TESTA WP2: Sampling

Roy Macarthur[1], Valerie Cockerell[2], Mauro Dal Pra[3], Ilaria Alberti[3], Maaike Bruinsma[4], Harrie Koenraadt[4], Jane Thomas[5]

[1]Fera,[2]SASA, [3]UNIMORE [4]Naktuinbouw, [5]NIAB

Testa WP2: Sampling

Part 1

- Goal of WP2
- A description of lots, sampling and testing
- Studies to provide estimates of the things that effect sampling

Part 2

- How to produce plans to meet goals for detection
- Some plans
- QA for plans

Goal of WP2

- Estimate the size of sources of variation that effect sampling and testing
- Produce sampling plans for the detection of quarantine pests and pathogens in seed lots
- Plans have associated reliable estimates for Limit of Detection for pest and pathogens in lots

What is sampling for?

- Testa is focussing on quarantine pathogens.
- Any finding leads to rejection
- The pest or pathogen should not be present at any level
- We want sampling and testing to provide high confidence that if a pest or pathogen is present at all it must be at a low level.

What is sampling for?

- Testa is focussing on quarantine pathogens.
- Any finding leads to rejection
- We want sampling and testing to provide high confidence that if a pest or pathogen is present at all it must be at a low level.
- We want sampling and test plans with a *reliably known* low limit of detection for pests and pathogens.

Designing plans is easy if…

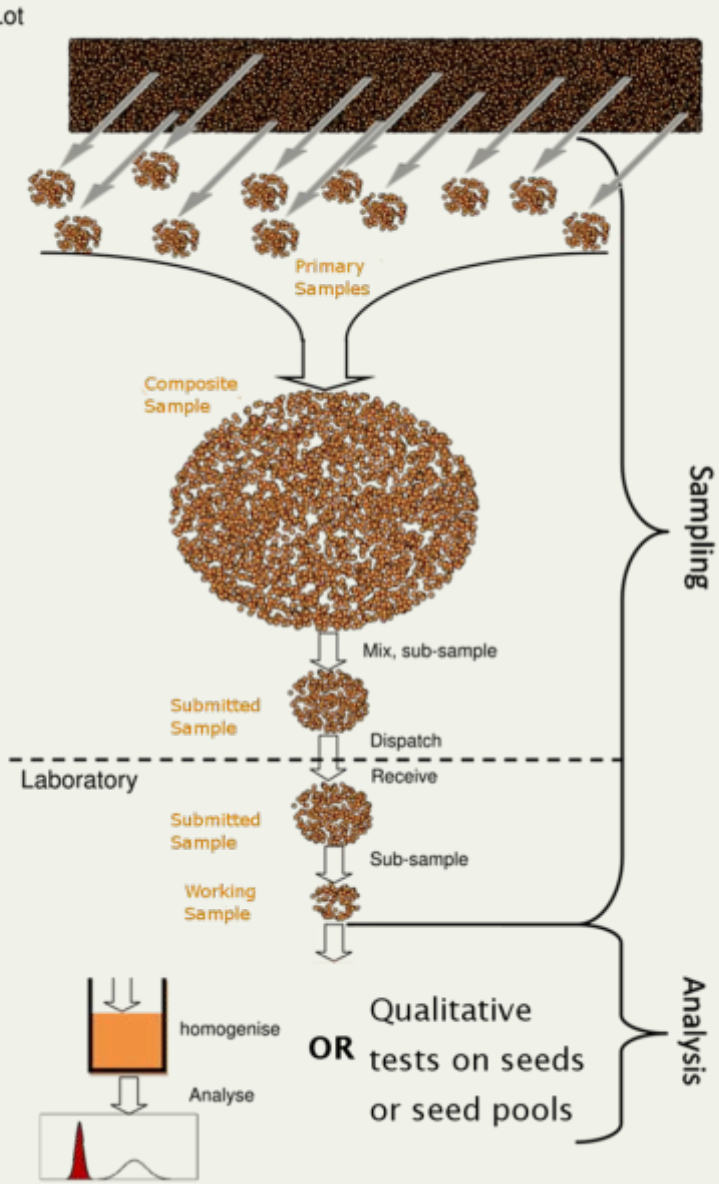
- Assume that seed lots are homogenous.
- Assume that the detection method is reliable

- There is a well known simple relation between sample size and limit of detection and confidence in the limit (eg ISPM31)

