```

```
## spse.dat
```

```
##    sampid        method lab test ref dilution     link
## 1       A Levy & Hadidi L01    0   0        0      NIC
## 2       B Levy & Hadidi L01    1   1        1      PIC
## 3       C Levy & Hadidi L01    1   1        1     SK59
## 4       D Levy & Hadidi L01    1   1        1  00-0034
## 5       E Levy & Hadidi L01    1   1        1  94-0025
## 10      J Levy & Hadidi L01    1   1       10 T+PPV#10
```

```
##      method
## lab   Levy & Hadidi Qualiplante test according to OImos Wetzel
##   L01            22                                   22     22
##   L03            22                                   22     22
##   L04            22                                   22     22
##   L05            22                                   22     22
##   L06             0                                   22      0
##   L07            22                                   22     22
##   L09            22                                   22     22
##   L10            22                                   22     22
##   L12            22                                   22     22
##   L13            22                                   22     22
##   L15            22                                    0     22
##   L16            22                                    0     22
##   L17            22                                   22     22
##   L18             0                                   22      0
##   NRC            22                                   22     22
```

Repeatability and reproducibility

As none of the measure of this test was repeated, classical repetability and reproducibility of the methods are impossible to assess.

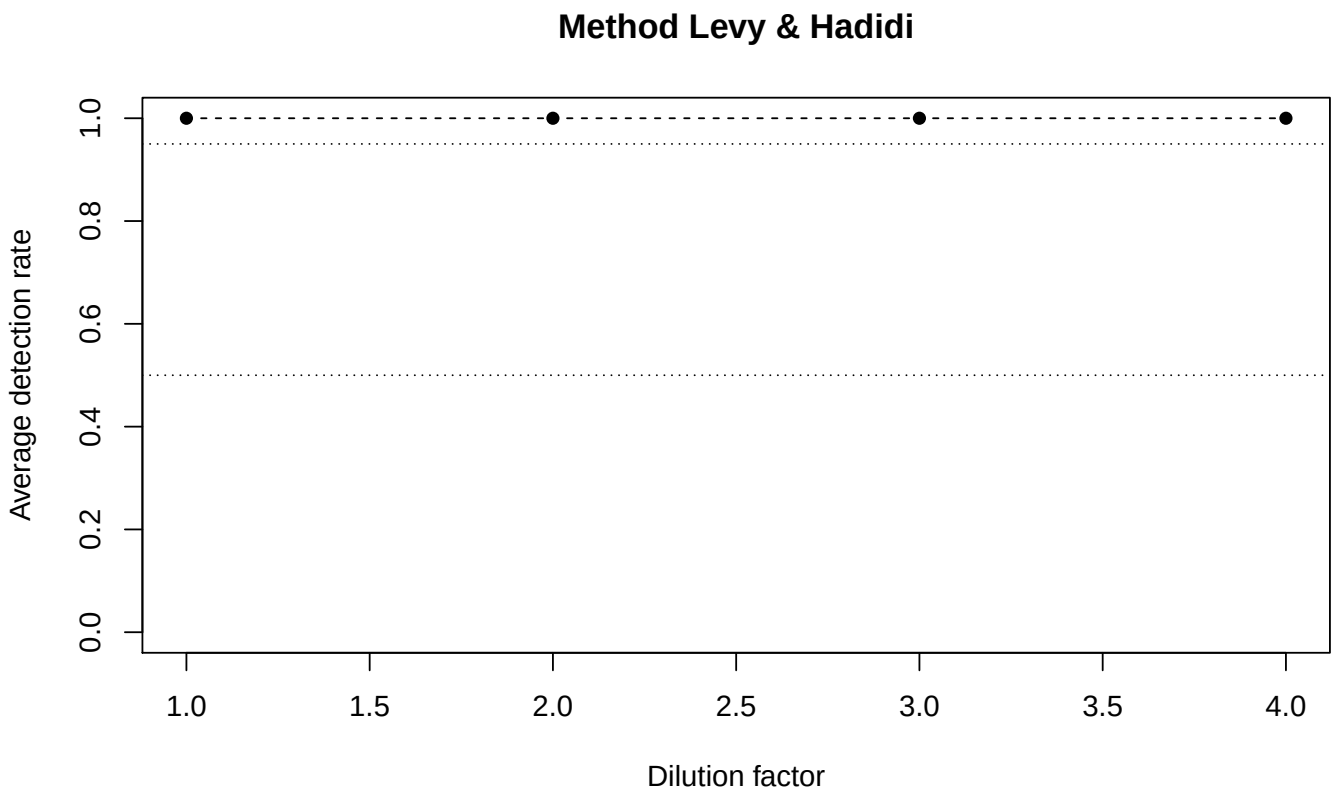
Analytical sensitivity

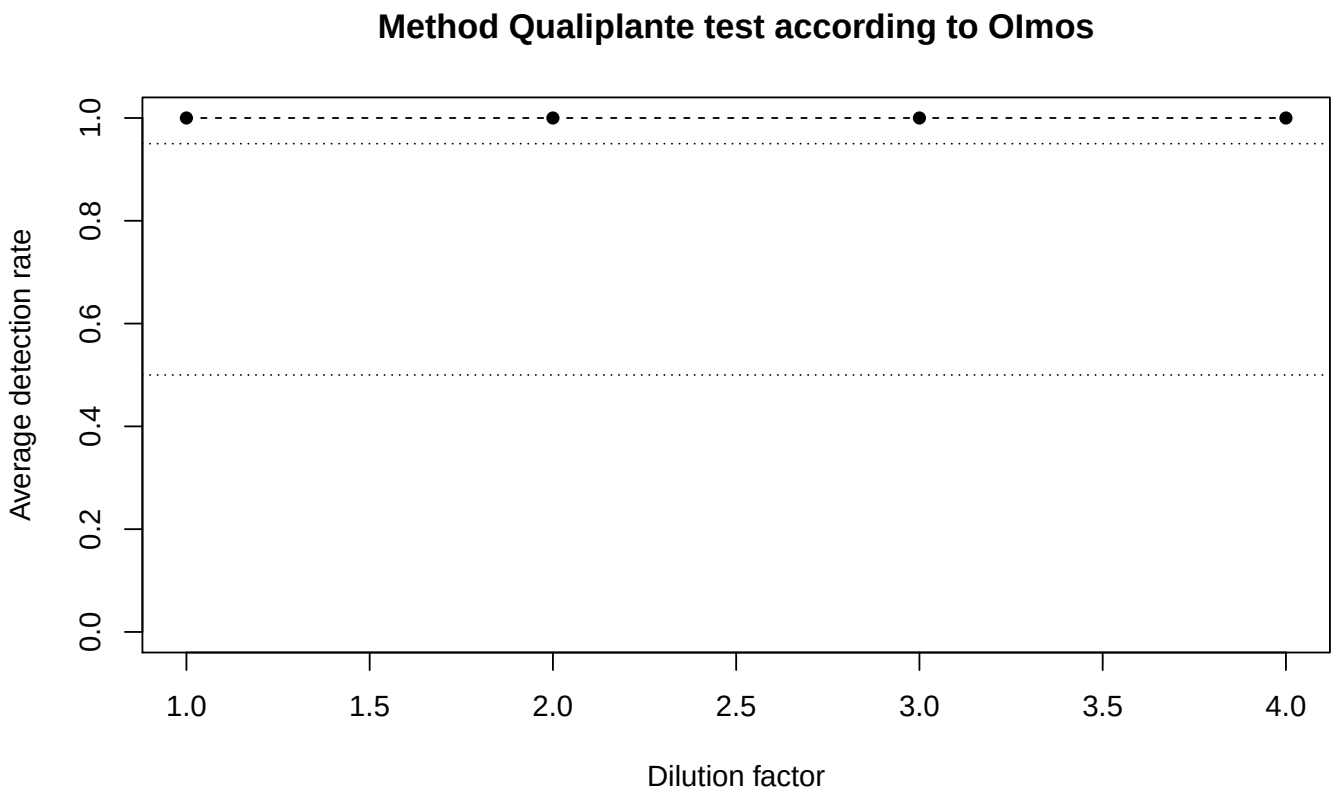
For each method , data of the diluted samples were used to adjust binomial generalized linear models (bGLM) with logit link between the dilution (expressed by the base 10 negative exponent of the corresponding dilution) and the detection status. The number of dilution level being very limited, the adjustement of bGLM is not always possible as this method require at least 5 levels, and the laboratory effect has been neglected as no replicates are available.

When possible, based on those models, dilutions corresponding to a 50% or 95% probability of detection have been calculated as an example of the possible LOD to report (LOD50 and LOD95).

All this have been calculated on the three dilution series available in the data.

Dilution serie 97-009




Method Qualiplante test according to Olmos
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Dilution factor

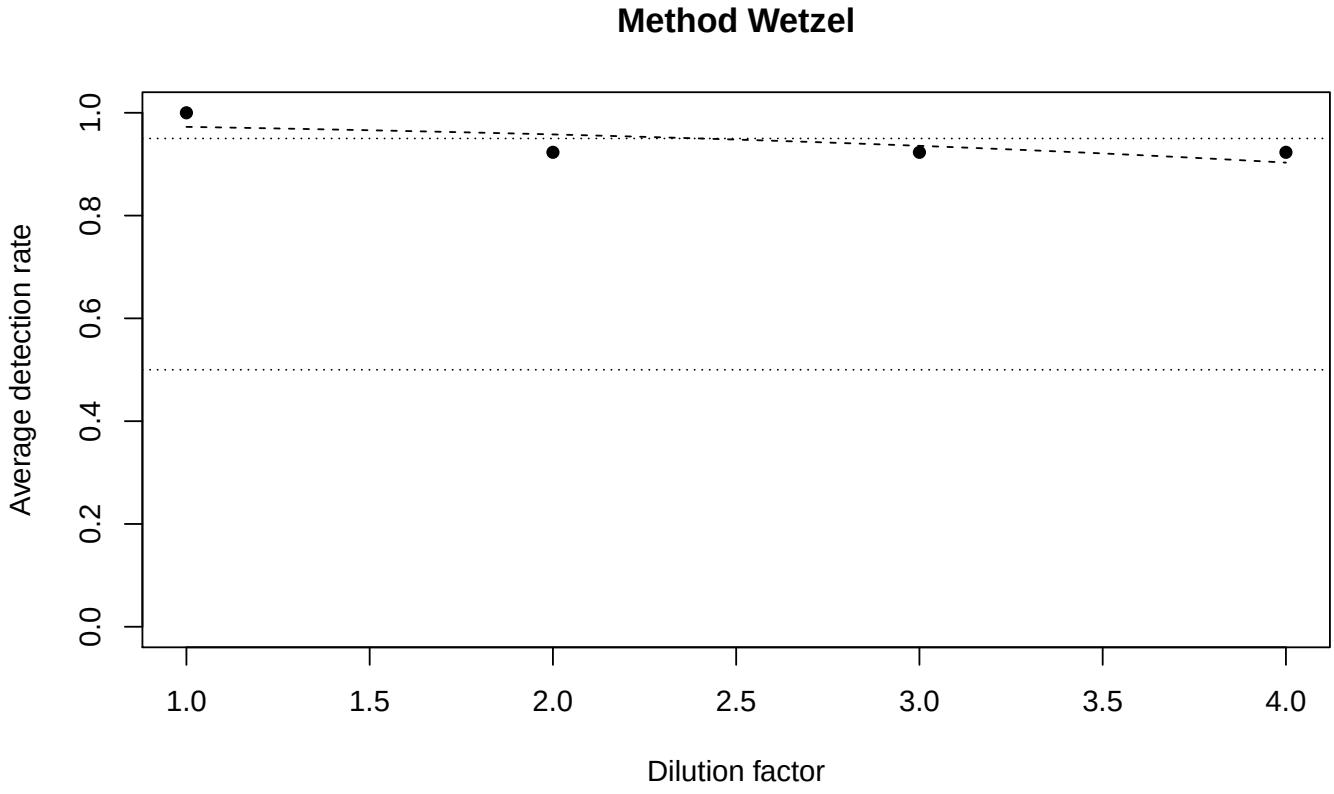

Method Wetzel
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Dilution factor

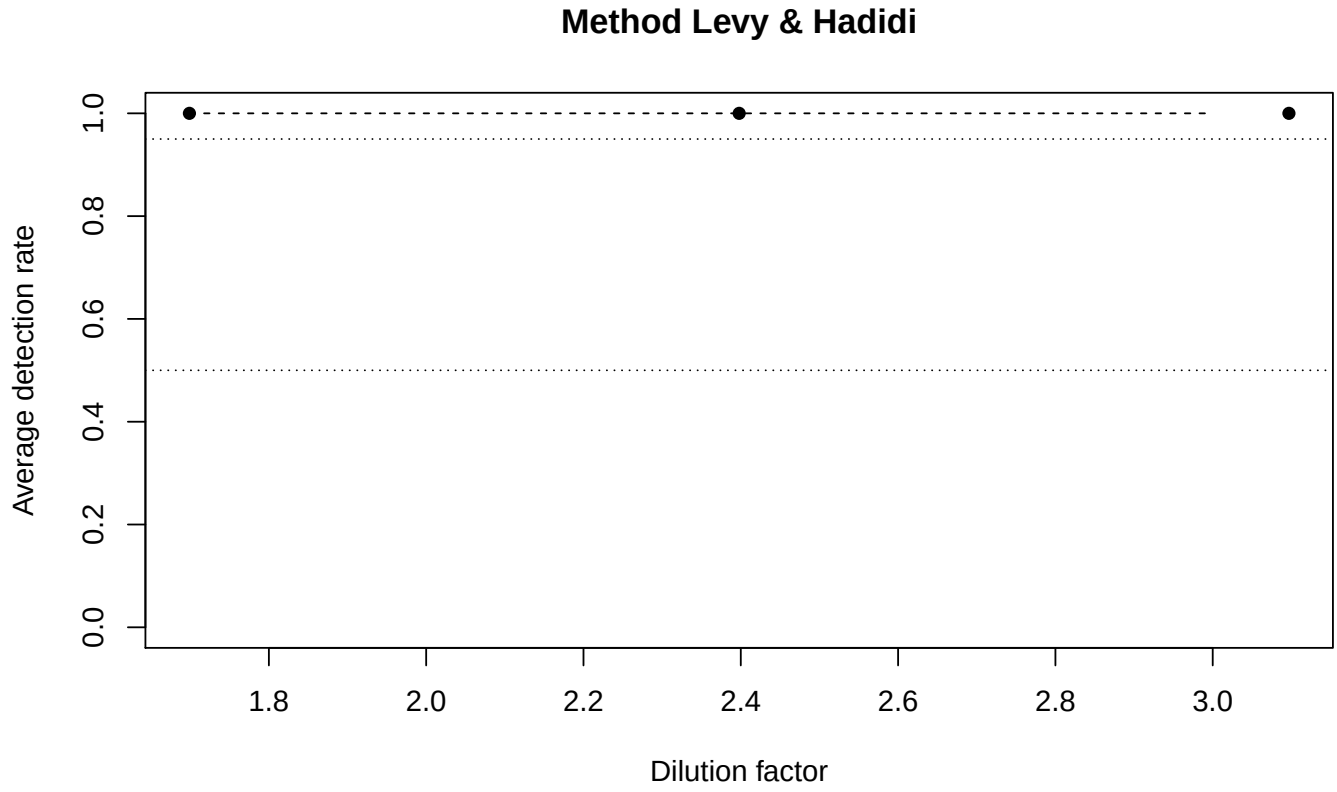
Table 1: Detection limits (log dilution factor) at 50% rate by methods

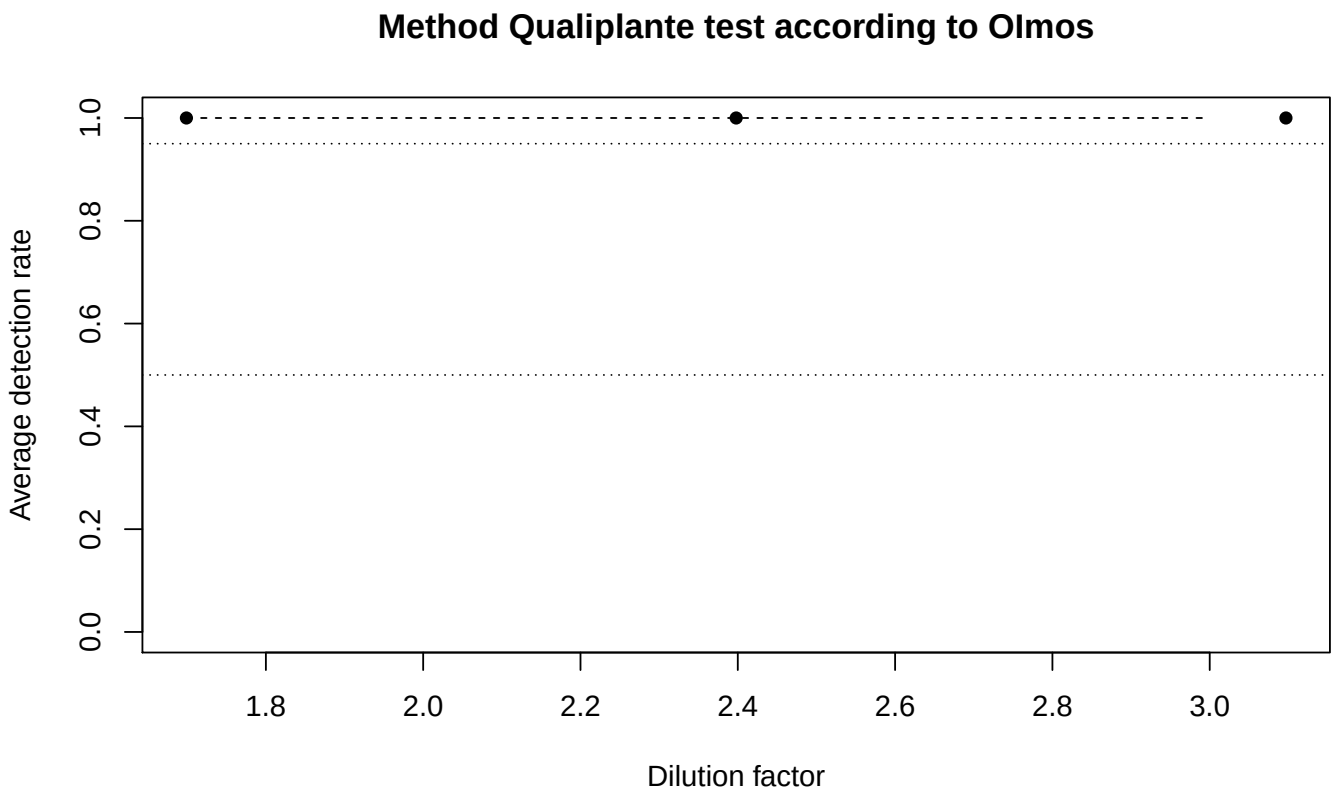
Levy & Hadidi	Qualiplante test according to OImos	Wetzel
NA	NA	NA

Table 2: Detection limits (log dilution factor) at 95% rate by methods

Levy & Hadidi	Qualiplante test according to OImos	Wetzel
NA	NA	2.4

Dilution serie T+PPV#23, 17-0125-05




Method Qualiplante test according to Olmos
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Dilution factor

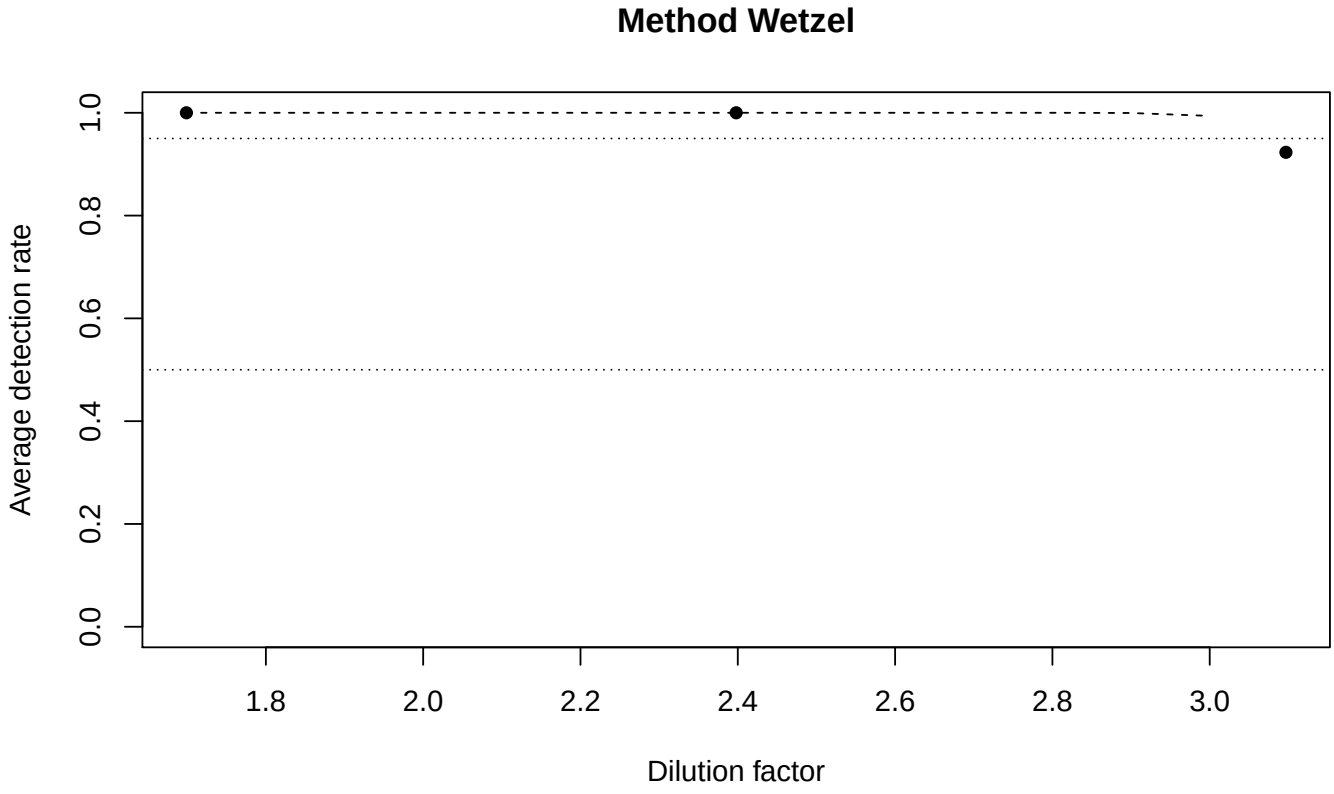

Method Wetzel
