

Interpretation of Efficacy Trials

The Need for Statistics

Jon Winfield

Head Human Health Branch

Regulatory requirements

- “Evidence must be submitted to show that the dose, timing and method of application recommended give adequate control....”
- “The design, analysis and reporting of trials must be in accordance with EPPO guidelines 152 and 181.”

Is a statistical test required?

- DIRECTIVE 91/414 – Annex III 6.2
- A statistical analysis of results amenable to such analysis must be carried out; where necessary the test guideline used must be adapted to enable such analysis.

Is a statistical test required?

- EPPO – PP1/152 (3)
- A decision on the need to conduct a statistical analysis of the results of a trial will depend on the results... and purpose
- Statistical analysis is not essential in all trials.

What are Efficacy trials for?

- Demonstrate effectiveness
- Justify the dose
- Demonstrate safety to the crop
- Demonstrate safety to other crops
- Support any other label claims

- In all cases
 - Across the range of conditions likely to be encountered in use
 - When used according to label advice

Demonstrating effectiveness

- Products must give *adequate* control
- There is no minimum level of acceptable control
- There is no requirement for comparative assessment (yet)

UK Criteria for Control Claims

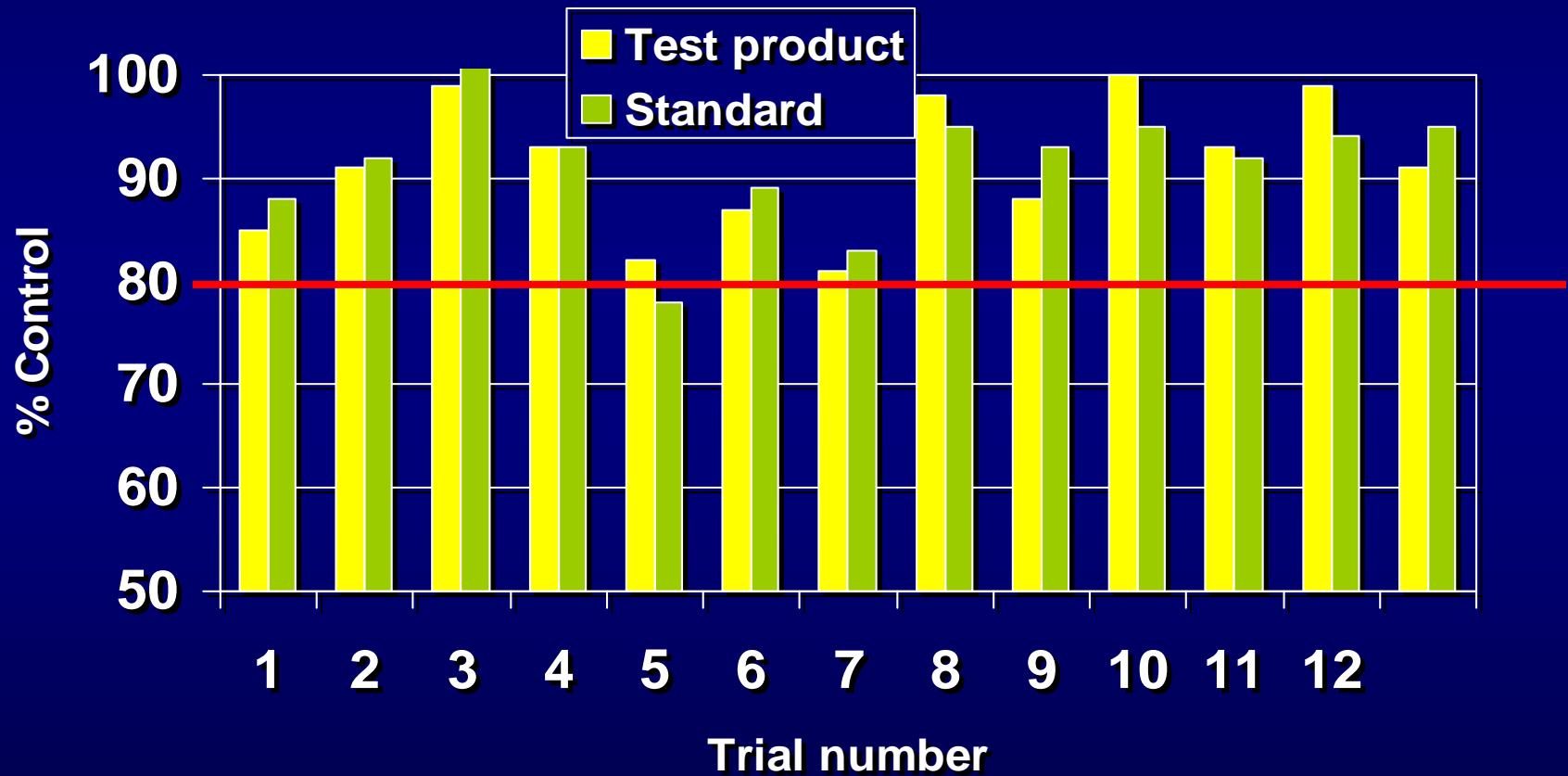
- Insecticides and Fungicides
 - Consistent control $\geq 80\%$
- Herbicides
 - Susceptible = Consistent control $\geq 85\%$
- Lower or less consistent control = lower claim