Level of pathogen may vary in different parts of the lot

We may want to detect low levels

We may have imperfect detection

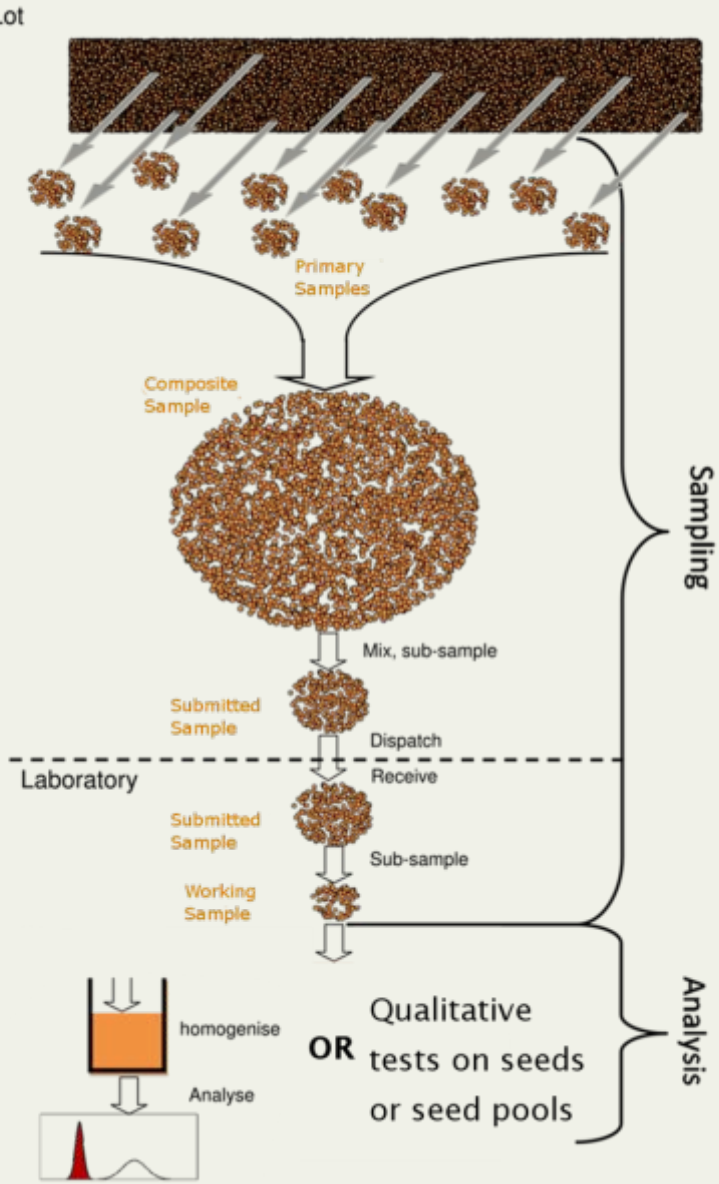


Level of pathogen may vary in different parts of the lot

We may want to detect low levels

Level of pathogen may vary "between seeds"