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
1.8
2.0
2.2
2.4
2.6
2.8
3.0
Dilution factor

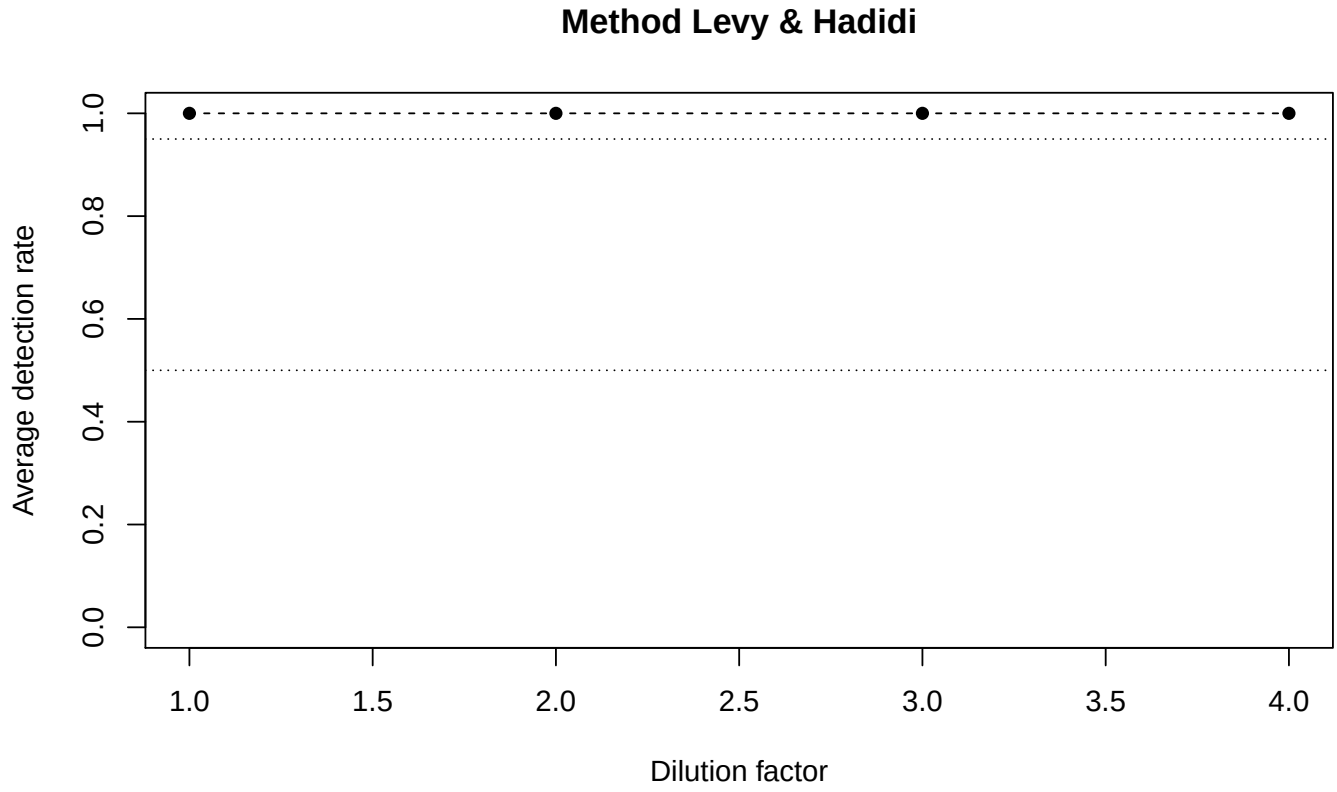
Table 3: Detection limits (log dilution factor) at 50% rate by methods

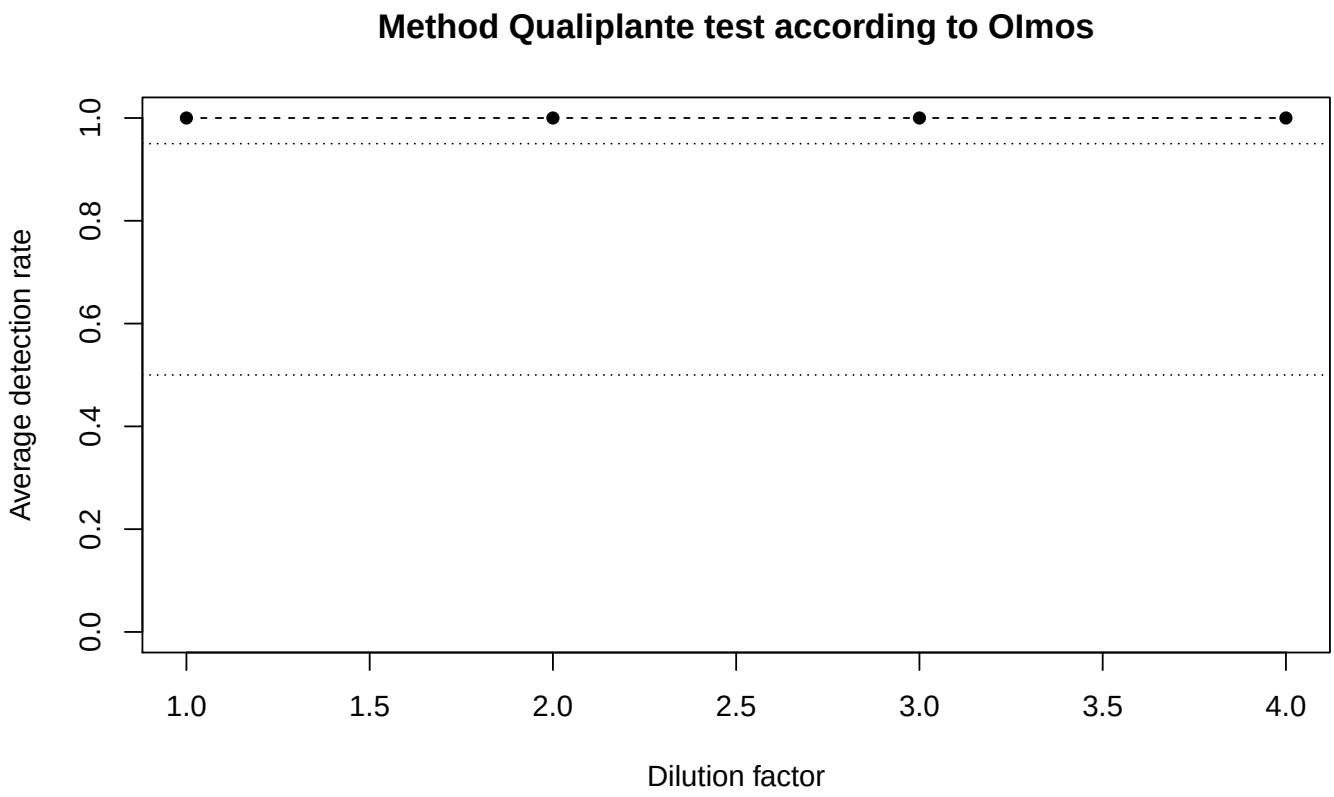
Levy & Hadidi	Qualiplante test according to OImos	Wetzel
NA	NA	NA

Table 4: Detection limits (log dilution factor) at 95% rate by methods

Levy & Hadidi	Qualiplante test according to OImos	Wetzel
NA	NA	3.1

Dilution 17-0125-05




Method Qualiplante test according to Olmos
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Dilution factor

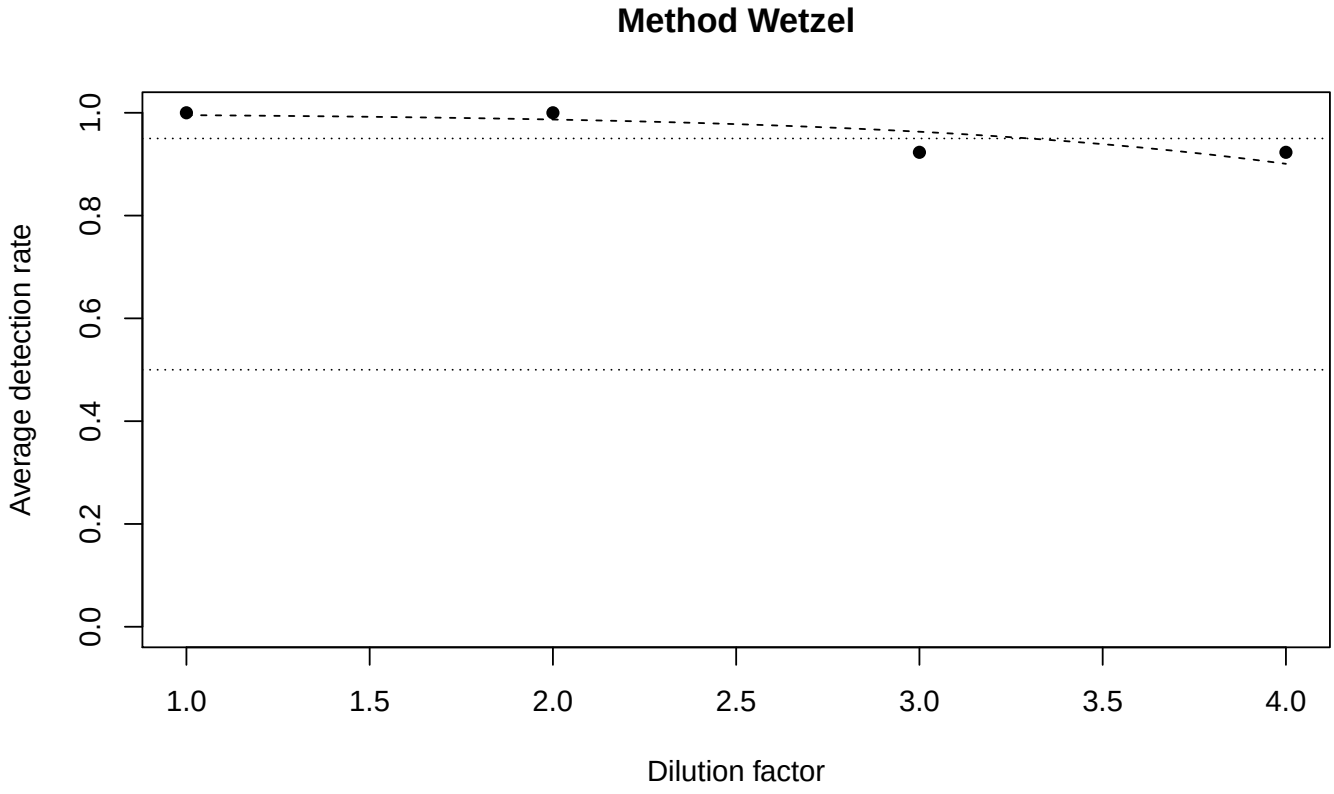

Method Wetzel
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Dilution factor

Table 5: Detection limits (log dilution factor) at 50% rate by methods

Levy & Hadidi	Qualiplante test according to OImos	Wetzel
NA	NA	NA

Table 6: Detection limits (log dilution factor) at 95% rate by methods

Levy & Hadidi	Qualiplante test according to OImos	Wetzel
NA	NA	3.3

Diagnostic sensitivity and specificity

Diagnostic sensitivity is estimated as the detection rate on samples with positive reference status, and the diagnostic specificity as the non detection rate on samples with negative reference status. Hence, those two parameters are heavily dependent on the choice of the reference positive and negative samples. Has the number of different samples vary from laboratories and methods, the comparison of the following estimates between methods has to be conducted with caution.

Table 7: Diagnostic sensitivity and specificity by methods

	Levy & Hadidi	Qualiplante test according to OImos	Wetzel
DSE	1.000	1.000	0.962
LCL	0.957	0.957	0.902
UCL	1.000	1.000	0.988
DSP	0.846	0.821	0.769
LCL	0.699	0.670	0.615
UCL	0.931	0.913	0.876

Likelihood ratios

Another drawback of a poor choice of reference samples is that 100% sensitivity or specificity estimates will impair the estimation of likelihood ratios, as they have either (1 - sensitivity) ou (1 - specificity) as denominator. Perfect specificity will then leads to infinite positive likelihood ratio, while perfect sensitivity leads to infinite negative likelihood ratio, obliterating the effect of the other parameter in the estimate. In this case, confidence limits can give some additional information by giving upper and lower bounds for this ratio, event in the case of an infinite estimation.

Table 8: Likelihood ratios by methods

	Levy & Hadidi	Qualiplante test according to OImos	Wetzel
LR+	6.500	5.571	4.167
LCL	3.364	3.061	2.503
UCL	13.798	11.139	7.612
LR-	Inf	Inf	20.000
LCL	23.738	23.015	7.971
UCL	Inf	Inf	51.742

Other diagnostic parameters

Table 9: Other diagnostic parameters by methods

	Levy & Hadidi	Qualiplante test according to OImos	Wetzel
Accuracy	0.958	0.951	0.909
LCL	0.910	0.901	0.850
UCL	0.983	0.978	0.947
Power	0.945	0.937	0.855
LCL	0.884	0.873	0.779
UCL	0.977	0.971	0.908
Rate True Positive	0.945	0.937	0.917
LCL	0.884	0.873	0.849
UCL	0.977	0.971	0.958
Rate True Negative	1.000	1.000	0.882
LCL	0.876	0.873	0.728
UCL	1.000	1.000	0.959

Outliers detection

Outliers can be detected by decomposing the calculation of previous parameters to the laboratory and/or the sample level, and looking for strong individual deviation among them. Those deviations can be investigated by expert, leading to the possible exclusion of the corresponding results from the general analysis if necessary.