UK Criteria for Control Claims

- Exceptions
 - Pernicious weeds such as *Alopecurus myosuroides* & *Avena fatua* >95% control
 - Seed borne diseases such as bunt and smut >98% control

Demonstrating Effectiveness

Control of Septoria in wheat

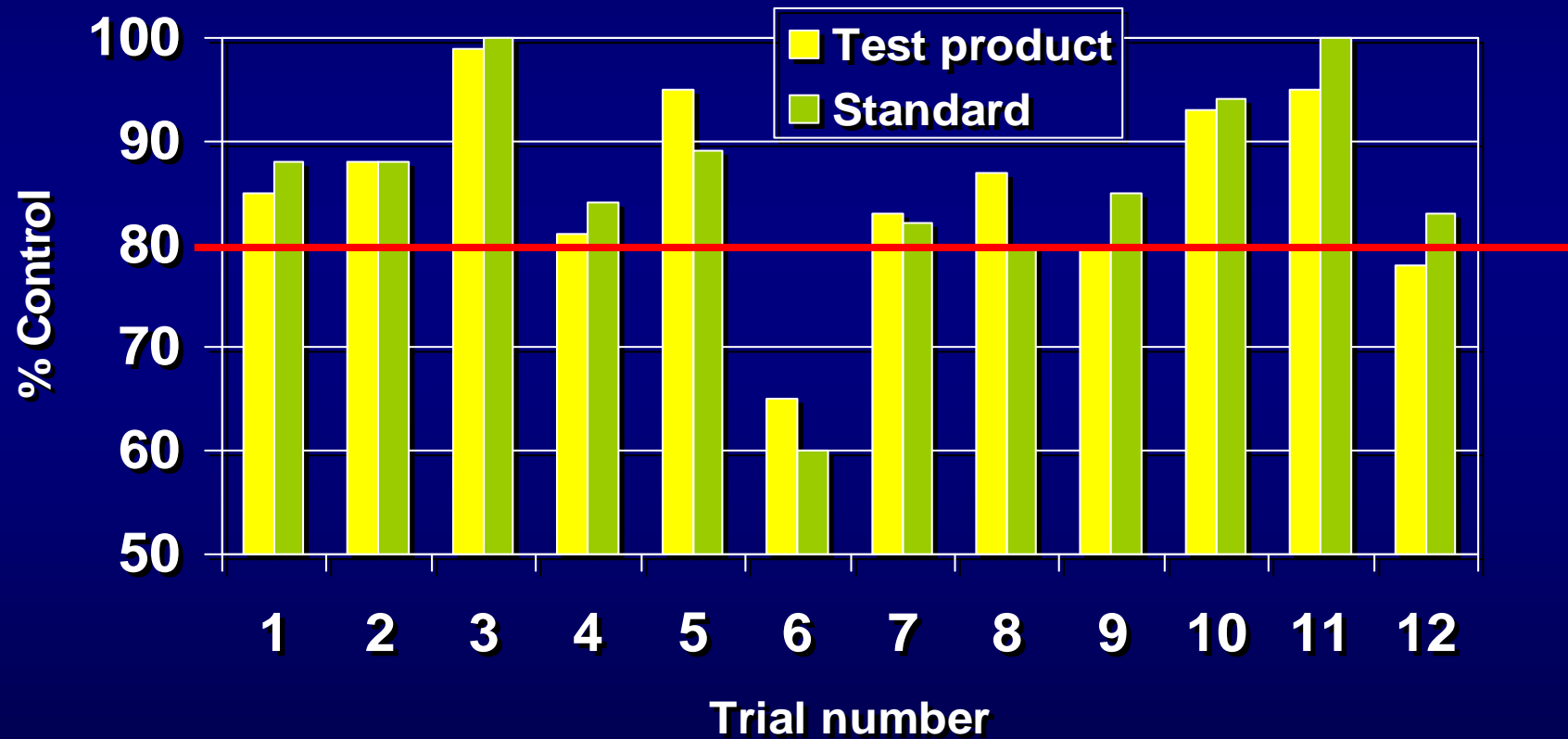


Statistical analysis

- Statistical comparison of Test Product and Standard is not required