We may have imperfect detection



Level of pathogen may vary in different parts of the lot

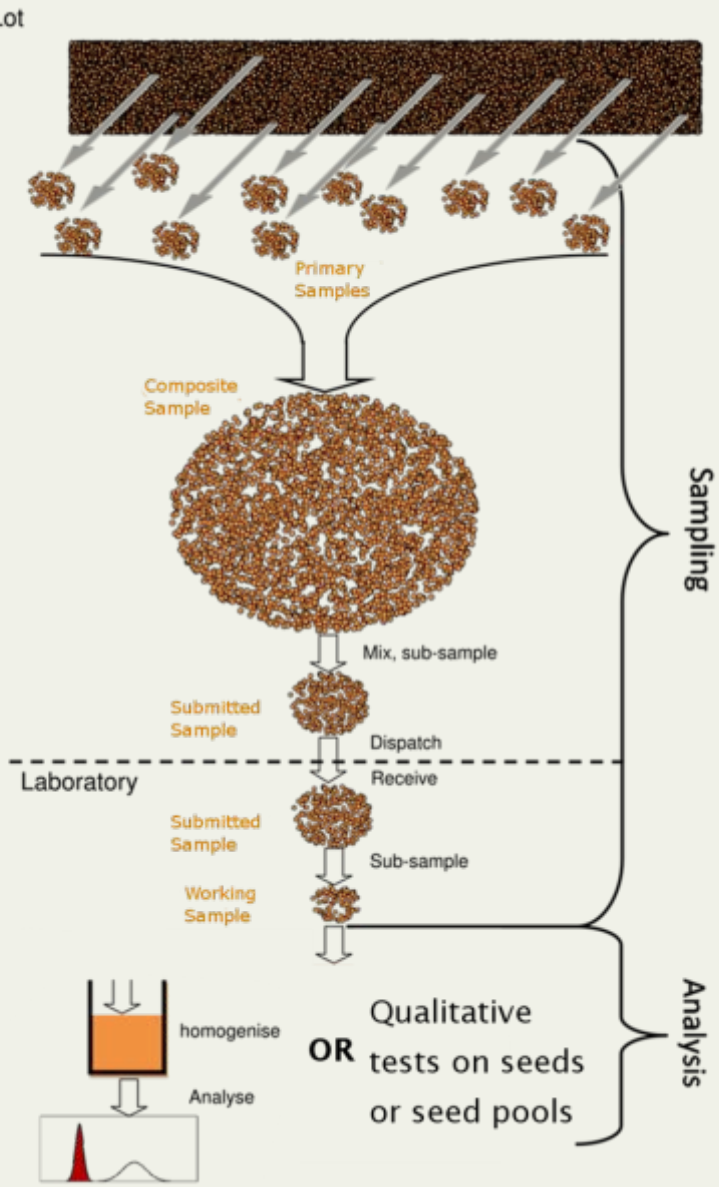
Take enough primary samples

We may want to detect low levels

Make the working sample large enough

We may have imperfect detection

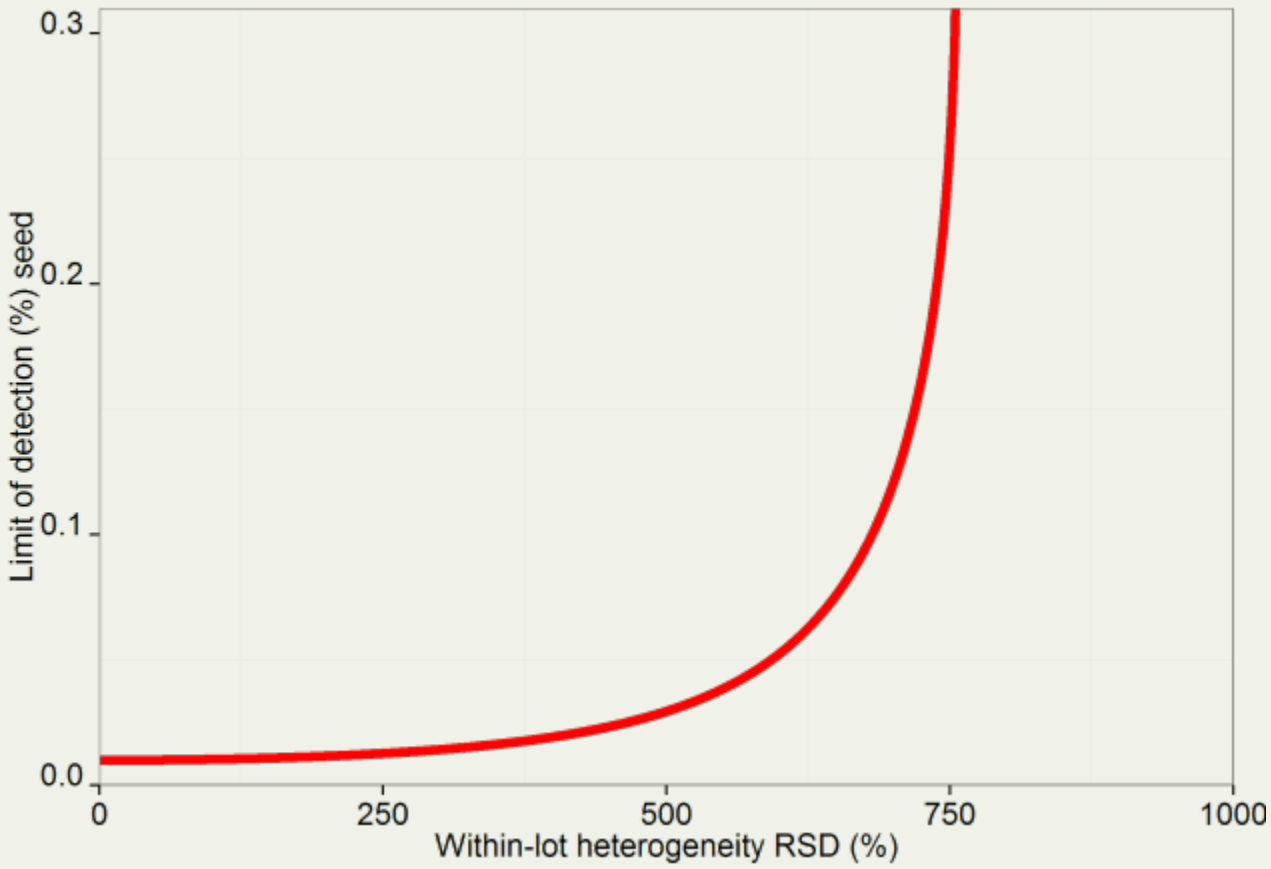
Modify method / replicate testing

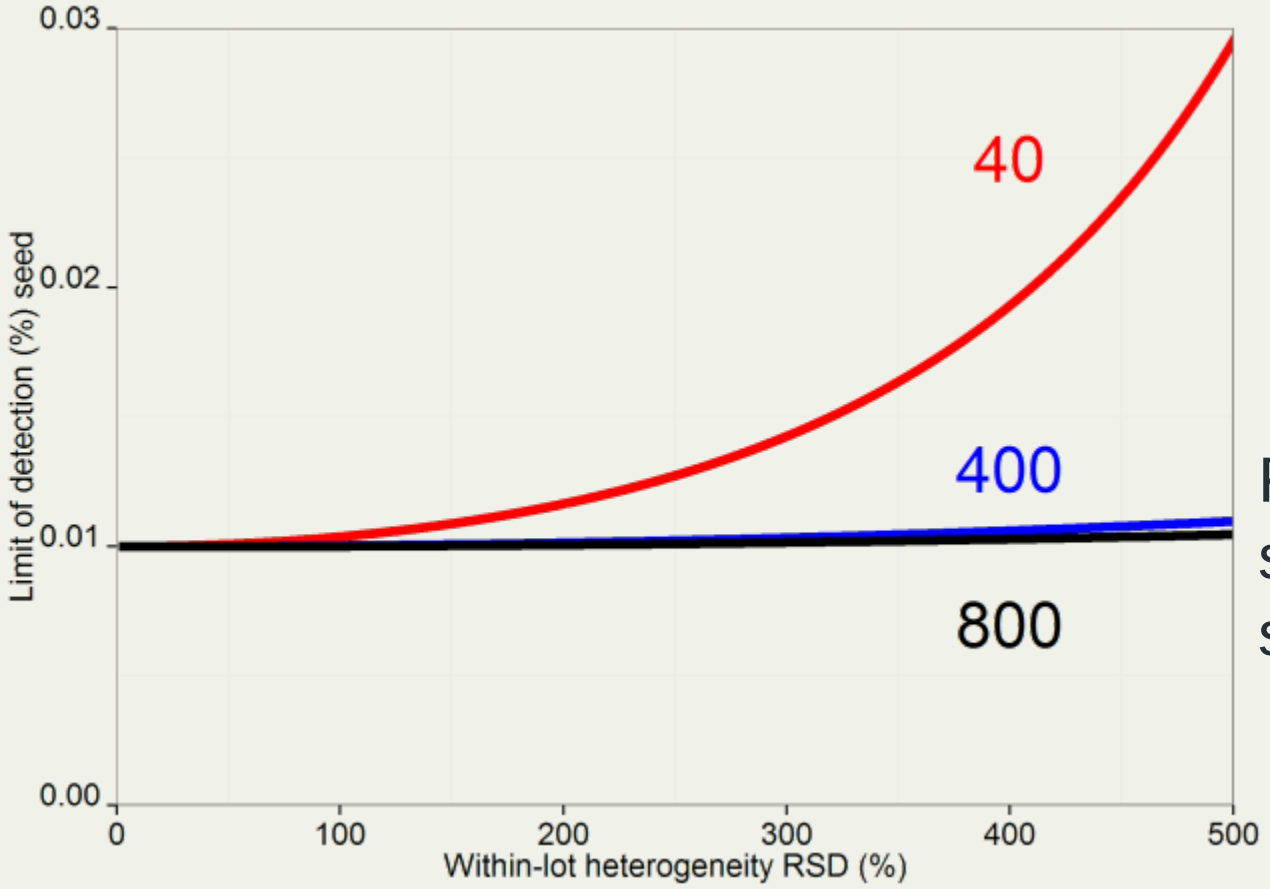


Factors the effect the sampling plan

FACTOR	SAMPLING PLAN
Variation in different parts of lot	Number of primary samples
Variation at small scale (seeds)	Size of working sample
Average level of pathogen	Size of analytical aliqout
LOD of analytical method	Size of working sample / Size of analytical aliquot / Replication