Repetability & reproducibility

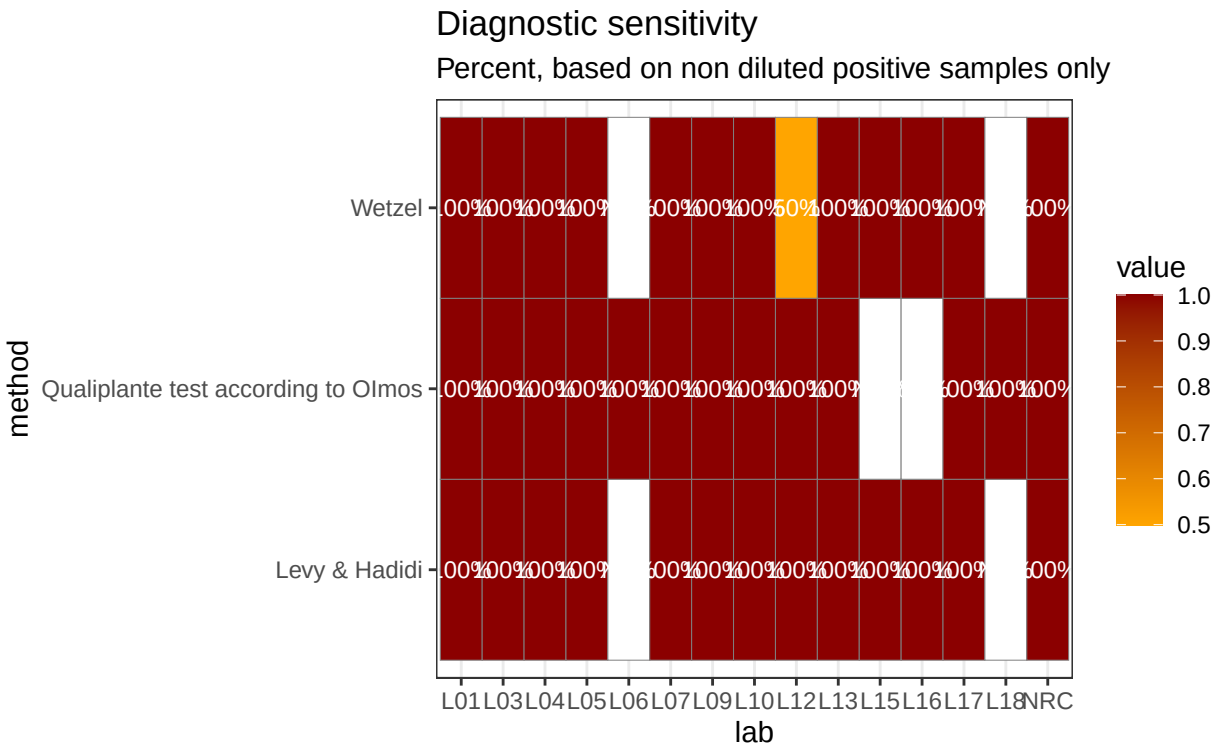
Impossible due to the absence of replicates

Analytical sensitivity

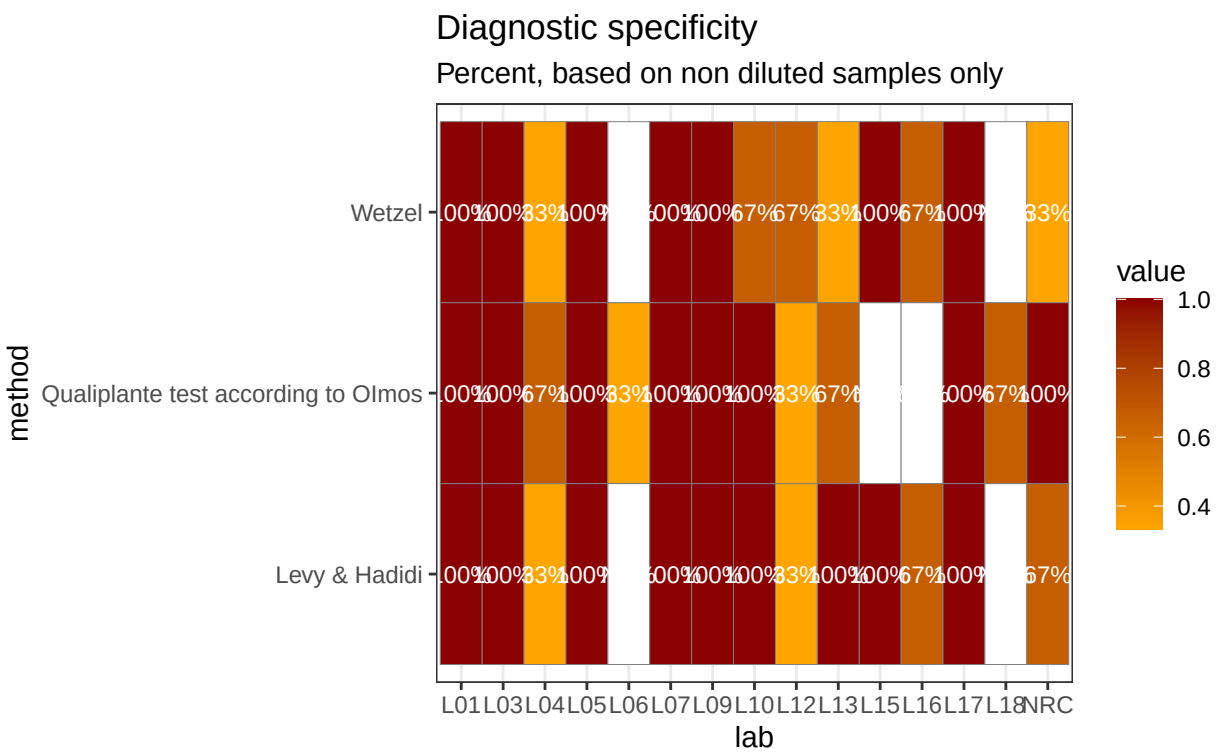
Impossible due to limited number of dilution levels

Diagnostic sensitivity and diagnostic specificity

Even in the case of very high diagnostic sensitivity or specificity, those parameters can still be usefull to detect outliers, which show below than average performance event in those situations. On the following figure, we can see that laboratory 12 had an issue in the application of the Wetzel method , as it only get 50% sensitivity on eight positive samples, while every other lab reaches 100%.



Negative samples with very low confusing sequences will lead to the same results for specificity (very frequent 100% estimation), but can also detect some poor performance of labs, methods or a combination of the two. Laboratory 12 shows again a poor diagnostic specificity, along with laboratory 04.



VALITEST TPS2 report

TSWVdata

Yves Brostaux

28 septembre 2021

Files reading and data preprocessing

TSWV data are stored in one csv files, with the following columns :

- Sample, Method, Laboratory, Replicat
- Results and Status (binary)
- Concentration/quantity/dilution
- Linked sample, Reference method and Sample info

Samples ID has been replaced by a new identification code which is unique to the biological sample across methods (A1-A22). Correspondance table with the sample info is shown below.

Samples A19 to A22 have a dilution information but are isolated from the dilution serie. They are then considered as raw samples, so are included in the general performance data (spse.dat) and excluded from the dilution serie data (dilu.dat). Samples A17 and A18, corresponding to the highest concentration in the dilution serie, is included in both dilution serie data and general performance data.

Status of samples A9 to A12 is reported as negative. It should be positive as status is independant from alleged method sensitivity.

Replicates structures seems to contains two levels, as 4 results are linked to one base sample for some laboratories, but those results were divided in two samples id with two replicates. Without further information, each sample id was considered as an independant sample for R&R study.

After correction , the date were combined into different data tables for the analysis.

- tidydata, 1 line per status result,
  - sampid, sample ID
  - method, method ID
  - lab, lab ID
  - test, result of the test (binary 0/1)
  - ref, reference status of each sample (binary 0/1)
  - dilution (dilution factor for diluted samples)
  - link, ID of the original sample of the dilution serie
- dilu.dat, subset of tidydata including only diluted samples, with additional column dilu (target concentration (log base))
- spse.dat, subset of tidydata excluding the dilution serie data, except its the highest concentration sample.

DISCLAIMER :

As it can be noted below, different dilutions of the samples A19-A22 has been used for different methods. As a result, those methods were not tested against the same samples and their performance results (except for specifity and analytical sensitivity, not impacted) CANNOT be compared across methods.

```
## tidydata
```

```
##   sampid  method lab replic test ref dilution link refmeth           sinfo dilu
## 1     A4 ELISA_1 L03      1    1   0       NA           No  ANSV (PV-1027)   NA
## 2     A4 ELISA_1 L03      2    1   0       NA           No  ANSV (PV-1027)   NA
```

```
## 3     A5 ELISA_1 L03      1    0   0       NA           No CSNV2 (PV-0529)   NA
## 4     A5 ELISA_1 L03      2    0   0       NA           No CSNV2 (PV-0529)   NA
## 5     A6 ELISA_1 L03      1    1   0       NA           No  GRSV (PV-0205)   NA
## 6     A6 ELISA_1 L03      2    1   0       NA           No  GRSV (PV-0205)   NA
```

```
## dilu.dat
```

```
##    sampid                  sinfo ref
## 25     A9 TSWV (PV1175) 1000000x   1
## 27    A10 TSWV (PV1175) 1000000x   1
## 29    A11  TSWV (PV1175) 100000x   1
## 31    A12  TSWV (PV1175) 100000x   1
## 33    A13   TSWV (PV1175) 10000x   1
## 35    A14   TSWV (PV1175) 10000x   1
## 37    A15    TSWV (PV1175) 1000x   1
## 39    A16    TSWV (PV1175) 1000x   1
## 41    A17     TSWV (PV1175) 100x   1
## 43    A18     TSWV (PV1175) 100x   1
```

```
## spse.dat
```

```
##      sampid                          sinfo ref
## 7        A1               healthy tomato 1   0
## 9        A2               healthy tomato 1   0
## 11       A3               healthy tomato 2   0
## 1        A4                 ANSV (PV-1027)   0
## 3        A5                CSNV2 (PV-0529)   0
## 5        A6                 GRSV (PV-0205)   0
## 13       A7                INSV2 (PV-0281)   0
## 15       A8              TCSV  (PV-0390)     0
## 41      A17             TSWV (PV1175) 100x   1
## 43      A18             TSWV (PV1175) 100x   1
## 17      A19    TSWV (PV-0182) 100x diluted   1
## 1330    A19     TSWV (PV-0182) 10x diluted   1
## 1887    A19   TSWV (PV-0182) 1000x diluted   1
## 19      A20    TSWV (PV-0182) 100x diluted   1
## 1331    A20     TSWV (PV-0182) 10x diluted   1
## 1889    A20   TSWV (PV-0182) 1000x diluted   1
## 21      A21  TSWV (PV-0389) 10000x diluted   1
## 1332    A21   TSWV (PV-0389) 1000x diluted   1
## 1891    A21 TSWV (PV-0389) 100000x diluted   1
## 23      A22  TSWV (PV-0389) 10000x diluted   1
## 1333    A22   TSWV (PV-0389) 1000x diluted   1
## 1893    A22 TSWV (PV-0389) 100000x diluted   1
```

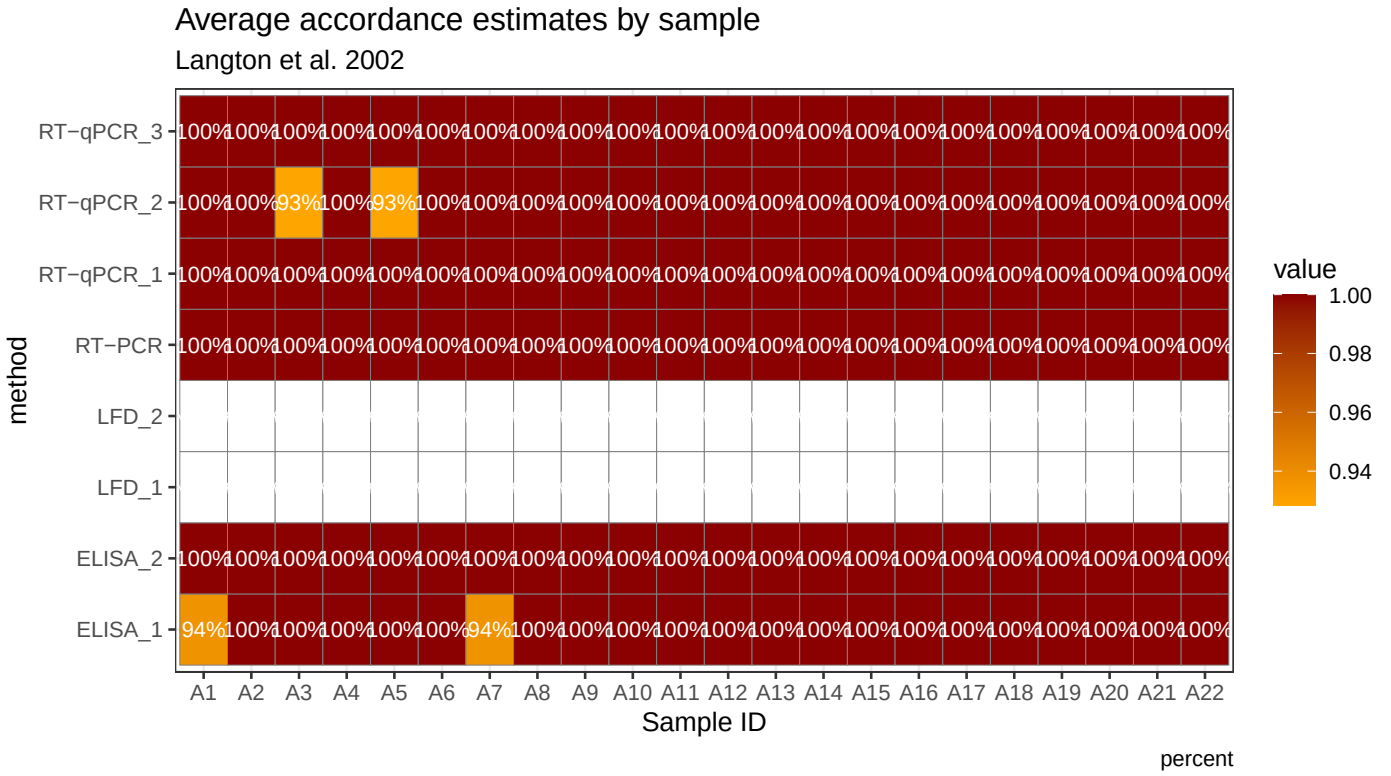
```
##      method
## lab   ELISA_1 ELISA_2 LFD_1 LFD_2 RT-PCR RT-qPCR_1 RT-qPCR_2 RT-qPCR_3
##   L01       0       0     0     0     44        44        44        44
##   L02       0       0     1     0      0         0         0         0
##   L03      44      44    21    22     44         0         0         0
##   L04      44      44    22    22     44        44        44        44
##   L05       0       0    22    22     44        44        44        44
##   L06      44      44     0     0      0         0         0         0
##   L07      44      44    22    22     44        44        44        44
##   L09      44      44    22    22     44         0        44         0
##   L10      44      44    22    22     44        44        44        44
##   L11      44      44     0     0     44        44        44        44
##   L12      44      44    22    22     44        44        44        44
##   L13      44       0     0     0      0        44        44        44
##   L14      44       0    22     0      0         0         0         0
##   L15      44      44     0     0      0        44        44        44
```

```
##    L17      44      44      22      22      44      44      44      44
##    L18      44      44      22      22       0       0       0       0
##    L19      44      44      22      22      44      44      44      44
##    L20      44      44      22      22      44      44      44      44
##    L21      44      44      22      22      44      44      44      44
```

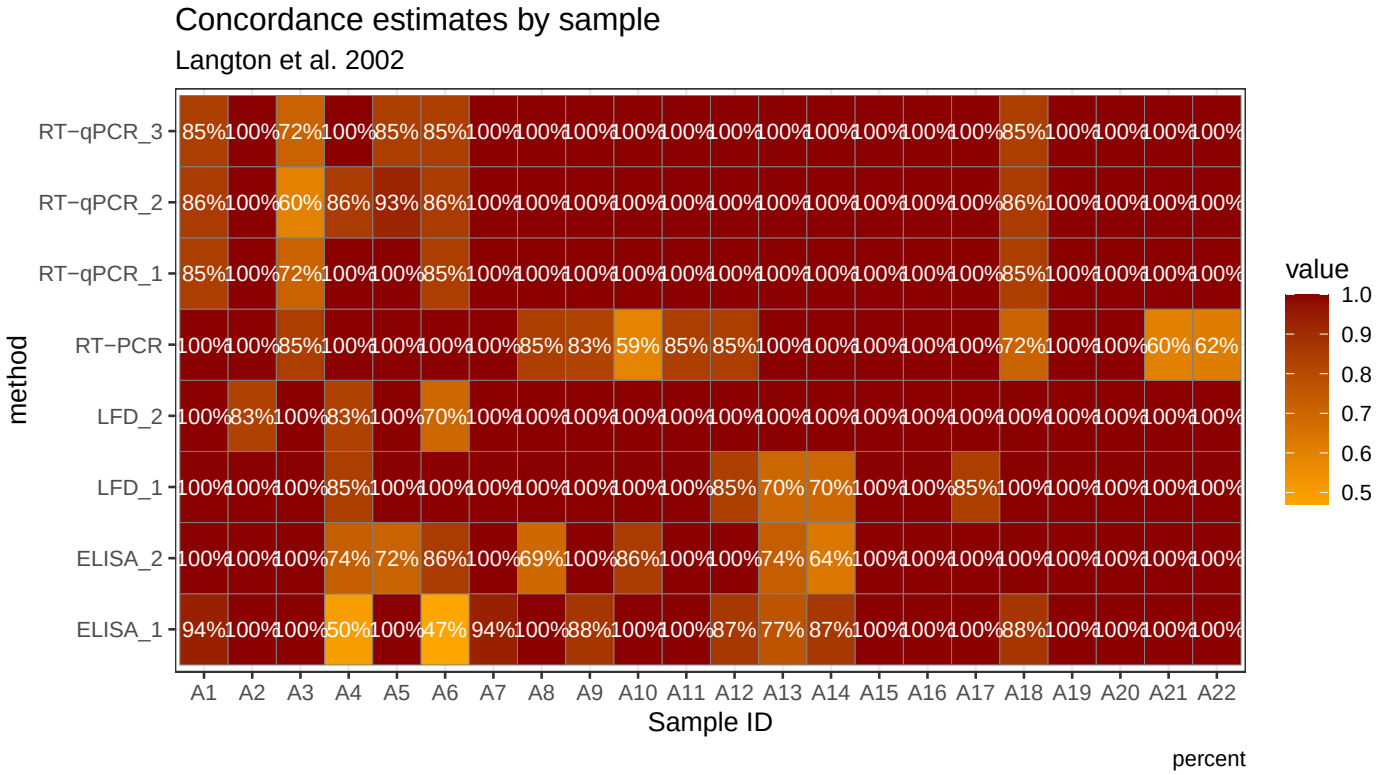
Repeatability and reproducibility

Accordance and concordance coefficients are used to estimate repeatability and reproducibility through laboratories, methods and samples. Unconclusive results are excluded from this analysis.