- You could demonstrate a significant difference between levels of disease in the treated and untreated.....
- But do you need to do the test?

Demonstrating Effectiveness

Control of Septoria in wheat



Is a test possible?

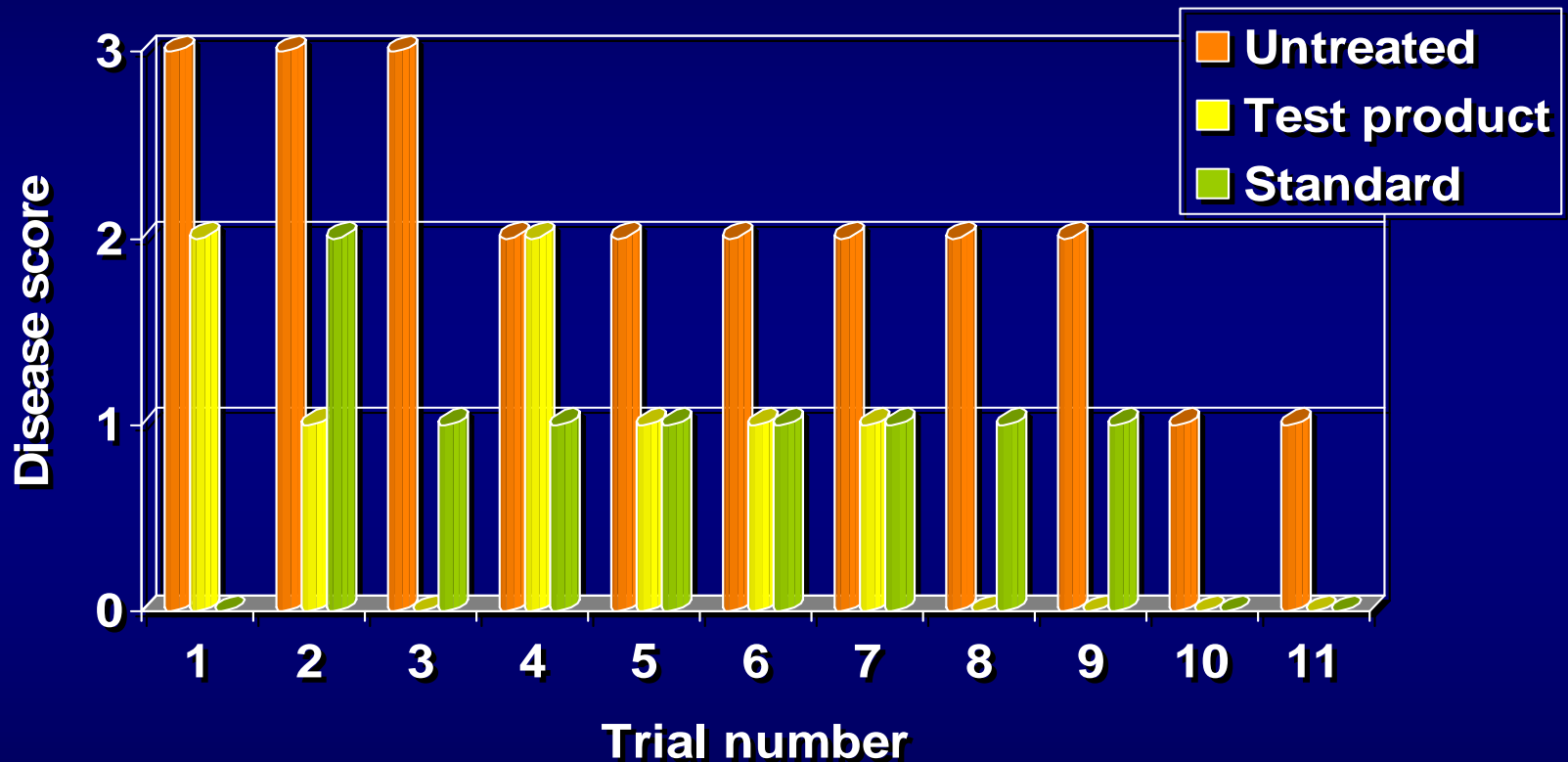
- Phoma stem canker in oilseed rape
- Assessed as scale
 - 0 = No disease
 - 1 = <20% disease
 - 2 = 20-50% disease
 - 3 = >50% disease