Effect of primary sample variation





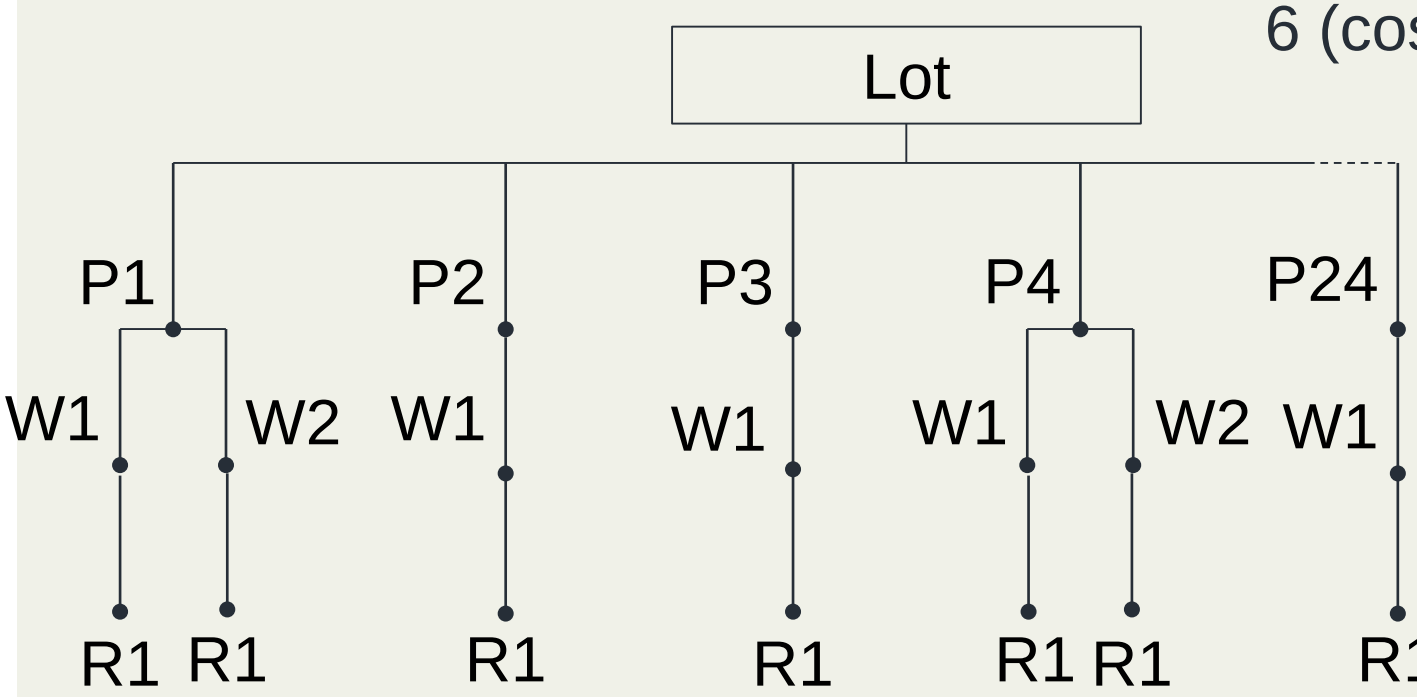
For 30 000 seed working sample

Testa studies on lots

- Four case studies.
- Fusarium species in wheat
- Ditylenchus sp. In field bean
- Xcc in brassica
- Tilletia in wheat
- Six lots per study

- Other scenarios using available information

Lot study design



6 (cost driven)