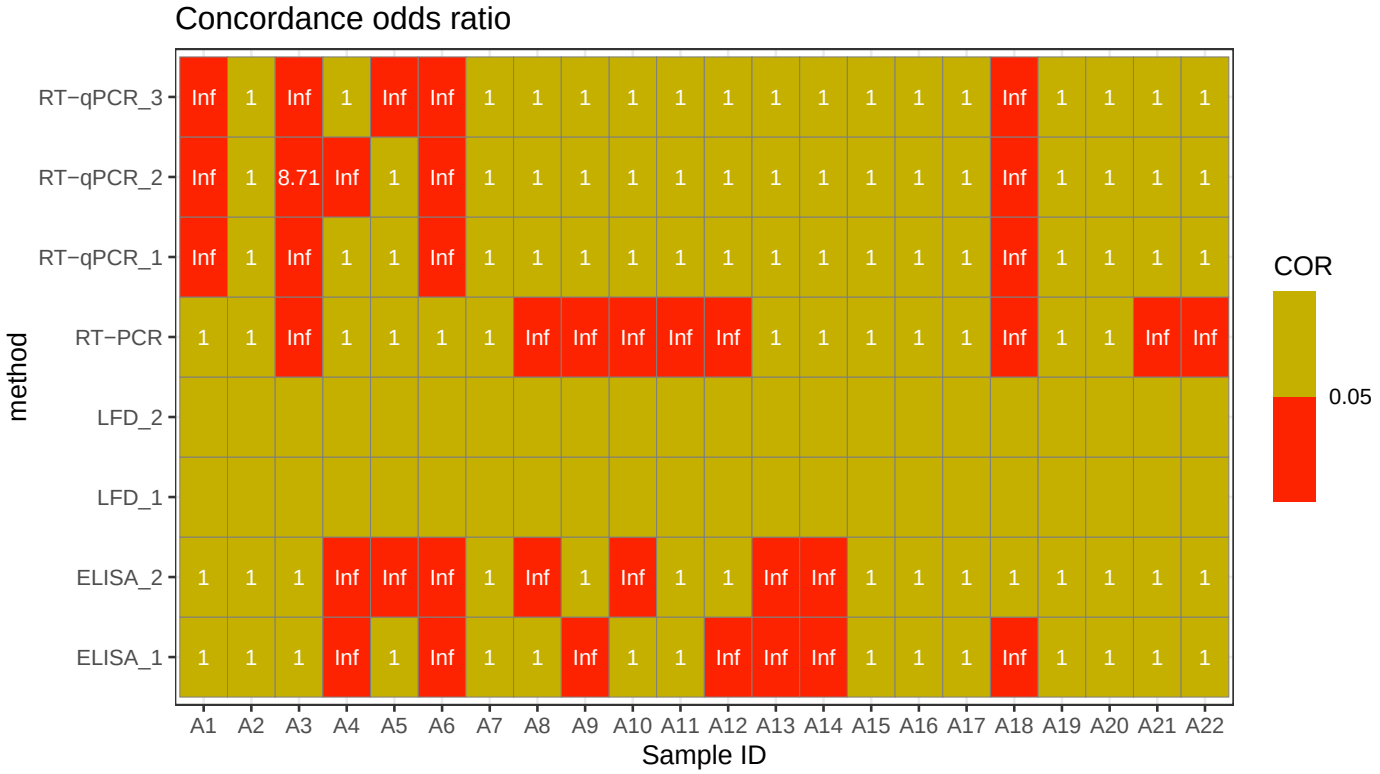
Accordance estimated values averaged by samples can give some insight about the high dependency of those value to the chosen sample.



Concordance estimates are equalily useful in this special case, because it highlights the samples which leads to high discrepancies between laboratories.



Concordance odds ratio (COR) can also assess the degree of interlaboratory variation ratio. COR removes the bias related to the accuracy of the results (i.e. numbers of true positive/negative and of false positive/negative) which are used to calculate the two parameters (i.e. concordance and accordance) taken separately. But because of the numerous 100% values, concordance odds ratios are of little help to discriminate all but the least efficient methods, as most of the estimates are either 1 or infinite values. This can be completed at the sample level by a Fisher's test, which test the hypothesis that there is a significant variation of the results between laboratories for a particular sample (the COR is significantly greater than one).



We also can characterize globally the repetability and reproducibility by calculating those parameters by method.

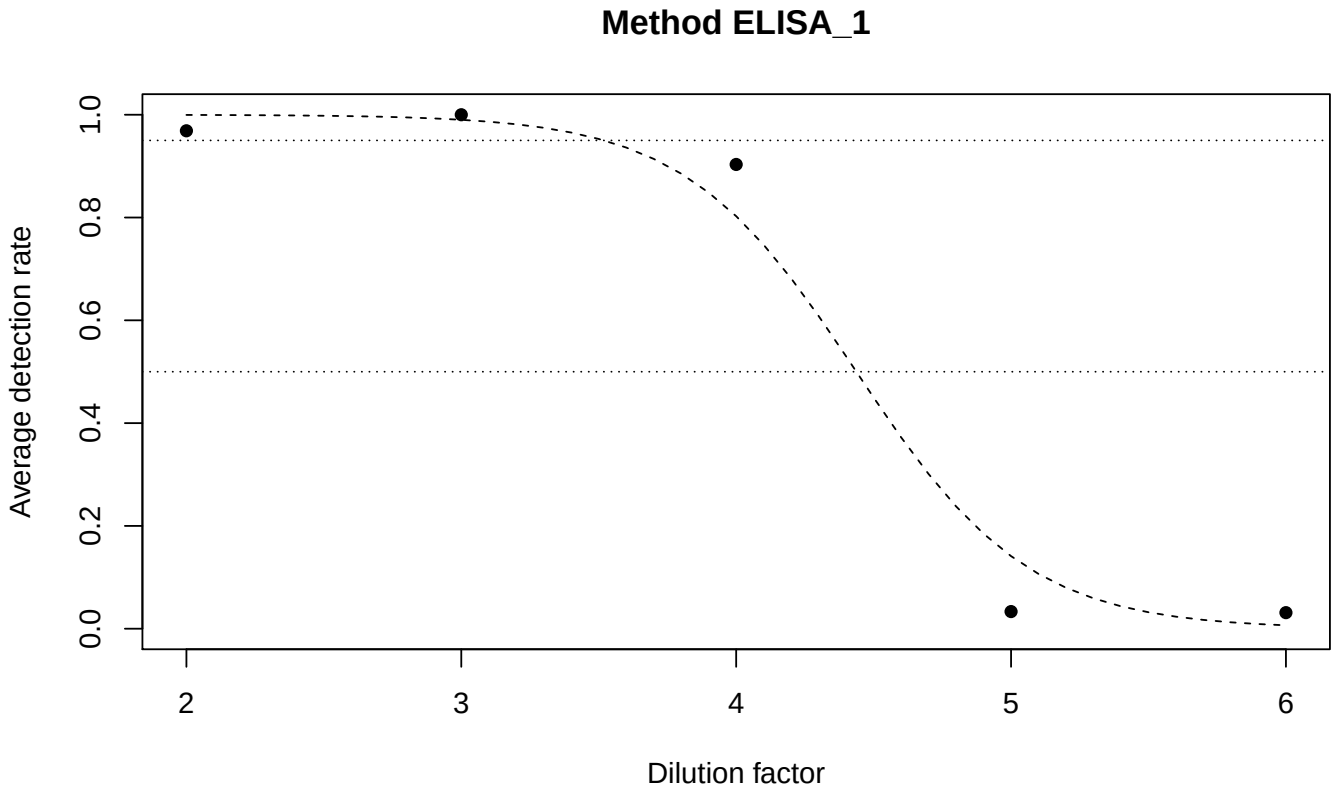
Table 1: Repeatabilty & Reproducibility by Method

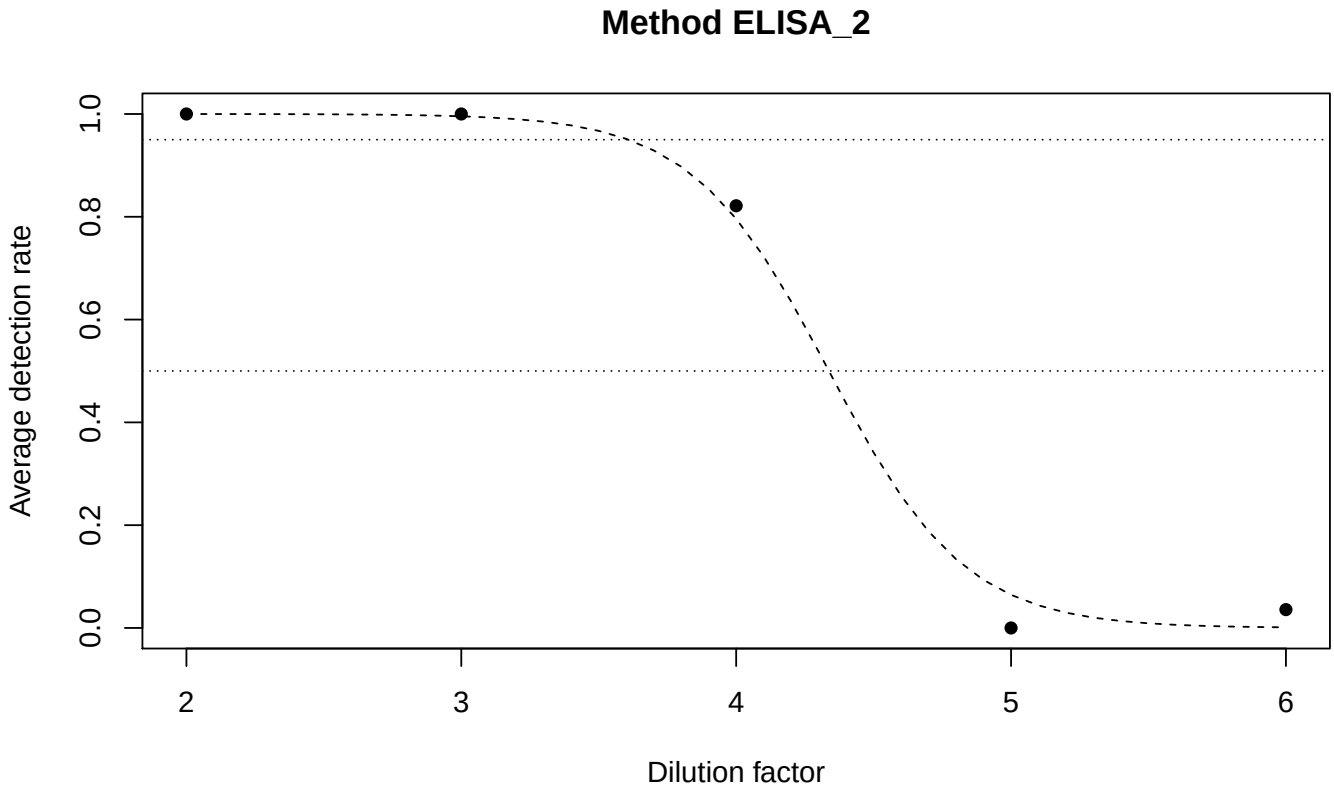
	ELISA_1	ELISA_2	LFD_1	LFD_2	RT-PCR	RT-qPCR_1	RT-qPCR_2	RT-qPCR_3
acc	0.994	1.00	NaN	NaN	1.000	1.000	0.994	1.000
conc	0.913	0.92	0.951	0.971	0.897	0.966	0.953	0.959
cor	16.600	Inf	NaN	NaN	Inf	Inf	7.622	Inf

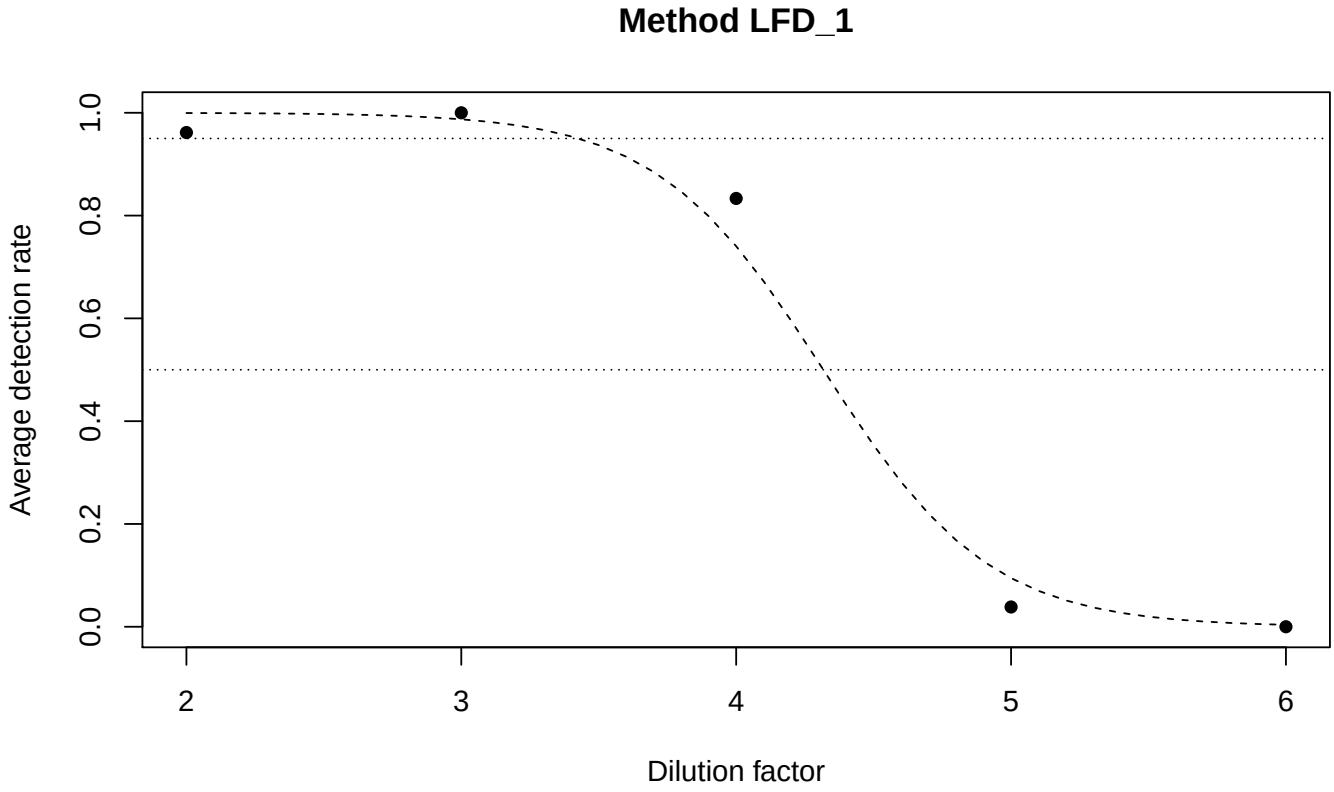
Analytical sensitivity

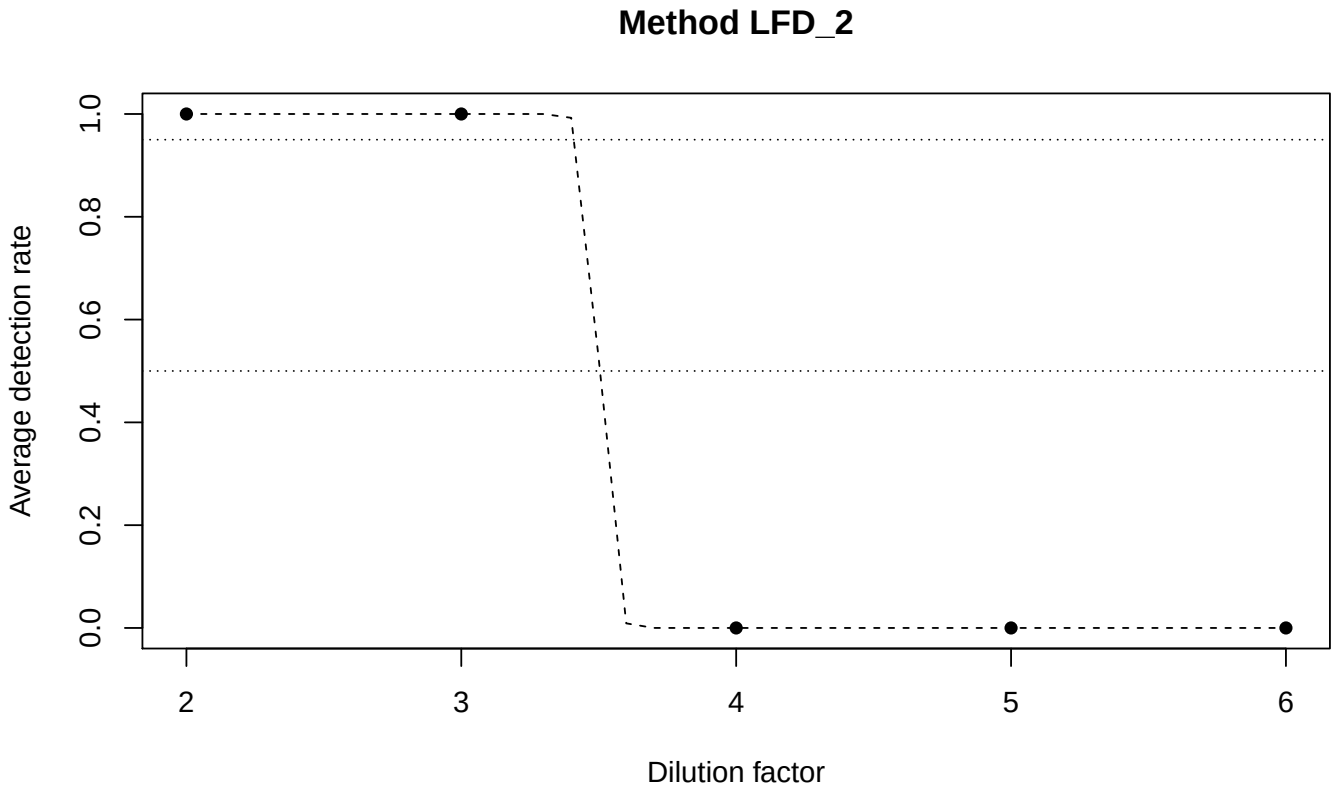
For each method , data of the diluted samples were used to adjust binomial generalized linear models (bGLM) with logit link between the dilution (expressed by the base 10 negative exponent of the corresponding dilution) and the detection status. The number of dilution level being very limited, the adjustement of bGLM is not always possible as this method require at least 5 levels, and the laboratory effect has been neglected.