Phoma stem canker in oilseed rape

- Typical trials results

Trial No.	1	2	3	4	5	6	7	8	9	10	11	Mean
	Disease level											
Untreated	2	1	2	3	2	3	2	2	1	3	2	2.09
Test Product	1	0	0	0	0	2	2	1	0	1	1	0.73
Standard	1	0	1	1	1	0	1	1	0	2	1	0.82

Phoma stem canker in oilseed rape



Where statistics is required?

- For summaries (e.g. averages*)
- Where results are not clear cut
 - Where a difference is small
 - To prove that a difference is insignificant
- Where you are trying to prove there is no difference
 - Formulation comparability
 - Tests under different conditions e.g. non-UK data

Effectiveness results that may not be clear cut

- Plant Growth Regulators
- Where levels of control are poor
 - e.g. biological control agents
- Indirect measures of effectiveness
 - e.g. yield
- Dose justification
- Comparison of results from different situations
 - e.g. climatic regions

Crop safety

- Statistics may be useful in showing
 - No significant phytotoxic effects
 - No significantly effect on quantity or quality of yield
 - No significant effects on propagating material
 - No adverse effects on following crops
 - No adverse effects on adjacent crops
- If NO phytotoxic effects, statistical analysis probably not required

Example 1

Insecticide

- No phytotoxic effect in any of 16 effectiveness trials or 4 crop safety screens
- Tests cover all applications timings, water rates and a wide range of varieties
- Average yield is similar to standards and disease free crops, > than infected & untreated crops
- Seed from treated crops shows >90% germination, similar to untreated and standards.
- **Statistical tests will add little value**

Example 2(a)

Herbicide

- No phytotoxic effect in any of 16 effectiveness trials or 4 crop safety screens
- Tests cover all applications timings, water rates and a wide range of varieties
- Average yield is similar to standards and weed free crops, > than infected & untreated crops
- Seed from treated crops shows >90% germination, similar to untreated and standards.
- **Statistical tests will add little value**

Example 2(b)

Herbicide

- Some trials show minor necrosis and chlorosis at N
- Trials at 2N show consistent 5-10% necrosis and chlorosis
- Yields are similar in all treatments
- **Statistical tests essential to show if effects on yield are significant**

What comparisons are useful?