24 primary samples

1 or 2 working samples per prima sample

36 analyses

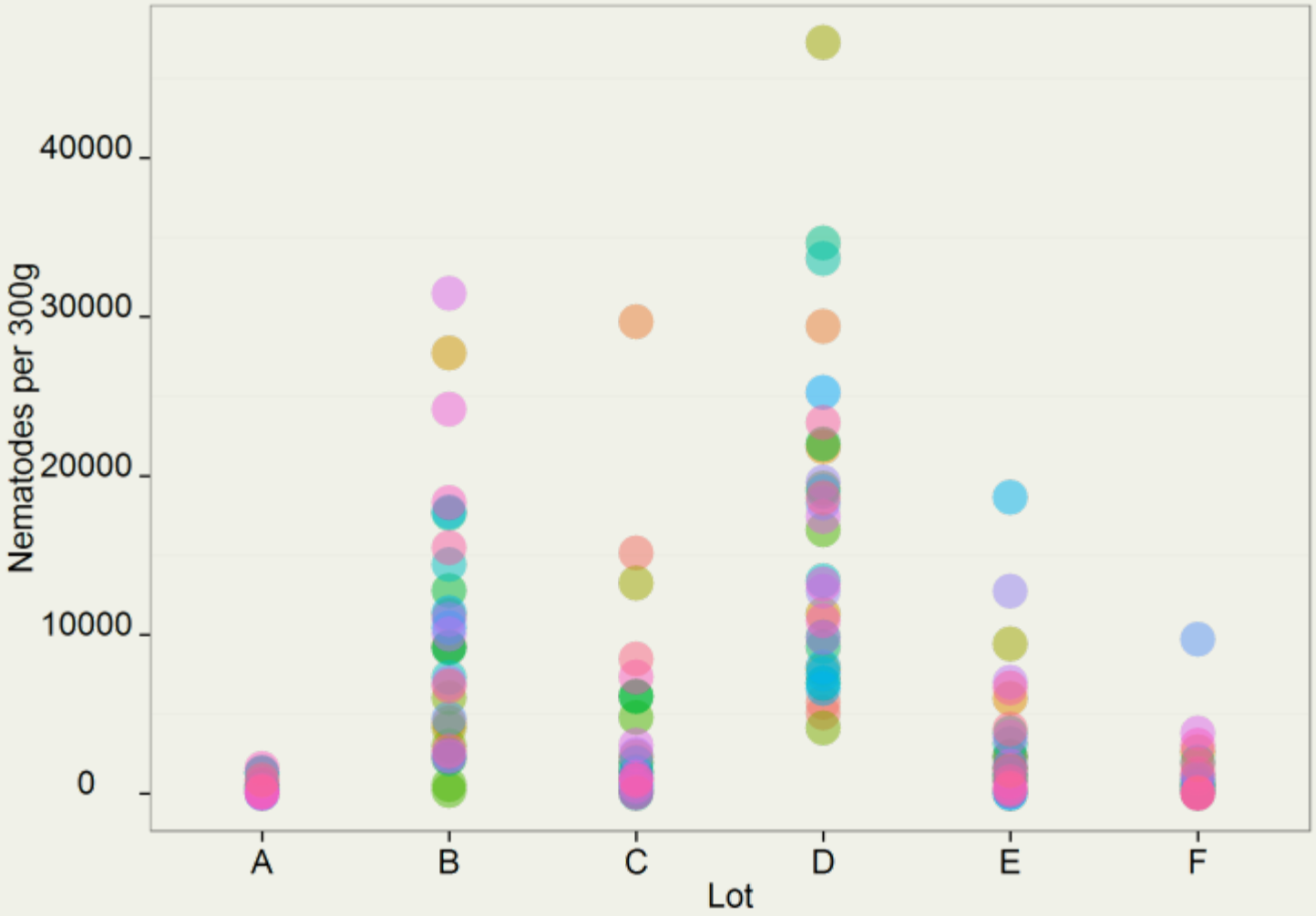
Ditylenchus: normal method

- Take 300g working sample from the submitted sample
- Extract into water
- Decant and examine water by microscopy
- Report positive if any are found (QUALITIATIVE TEST)

Ditylenchus: lot study method

- Take 300g working sample from the submitted sample. For half of the samples take duplicate 300g working samples (REPLICATION)
- Extract into water
- Decant and examine water by microscopy
- If nematodes are found, count them and report the number (QUANTITATIVE TEST)