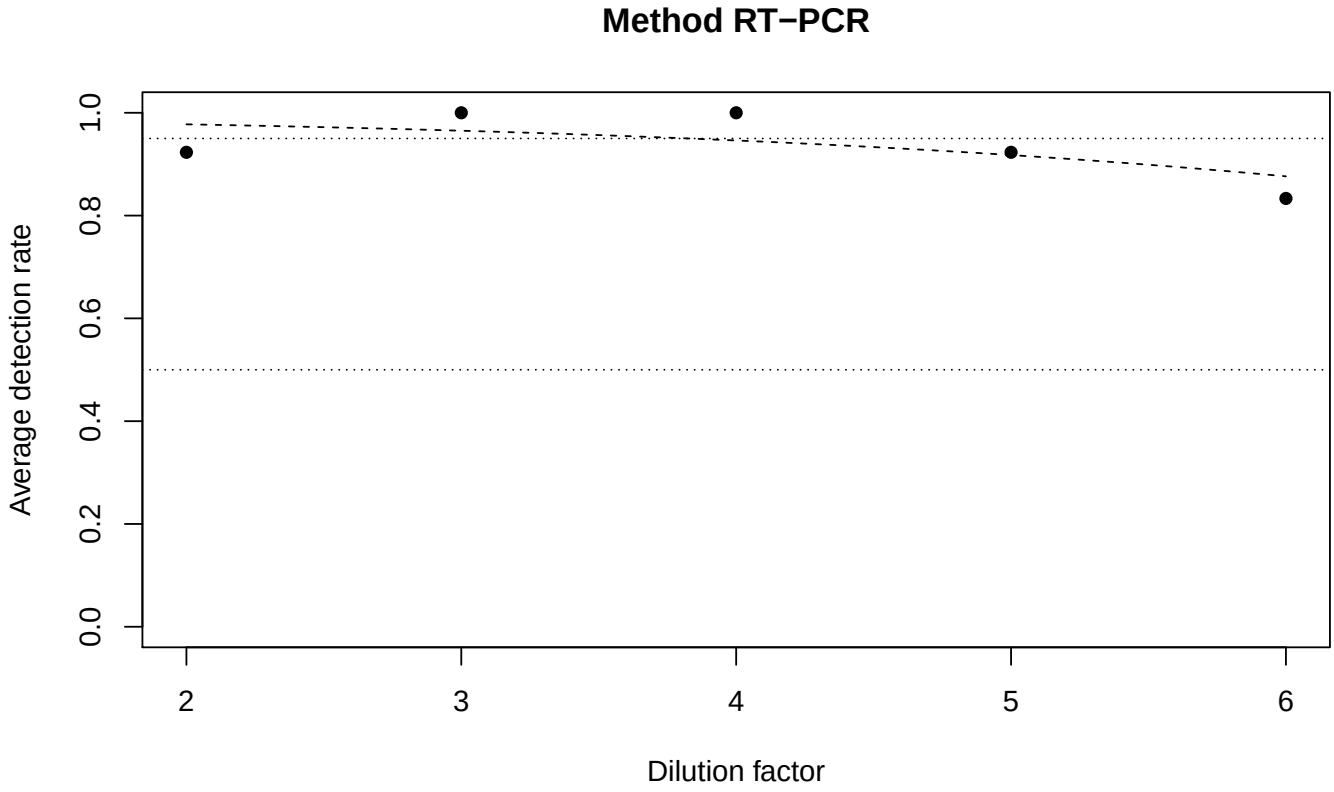
When possible, based on those models, dilutions corresponding to a 50% or 95% probability of detection have been calculated as an example of the possible LOD to report (LOD50 and LOD95).

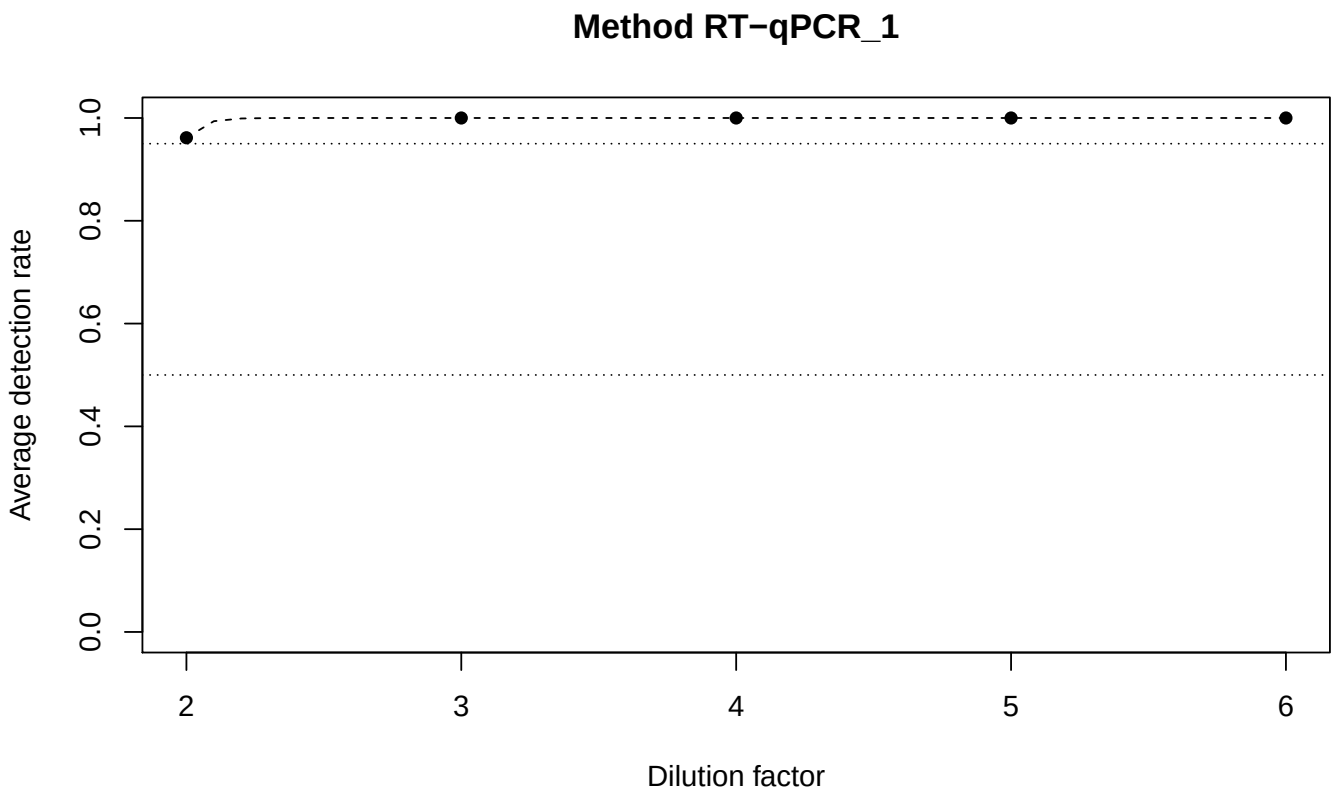


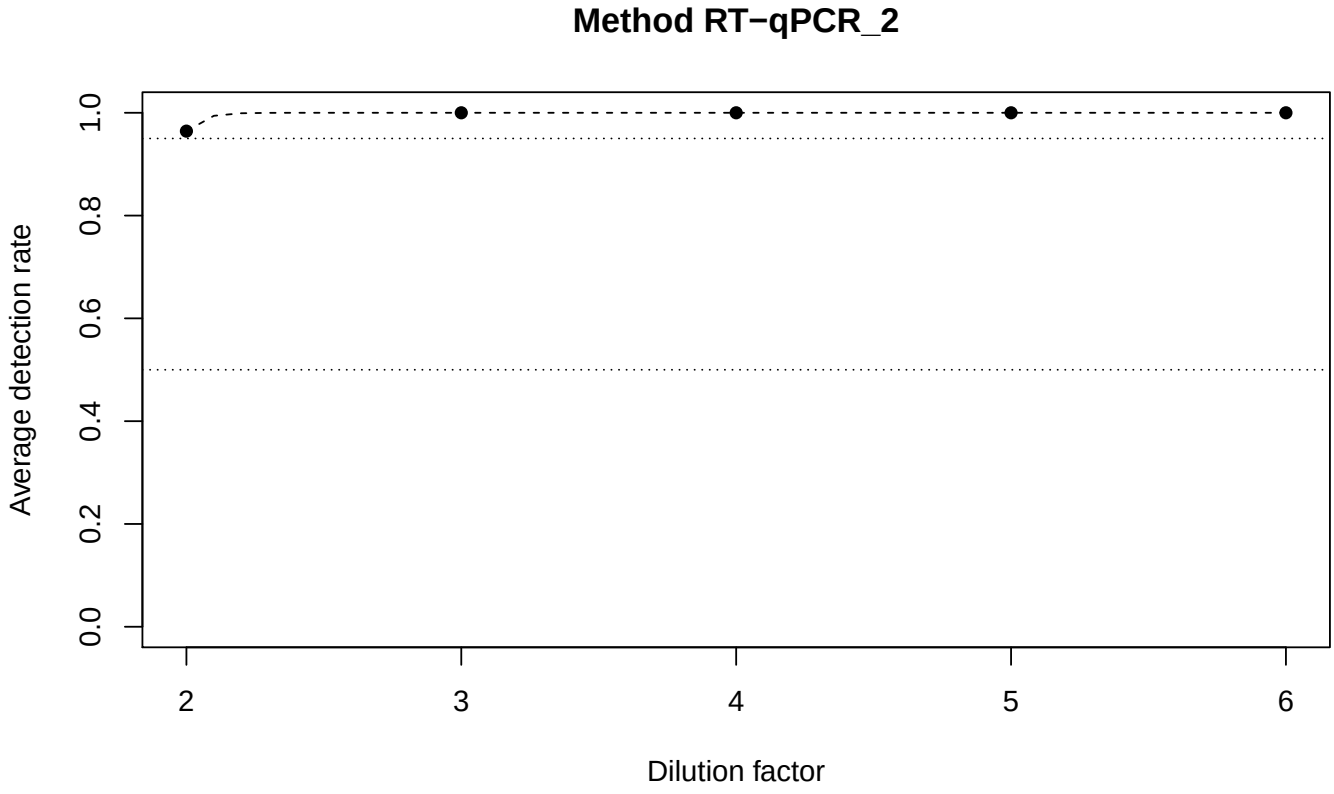

Method ELISA_2
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
2
3
4
5
6
Dilution factor


Method LFD_1
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
2
3
4
5
6
Dilution factor


Method LFD_2
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
2
3
4
5
6
Dilution factor


Method RT−PCR
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
2
3
4
5
6
Dilution factor


Method RT−qPCR_1
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
2
3
4
5
6
Dilution factor


Method RT−qPCR_2
Average detection rate
0.0
0.2
0.4
0.6
0.8
1.0
2
3
4
5
6
Dilution factor

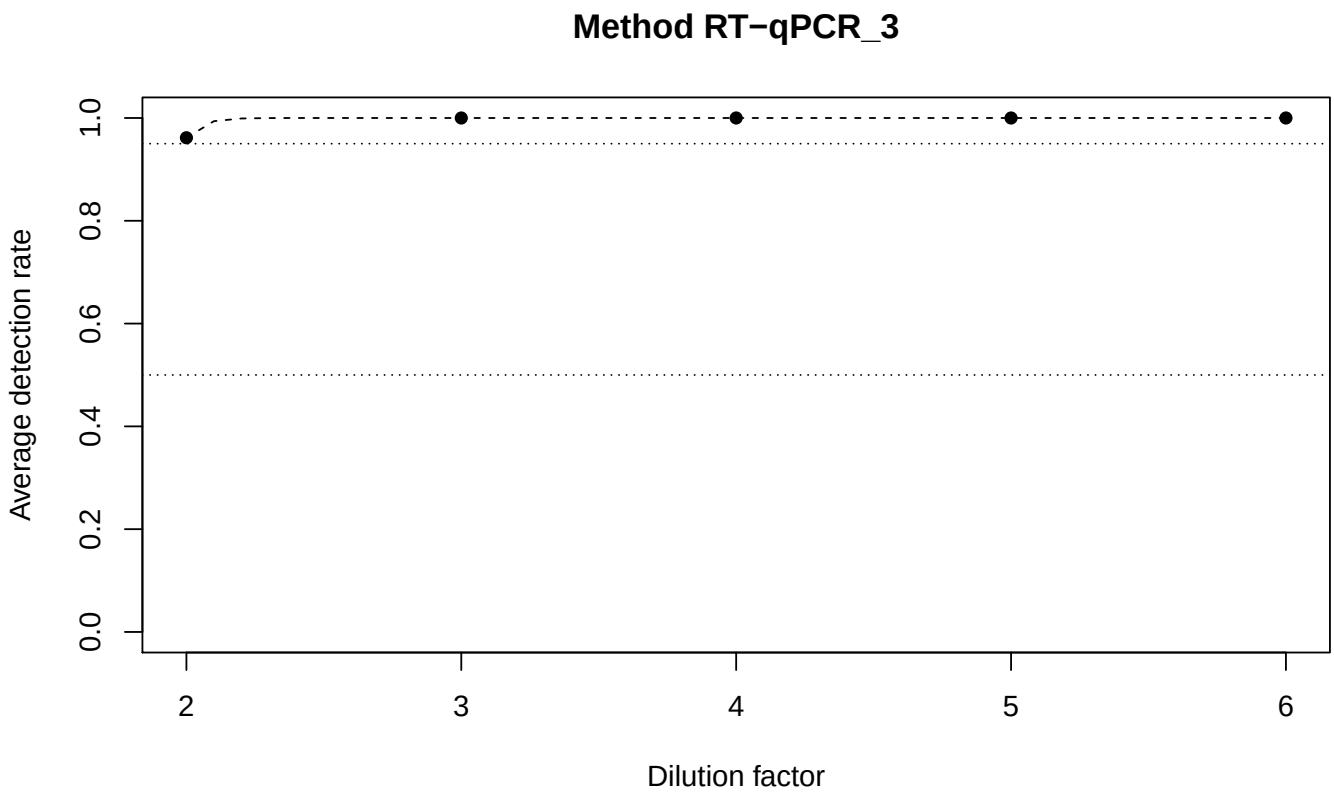


Table 2: Detection limits (log dilution factor) at 50% rate by methods

ELISA_1	ELISA_2	LFD_1	LFD_2	RT-PCR	RT-qPCR_1	RT-qPCR_2	RT-qPCR_3
4.4	4.3	4.3	3.5	NA	NA	NA	NA

Table 3: Detection limits (log dilution factor) at 95% rate by methods

ELISA_1	ELISA_2	LFD_1	LFD_2	RT-PCR	RT-qPCR_1	RT-qPCR_2	RT-qPCR_3
3.5	3.6	3.4	3.4	3.8	NA	NA	NA

Diagnostic sensitivity and specificity

Diagnostic sensitivity is estimated as the detection rate on samples with positive reference status, and the diagnostic specificity as the non detection rate on samples with negative reference status. Hence, those two parameters are heavily dependent on the choice of the reference positive and negative samples. Has the number of different samples vary from laboratories and methods, the comparison of the following estimates between methods has to be conducted with caution.

Table 4: Diagnostic sensitivity by methods

	ELISA_1	ELISA_2	LFD_1	LFD_2	RT-PCR	RT-qPCR_1	RT-qPCR_2	RT-qPCR_3
DSE	0.99	1.000	0.987	1.000	0.891	0.987	0.988	0.987
LCL	0.96	0.973	0.924	0.939	0.832	0.951	0.955	0.951
UCL	1.00	1.000	1.000	1.000	0.932	0.999	1.000	0.999

Table 5: Diagnostic specificity by methods

	ELISA_1	ELISA_2	LFD_1	LFD_2	RT-PCR	RT-qPCR_1	RT-qPCR_2	RT-qPCR_3
DSP	0.883	0.554	0.635	0.958	0.976	0.962	0.938	0.952
LCL	0.837	0.488	0.539	0.894	0.943	0.925	0.897	0.913
UCL	0.917	0.617	0.721	0.987	0.991	0.982	0.963	0.975

Likelihood ratios

Another drawback of a poor choice of reference samples is that 100% sensitivity or specificity estimates will impair the estimation of likelihood ratios, as they have either (1 - sensitivity) ou (1 - specificity) as denominator. Perfect specificity will then leads to infinite positive likelihood ratio, while perfect sensitivity leads to infinite negative likelihood ratio, obliterating the effect of the other parameter in the estimate. In this case, confidence limits can give some additional information by giving upper and lower bounds for this ratio, event in the case of an infinite estimation.