Ditylenchus sp.- results



Level of pathogen may vary in different parts of the lot

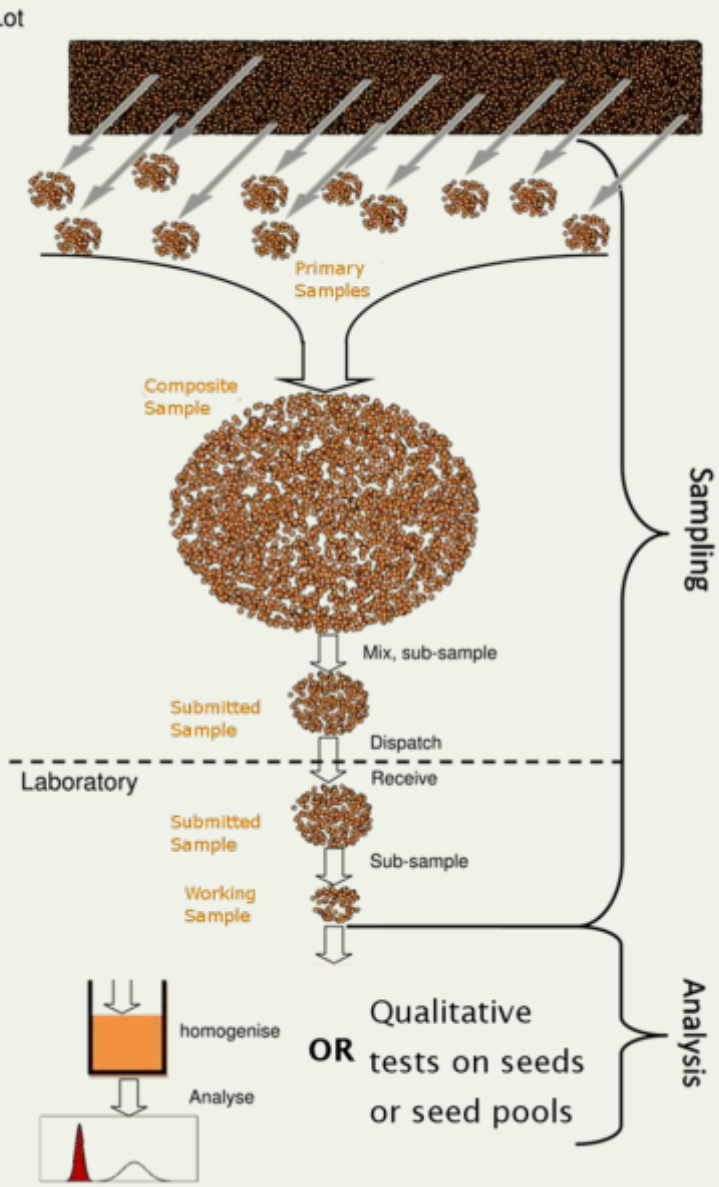
Take enough primary samples

We may want to detect low levels

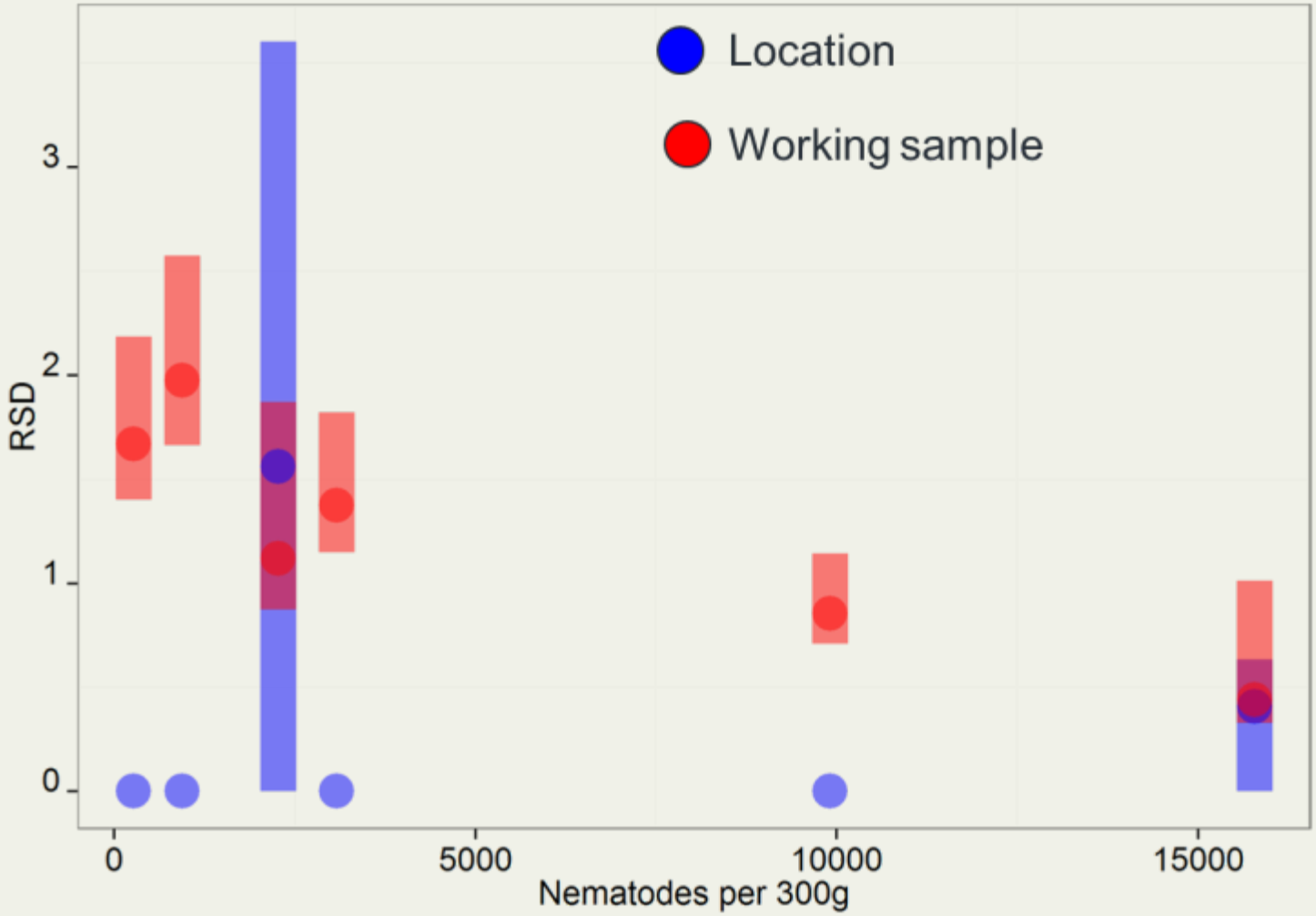
Make the working sample large enough

We may have imperfect detection

Modify method / replicate testing



Between location and between working sample variation



Estimates of factors that effect sampling: Ditylenchus

Mean nematodes (per 300g)	Parameter estimates RSD primary sample	RSD working sample (300 g)
269	0.001	1.670
944	0.000	1.975
3078	0.001	1.374
4201	1.558	1.118
9907	0.000	0.854
17048	0.407	0.441

Estimates of factors that effect sampling: tilletia

Mean spores (per seed)	Parameter estimates RSD primary sample	RSD working sample (100 seeds)
0.0446	0.364	3.65
0.707	0.678	0.685
1.14	0.730	0.000
2.33	0.906	0.104
2.80	0.265	0.168
6.93	1.01	0.0286

Estimates of factors that effect sampling: fusariums

Lot	Pathogen	Estimate on logit scale Estimate	se	s.d.	Mean prevalence (%) 95% C.I.		Primary sampling RSD
B	F.graminearum	-5.83	0.22	0.53	0.19	0.45	0.56
	F.poae	-6.74	0.24	0.00	0.07	0.19	0.00
D	F.graminearum	-7.02	0.49	0.15	0.03	0.23	0.15
	F.poae	-5.73	0.20	0.49	0.22	0.48	0.52
E	F.graminearum	-7.03	0.38	0.48	0.04	0.19	0.51
	F.poae	-8.19	0.50	0.00	0.01	0.07	0.00
H	F.graminearum	-7.15	0.42	0.67	0.03	0.18	0.75
	F.poae	-6.34	0.26	0.48	0.11	0.30	0.51
L	F.graminearum	-7.09	0.29	0.00	0.05	0.15	0.00
	F.poae	-5.75	0.15	0.00	0.24	0.42	0.00
M	F.graminearum	-6.94	0.27	0.00	0.06	0.16	0.00
	F.poae	-6.24	0.19	0.00	0.13	0.28	0.00

Estimates of factors that effect sampling: XCC

- Results expressed as cfu per sample
- Sparse results (lots of zeroes) with some very high values
- Analysed as between-sample variation. Assumed to be driven by between seed variation

Estimates of factors that effect sampling: XCC

Lot	Mean cfu per 2 seeds	Between 100-seed-sample RSD
A	0.0750	3.69
B	0.175	1.89
C	228	7.71

Effect of detection: method tilletia

- Two methods
  - Microscopy of an aliqout of extract: effective sample size = 9.07 seeds
  - Centrifugation and examination of whole extract effective sample size = 900 seed
- But during experimental comparison recovery from centrifuged samples is only 19.7% of the expected value: effective sample size 177.3 seeds

Effect of detection method: general approach

- Eg: For XCC take 10 000 seeds working sample soak in 100 ml, take 100µl aliqout for testing
- A perfect test for presence:
  - Sample size for effect of between seed variation = 10 000 seeds
  - Sample size for detecting low mean level in seeds = 10 seeds

Effect of detection method: general approach

- Eg: For XCC take 10 000 seeds working sample soak in 100 ml, take 100µl aliquot for testing
- A test for presence with LOD of 10 cfu (95% probability)
  - Sample size for effect of between seed variation = 10 000 seeds
  - Sample size for detecting low mean level in seeds = 2.59 seeds
- A test for presence with LOD of 100 cfu (95% probability)
  - Sample size for effect of between seed variation = 10 000 seeds
  - Sample size for detecting low mean level in seeds = 0.295 seeds

Effect of detection method: general approach

- Eg: For XCC take S seeds working sample soak in V_1, take V_2 aliquot for testing
- Test for presence with LOD of L cfu (p_d% probability)
  - Sample size for effect of between seed variation = S seeds
  - Sample size E for detecting low mean level in seeds :

$$E = S\frac{V_2}{V_1}\left(1-(1-p_d)^{1/L}\right)$$

QUESTIONS ON STUDIES?

Level of pathogen may vary in different parts of the lot

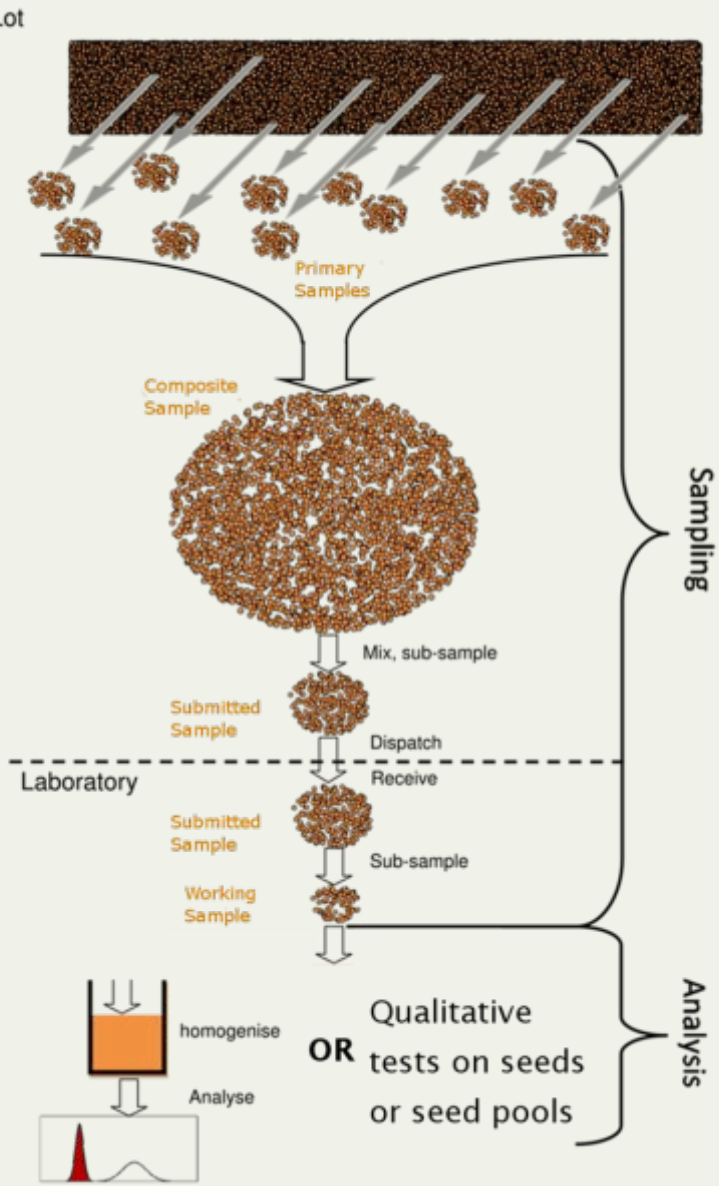
Take enough primary samples

We may want to detect low levels

Make the working sample large enough

We may have imperfect detection

Modify method / replicate testing



Elements of the plan

FACTOR	SAMPLING PLAN
Variation in different parts of lot	Number of primary samples
Variation at small scale (seeds)	Size of working sample
Average level of pathogen	Size of analytical aliqout
LOD of analytical method	Size of working sample / Size of analytical aliquot / Replication

Elements of the plan

FACTOR	SAMPLING PLAN
Variation in different parts of lot	Number of primary samples
Variation at small scale (seeds)	Size of working sample
Average level of pathogen	Size of analytical aliqout
LOD of analytical method	Size of working sample / Size of analytical aliquot / Replication

$$primary\ samples\ \geq \left(\frac{R_L}{0.18}\right)^2$$

Elements of the plan

FACTOR	SAMPLING PLAN
Variation in different parts of lot	Number of primary samples
Variation at small scale (seeds)	Size of working sample
Average level of pathogen	Size of analytical aliqout
LOD of analytical method	Size of working sample / Size of analytical aliquot / Replication

$$L_D = \frac{V_1\left((1-p_d)^{\frac{-R^2}{S}} - 1\right)}{V_2(1-(1-p_a)^{1/L_a})R^2}$$