Table 6: Likelihood ratios by methods

	ELISA_1	ELISA_2	LFD_1	LFD_2	RT-PCR	RT-qPCR_1	RT-qPCR_2	RT-qPCR_3
LR+	8.444	2.240	2.702	24.000	37.067	25.667	15.810	20.533
LCL	6.093	1.953	2.136	9.777	16.163	13.329	9.669	11.446
UCL	11.879	2.612	3.540	61.272	86.553	50.337	26.292	37.514
LR-	84.750	Inf	49.500	Inf	8.956	75.000	78.750	74.250
LCL	23.729	24.755	9.134	18.919	5.820	21.111	22.123	20.899
UCL	308.718	Inf	281.047	Inf	14.118	272.998	286.723	270.283

Other diagnostic parameters

Table 7: Other diagnostic parameters by methods

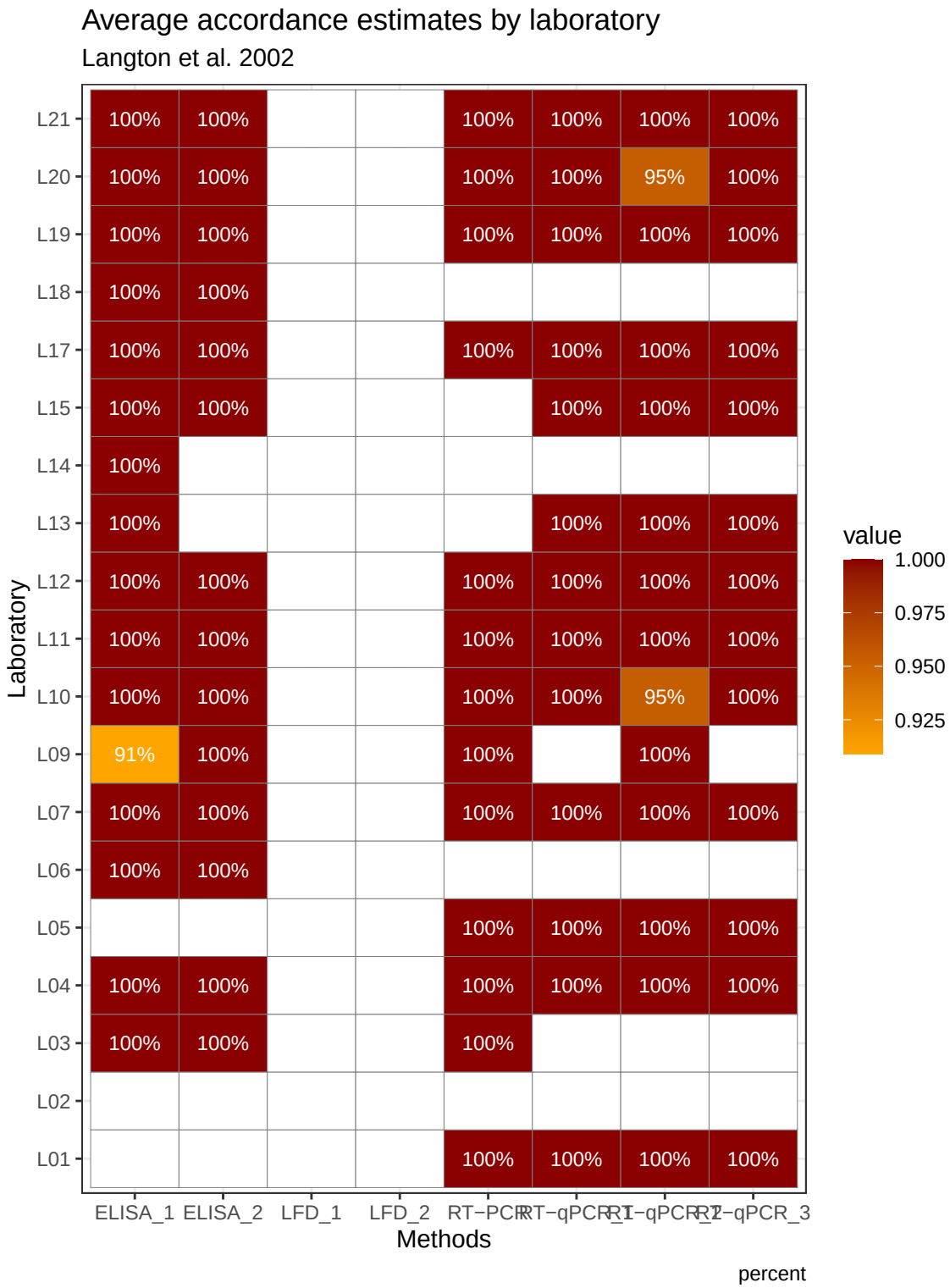
	ELISA_1	ELISA_2	LFD_1	LFD_2	RT-PCR	RT-qPCR_1	RT-qPCR_2	RT-qPCR_3
Accuracy	0.929	0.745	0.786	0.976	0.940	0.973	0.959	0.967
LCL	0.901	0.699	0.720	0.938	0.910	0.949	0.934	0.943
UCL	0.949	0.786	0.839	0.993	0.960	0.986	0.975	0.982
Power	0.848	0.627	0.658	0.947	0.781	0.928	0.902	0.917
LCL	0.795	0.568	0.568	0.868	0.714	0.877	0.850	0.864
UCL	0.890	0.683	0.738	0.983	0.836	0.959	0.938	0.951
Rate True Positive	0.864	0.627	0.670	0.947	0.965	0.951	0.922	0.939
LCL	0.812	0.568	0.579	0.868	0.919	0.904	0.873	0.890
UCL	0.903	0.683	0.749	0.983	0.987	0.976	0.954	0.968
Rate True Negative	0.991	1.000	0.985	1.000	0.923	0.990	0.991	0.990
LCL	0.966	0.964	0.912	0.952	0.879	0.962	0.964	0.962
UCL	1.000	1.000	1.000	1.000	0.952	1.000	1.000	1.000

Outliers detection

Outliers can be detected by decomposing the calculation of previous parameters to the laboratory and/or the sample level, and looking for strong individual deviation among them. Those deviations can be investigated by expert, leading to the possible exclusion of the corresponding results from the general analysis if necessary.

Repetability & reproducibility

Detailed accordance splitted by laboratories can help detecting labs which had some trouble with their repetability.



Analytical sensitivity

For each method and each laboratory, data of the diluted samples were used to adjust binomial generalized linear models (bGLM) with logit link between the target concentration (expressed in log scale) and the detection status.

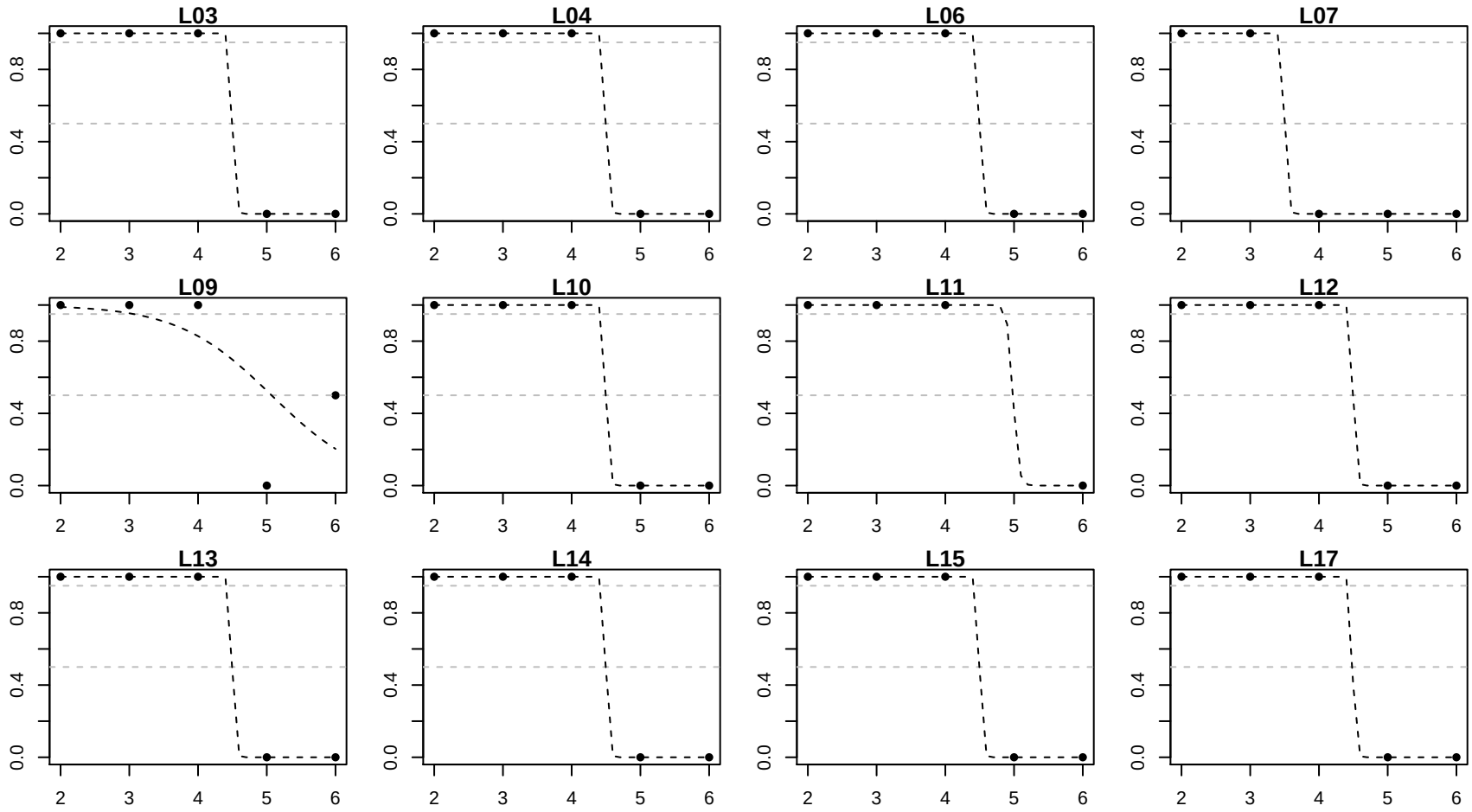
From those results, some warnings must be addressed when using bGLM to calculate LOD. Those models are based on the acceptable assumption that the probability of detection is decreasing when the dilution level increases. At the laboratory level, this hypothesis can be violated in different cases :

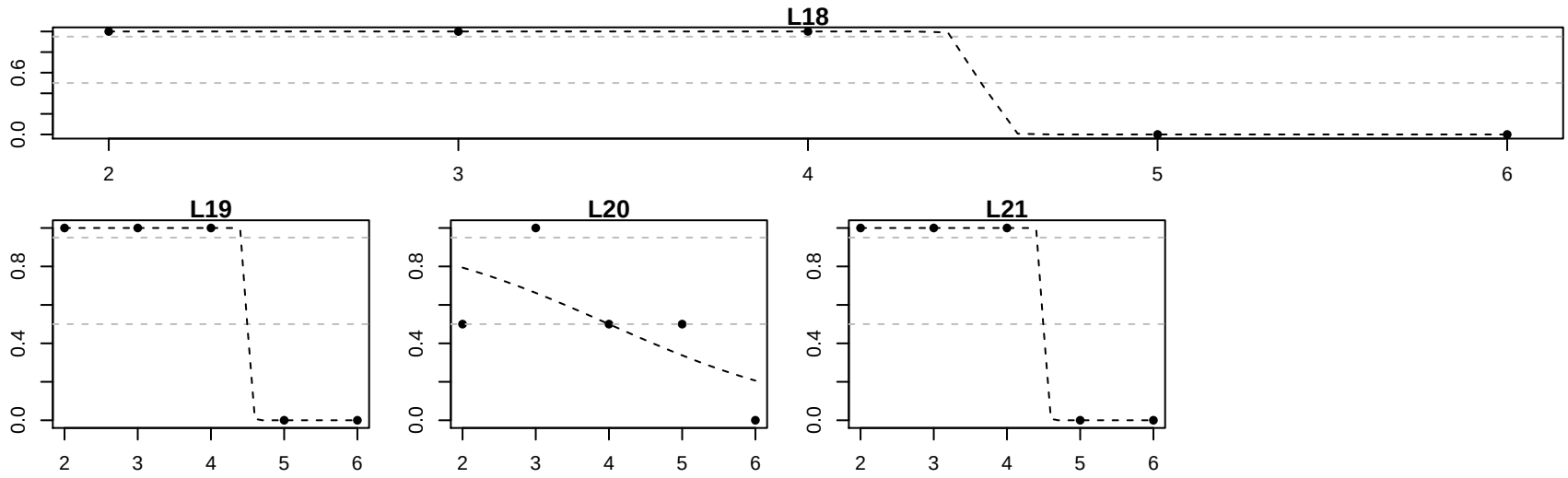
- when all the samples show the same status whatever the dilution level;
- when the observed detection rate shows non monotonous behaviour (e.g. decrease then increase again).

When this happens, the calculated LODxx calculated from the model cannot be trusted. Hence, expert verification of the fit of the model is mandatory before interpreating the calculated LOD value.

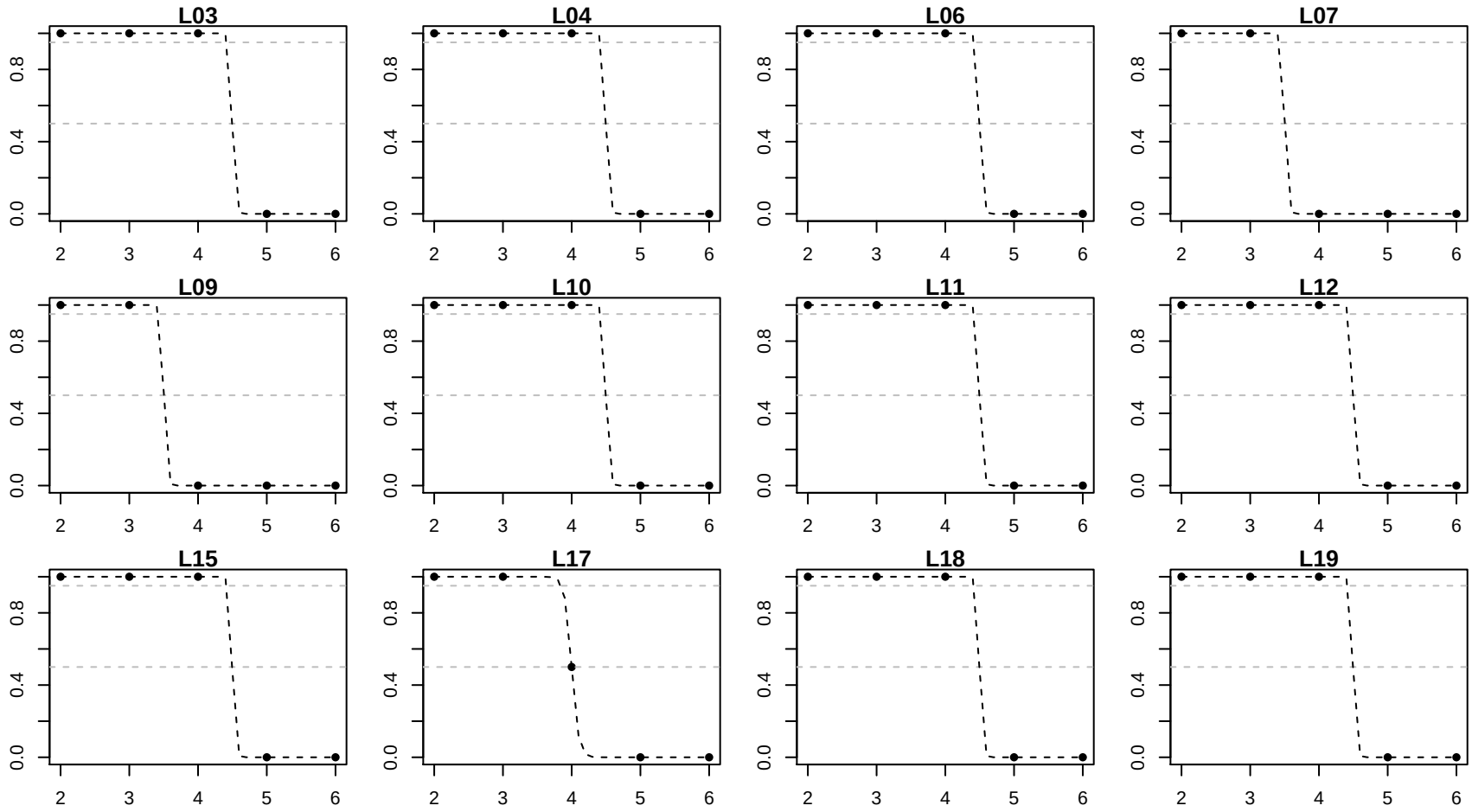
On the following figures, observed status and adjusted bGLM are shown for each method and laboratories. The problematic cases are very easy to identify on such graphical representation.

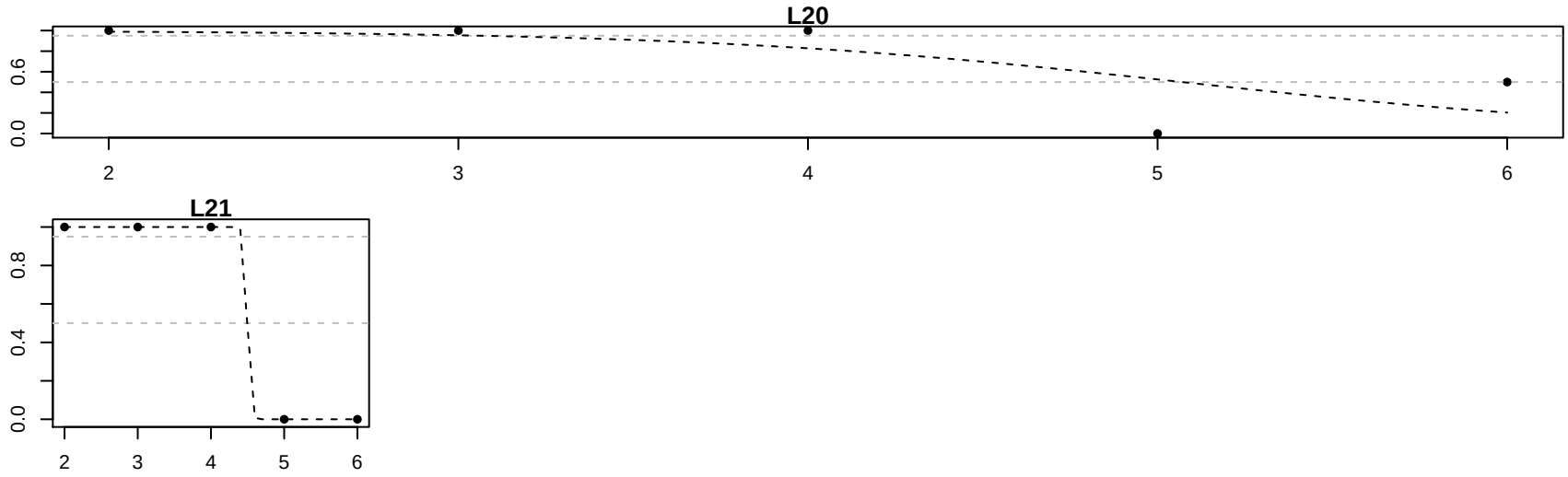
Method ELISA_1



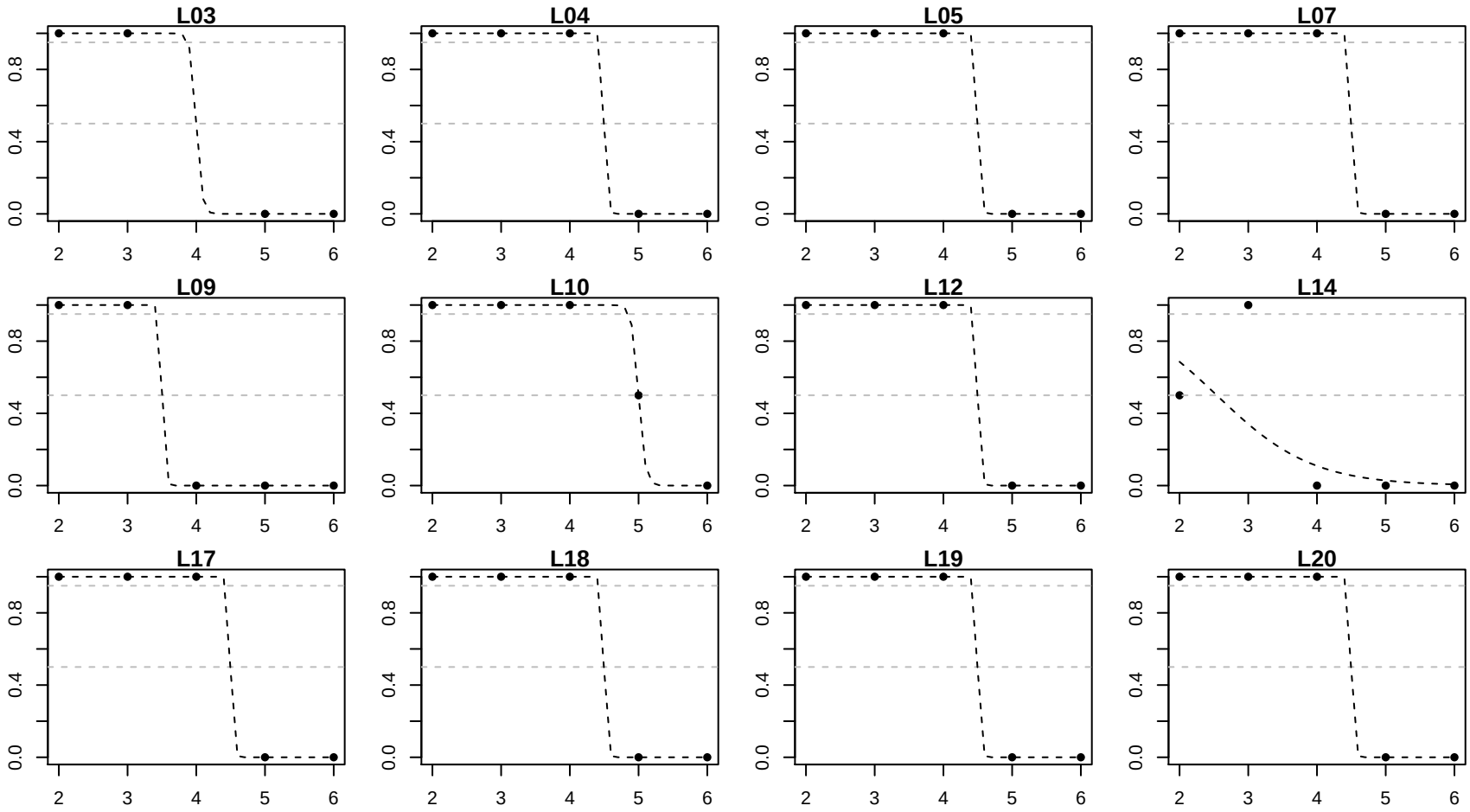

L18
0.0
0.6
2
3
4
5
6
L19
0.0
0.4
0.8
2
3
4
5
6
L20
0.0
0.4
0.8
2
3
4
5
6
L21
0.0
0.4
0.8
2
3
4
5
6

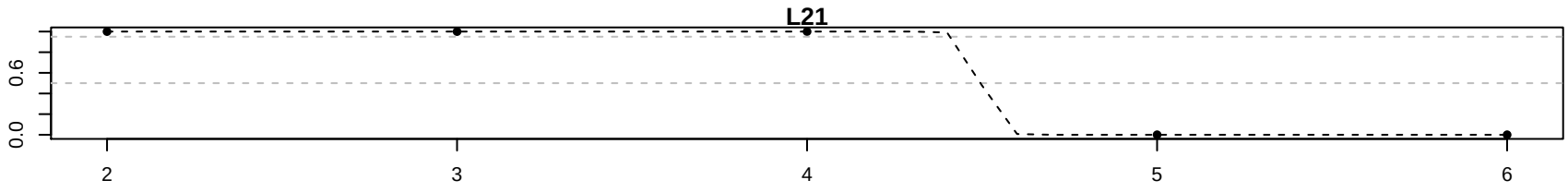
Method ELISA_2



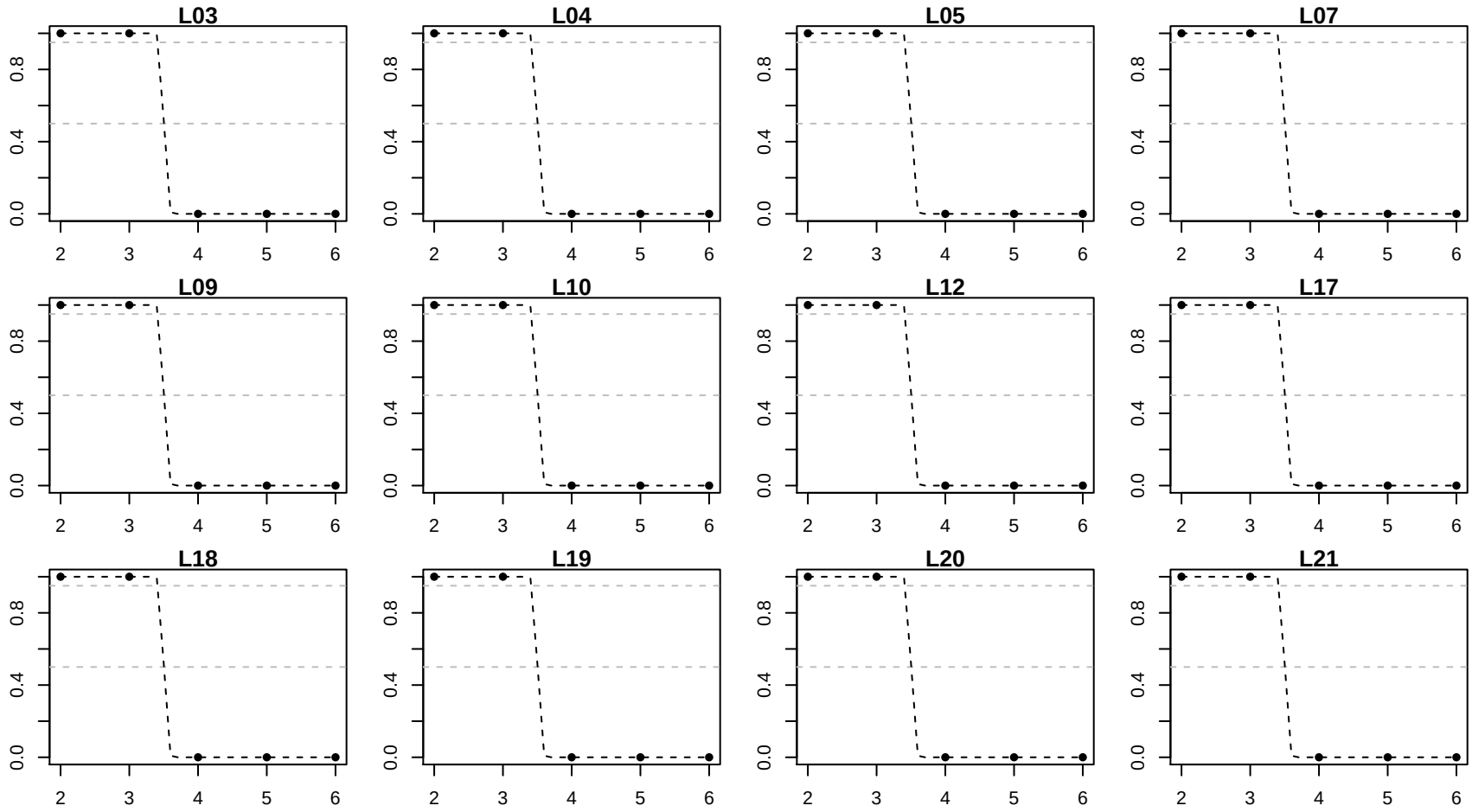

L20
0.0
0.6
2
3
4
5
6
L21
0.0
0.4
0.8
2
3
4
5
6

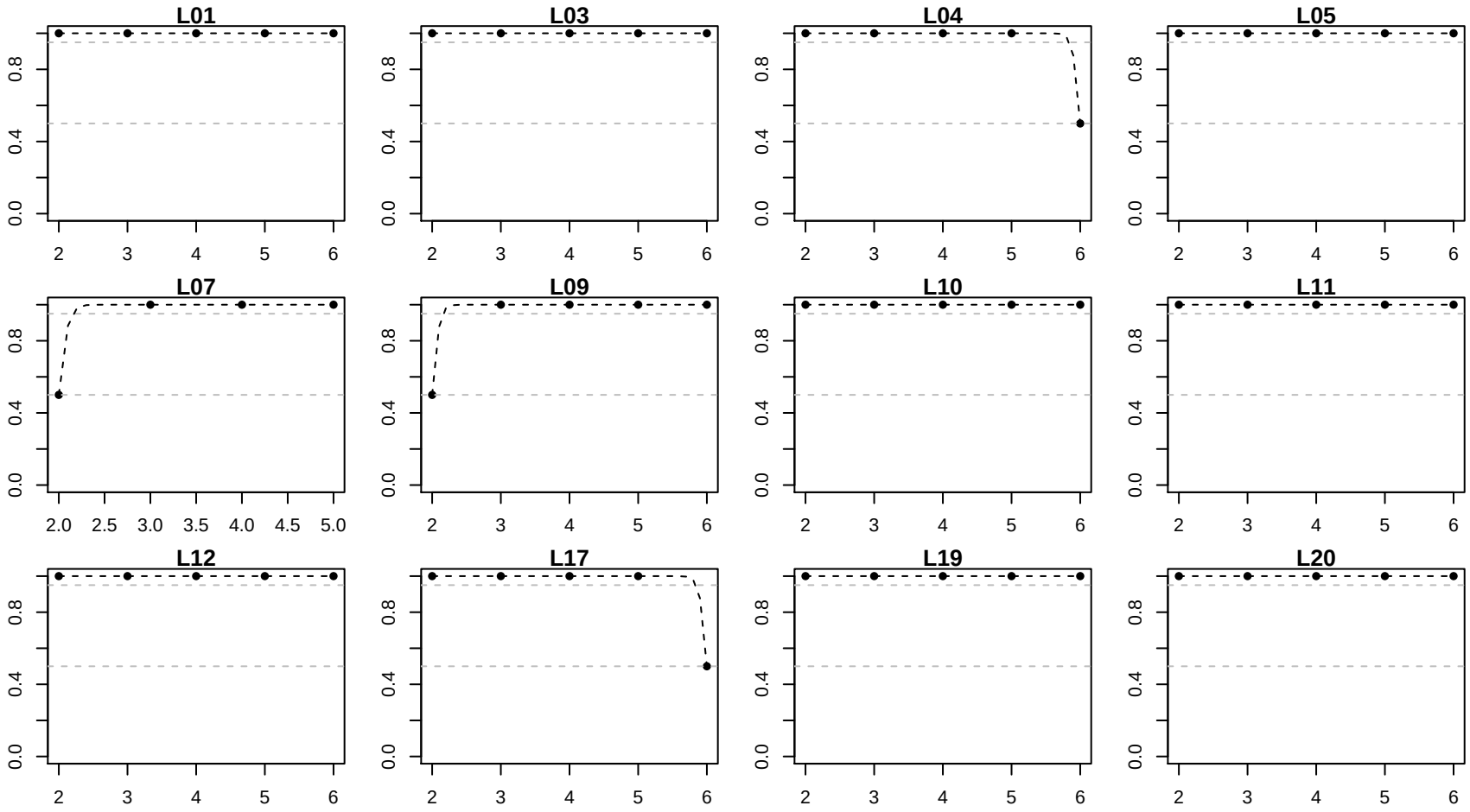
Method LFD_1

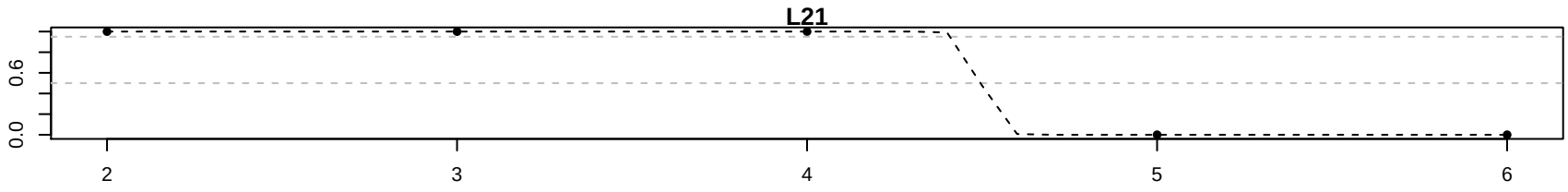



L21
0.0
0.6
2
3
4
5
6

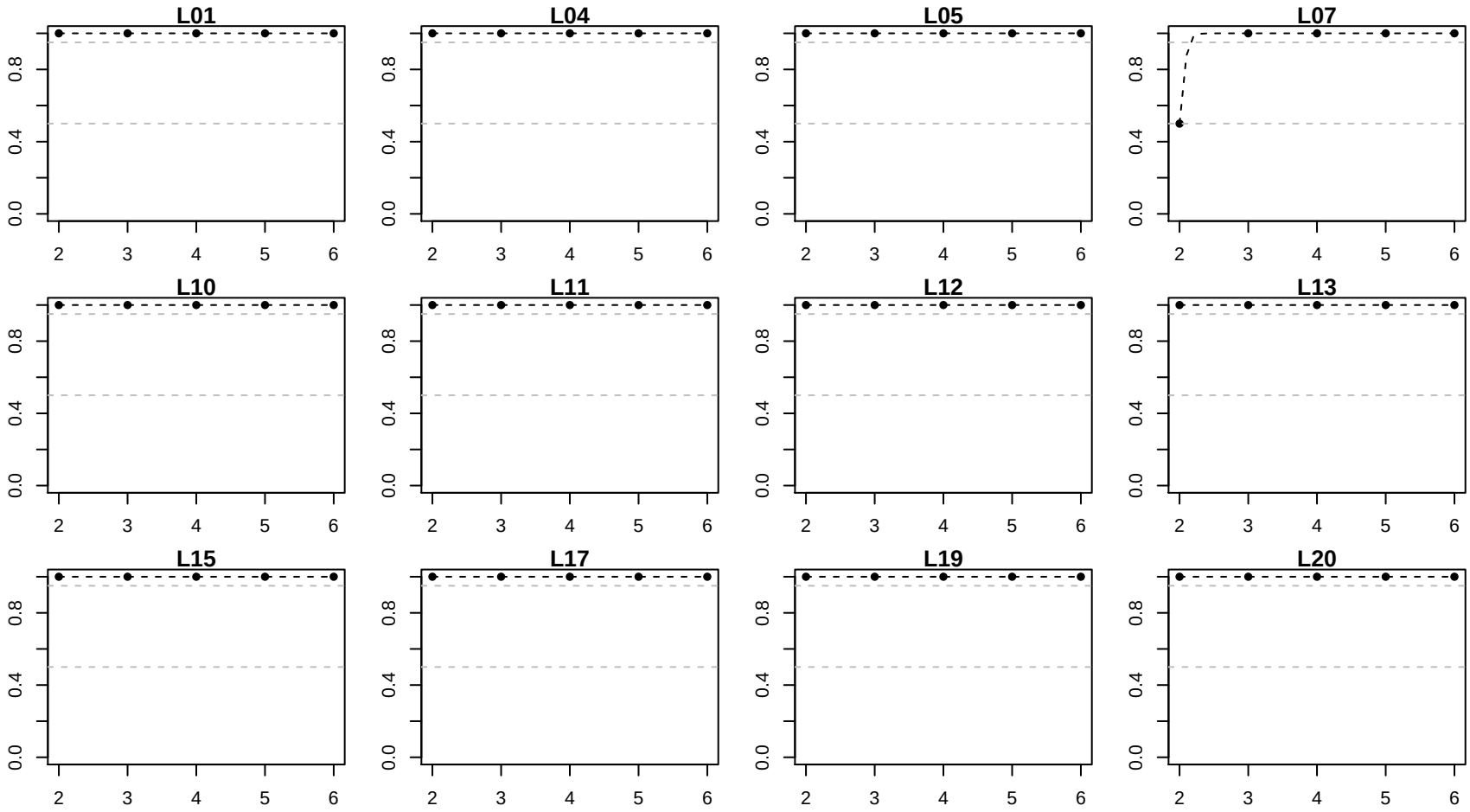
Method LFD_2

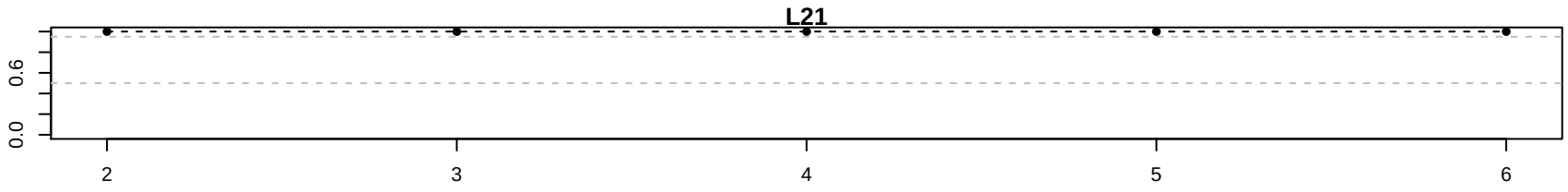



Method RT−PCR
L01
L03
L04
L05
L07
L09
L10
L11
L12
L17
L19
L20

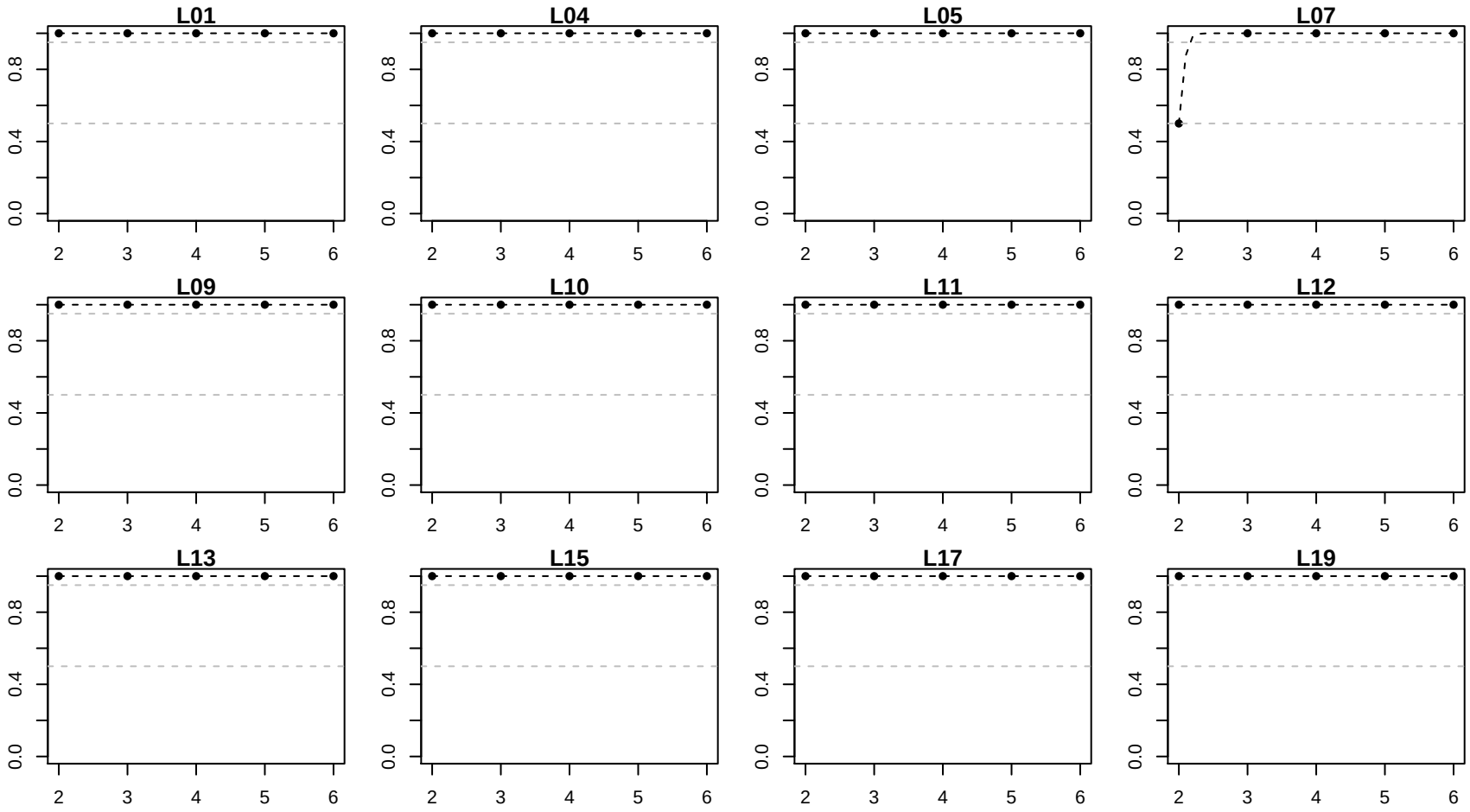

L21
0.0
0.6
2
3
4
5
6

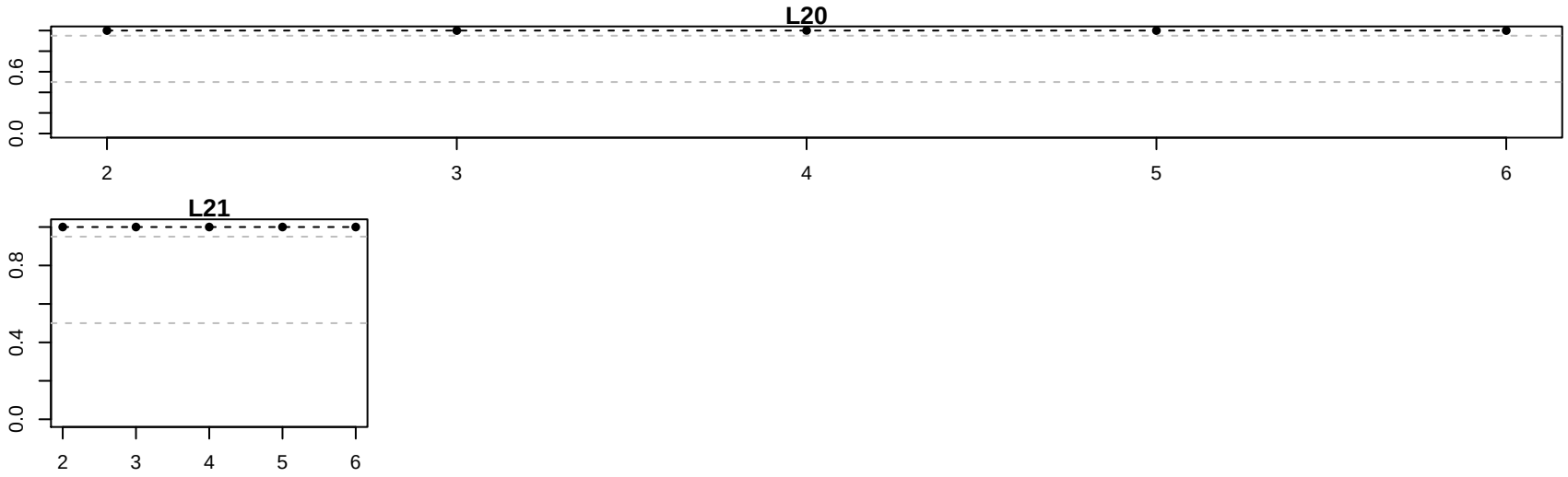
Method RT−qPCR_1