Elements of the plans

$$primary\ samples\ \geq \left(\frac{R_L}{0.18}\right)^2$$

$$L_D = \frac{V_1\left((1-p_d)^{\frac{-R^2}{S}} - 1\right)}{V_2(1-(1-p_a)^{1/L_a})R^2}$$

L_D	**limit of detection of sampling plan (spores, cfu, pests per seed) with probability of detection p_d**
R_L	Between location variation in the expected level of pathogen in the lot
R	between seed variation in level of pest and pathogen expressed as RSD
S	**Number of seeds in working sample**
V_1	volume of extract or homogenate
V_2	volume of portion of extract or homogenate analysed or examined
L_a	Limit of detection of analytical method (spores, cfu, pests) with probability of detection p_a

Elements of the plans: Ditylenchus

$$primary\ samples \geq \left(\frac{R_L}{0.18}\right)^2$$

$$L_D = \frac{V_1\left((1-p_d)^{\frac{-R^2}{S}} - 1\right)}{V_2(1-(1-p_a)^{1/L_a})R^2}$$

L_D	**limit of detection of sampling plan (nematodes per gram) with 95% probability of detection**
R_L	up to RSD = 1.558
R	Variation between 300g samples up to RSD = 1.975
S	**Number of seeds in working sample**
V_1	100 ml
V_2	100 ml
p_a, L_a	100% probability of detection

Elements of the plans: Ditylenchus

$$primary\ samples\ \geq 74$$

Working sample size (g)	Limit of detection (nematodes per 100 g)
300	10152
600	29
900	4.1
1200	1.5
1500	0.80

Elements of the plans: Tilletia

$$primary\ samples \geq \left(\frac{R_L}{0.18}\right)^2$$

$$L_D = \frac{V_1\left((1-p_d)^{\frac{-R^2}{S}} - 1\right)}{V_2(1-(1-p_a)^{1/L_a})R^2}$$

L_D	**limit of detection of sampling plan (spores per seed) with 95% probability of detection**
R_L	up to RSD = 1.01
R	Variation between 300-seed samples up to RSD = 3.65
S	**Number of seeds in working sample**
V_1	62.5: 5.1
V_2	1 (based on filed of view of microscope): 1(centrifuge)
p_a, L_a	100% probability of detection

Elements of the plans: Tilletia

$$primary\ samples\ \geq 32$$

Working sample size (seeds)	Limit of detection (spores per seed)	
	Microscopy	Centrifuge
900	9372	761
1200	337	27
1500	46	3.7
1800	12	0.98
2100	4.7	0.38
2400	2.3	0.19
2700	1.3	0.11
3000	0.80	0.067

Elements of the plans: XCC

$$primary\ samples\ \geq \left(\frac{R_L}{0.18}\right)^2$$

$$L_D = \frac{V_1\left((1-p_d)^{\frac{-R^2}{S}} - 1\right)}{V_2(1-(1-p_a)^{1/L_a})R^2}$$

L_D	**limit of detection of sampling plan (cfu per seed) with 95% probability of detection**
R_L	not known
R	Variation between 100-seed samples up to RSD = 7.71
S	**Number of seeds in working sample**
V_1	100 ml
V_2	100 µl
p_a, L_a	1 to 10 cfu 95% probability of detection

Elements of plans: XCC

$$primary\ samples: up\ to\ 40\ (ISTA)$$

Working sample size (seeds)	Limit of detection (cfu per 1000 seed)	
	Analytical LOD = 1 cfu	Analytical LOD = 10 cfu
10 000	874	3207
15 000	403	1480
20 000	254	933
25 000	184	675
30 000	144	527
35 000	117	431
40 000	99	364
45 000	86	315

Elements of the plans: fusariums

$$primary\ samples \geq \left(\frac{R_L}{0.18}\right)^2$$

L_D Estimated numerically from beta-binomial distribution

L_D	**limit of detection of sampling plan proportion of infected seeds; 95% probability of detection**
R_L	up to RSD=0.75
R	Qualitative: infected, not infected
S	**Number of seeds in working sample**
V_1	1
V_2	1
p_a, L_a	100% probability of detection

Elements of plans: fusariums

$primary\ samples: 20$

Working sample size (seeds)	Limit of detection (proportion infected seeds)
100	3.1%
150	2.1%
200	1.5%
300	1.0%
400	0.78%

QA for sampling plans

Produce INDEPENDENT duplicate working samples.

- Test each working sample using the more quantitative version of available tests.
- Proportion positive or number of spore/cfu/pest for each working sample should be equivalent
- Standard approaches are:
- Little R.J.A., 1989, Testing the equality of two independent binomial proportions, The American Statistician, 43(4), 283-288 [in “Slippery Approach to Bayesianism”]
- Przyborowski J and Wilenski H, 1940, Homogeneity of Results in Testing Samples from Poisson Series: With an Application to Testing Clover Seed for Dodder, Biometrika, 31(3-4), 313-323

Examples of QA results that show unexpected variation

Counting pests, spores cfu

Lower count	Higher count
0	8
1	11
2	13
3	15
4	17
5	19
6	20
7	22
8	24
9	25
10	27

Counting positive sub-samples

Number of subsamples	Lower number of positives	Higher number of positives
4	0	4
5	0	5
6	0	6
6	1	6
7	0	7
7	1	7
7	2	7
7	0	6
7	1	6

Testa sampling WP

- Provides estimate of LOD against sampling effort for a number of specific scenarios.
- Provides some methods that can be used to make estimates of LOD and sampling effort for any scenario.
- Integrates the effects of sampling and analysis

- Testing lots in enough detail to estimate parameters can be expensive
- Getting hold of the right lots can be challenging

The team

UNIVERSITÀ DEGLI STUDI
DI MODENA E REGGIO EMILIA

CRA
CONSIGLIO PER LA RICERCA
IN AGRICOLTURA E L'ANALISI
DELL'ECONOMIA AGRARIA

Testa