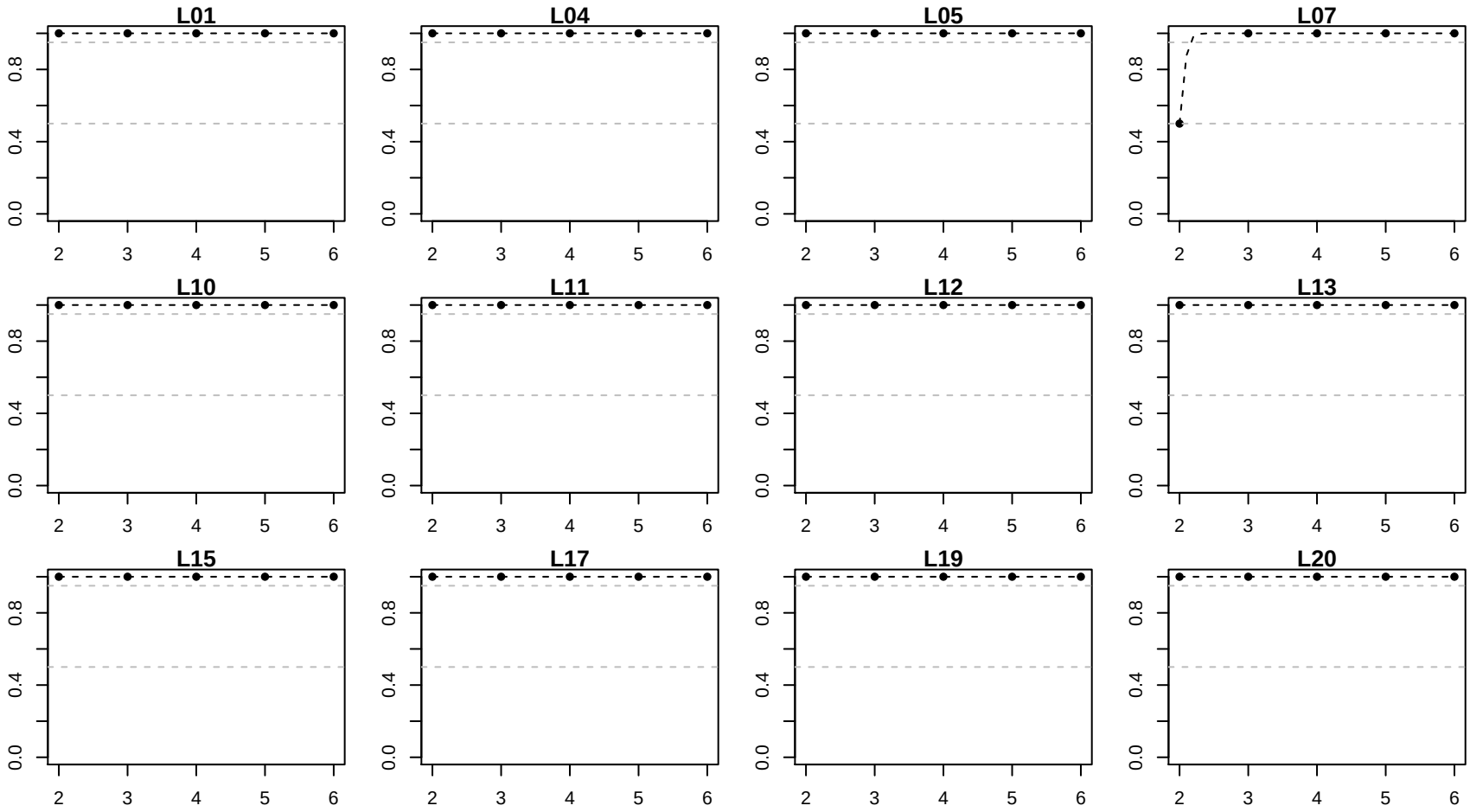

L21
0.0
0.6
2
3
4
5
6

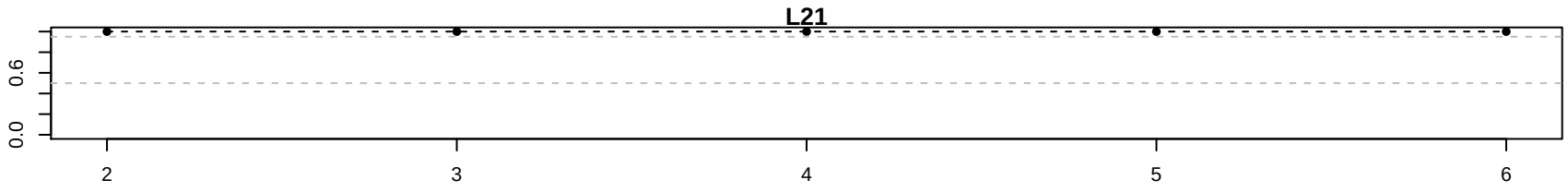
Method RT−qPCR_2




L20
0.0
0.6
2
3
4
5
6
L21
0.0
0.4
0.8
2
3
4
5
6

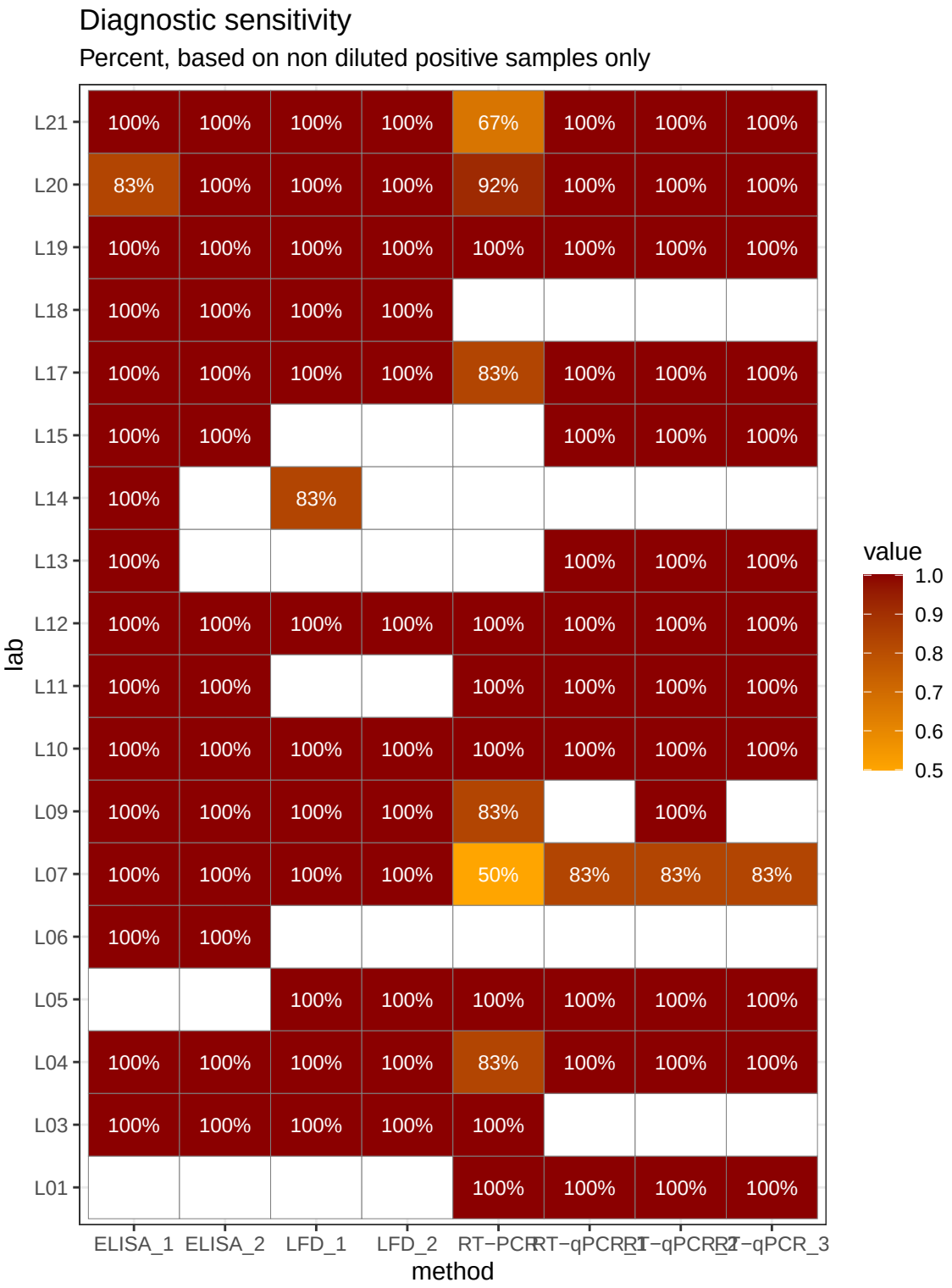
Method RT−qPCR_3




L21
0.0
0.6
2
3
4
5
6

Diagnostic sensitivity and diagnostic specificity

Even in the case of very high diagnostic sensitivity or specificity, those parameters can still be usefull to detect outliers, which show below than average performance event in those situations.



Negative samples with very low confusing sequences will lead to the same results for specificity (very frequent 100% estimation), but can also detect some poor performance of labs, methods or a combination of the two.

Diagnostic specificity

Percent, based on non diluted samples only