Example

- New herbicide (New) mixing existing residual and contact actives (Res and Con)
- No new label claims,
- but BAD claims faster more reliable effect than products containing Res alone

Test products; Effectiveness trials

Product	RES (g/ha)	CON (g/ha)	Form ⁿ	Doses
New	2000	50	EC	0.5,0.8,N
New WDG	2000	50	WDG	0.5,0.8,N
Std 1	2000		WP	0.5,0.8,N
Std 2	2000		WDG	0.5,0.8,N
Std 3		150	EC	1/3 N
Std 2&3	2000	50		N + 1/3 N

Site reference		1		2		3		4		5	
Crop/variety		Wizard		Wizard		Target		Fuege		Dilco	
Application Date		5.4.06		5.4.06		4.3.06		12.5.06		5.12.05	
Crop GS (BBCH) at application		14		18		12		16		15	
Assessment date		17.5.06	6.6.06	5.5.06	31.5.06	20.4.06	3.6.06	26.6.06	19.7.06	23.1.06	10.4.06
Weed GS (BBCH) at assessment		59	65	32	65	13-26	65	32	65	13-21	23-31
No weeds/m ²		10.8%	27.5%	7.5%	15%	64	144	66	89	11	2.25
Treatment – evaluation interval (days)		42	62	30	56	47	91	45	68	43	120
Treatment	Rate of use										
1. Untreated	-	0a	0a	0.0 d	0.0 c	0.0c	0.0 c	0.0 b	0 a	0.0 b	0.0 b
2. New	7.0 l/ha	69.8 bc	47.3 a	17.5 b	10.0 bc	90.8 a	86.3 ab	25.0 a	95.0 b	100.0 a	85.7 a
3. New	5.6 l/ha	60.5 ab	54.5 a	17.5 b	12.5 bc	85.8 a	80.0 ab	25.0 a	72.5 b	100.0 a	55.4 ab
4. New	3.5 l/ha	25.6 bc	22.7 a	10.0 c	12.5 bc	82.0 a	73.8 ab	22.5 a	77.5 b	100.0 a	60.7 ab
5. New WDG	3.5 kg/ha	67.4 bc	36.4 a	12.5 bc	12.5 bc	90.8 a	88.8 a	28.8 a	82.5 b	100.0 a	73.2 ab
6. New WDG	2.8 kg/ha	72.1 bc	57.3 a	10.0 c	15.0 bc	82.0 a	71.3 ab	31.3 a	67.5 b	100.0 a	58.9 ab
7. New WDG	1.75 kg/ha	39.5 c	36.4 a	10.0 c	12.5 bc	73.3 a	40.0 abc	26.3 a	55.0 b	100.0 a	85.7 a
8. Std 1	3.5 kg/ha	79.1 bc	54.5 a	17.5 b	22.5 b	94.5 a	80.0 ab	23.8 a	87.5 b	100.0 a	100.0 a
9. Std 1	2.8 kg/ha	55.8 ab	57.3 a	10.0 c	17.5 bc	68.8a	63.8 ab	26.2 a	80.4 b	100.0 a	87.5 a
10. Std 1	1.75 kg/ha	27.9 bc	13.6 a	10.0 c	12.5 bc	80.0 a	81.3 ab	21.2 a	57.1 b	100.0 a	75.0 ab
11. Std 2	3.0 kg/ha	74.4 bc	36.4 a	12.5 bc	12.5 bc	67.5 a	63.8 ab	23.8 a	92.5 b	100.0 a	60.7 ab
12. Std 2	2.4 kg/ha	48.8 bc	36.4 a	10.0 c	15.0 bc	84.5 a	83.8 ab	25.0 a	65.0 b	100.0 a	73.2 ab
13. Std 2	1.5 kg/ha	41.9 bc	36.4 a	10.0 c	10.0 bc	74.7 a	37.5 abc	23.8 a	72.5 b	100.0 a	48.2 ab
14. Std 3	0.5 l/ha	44.2 c	48.2 a	90.0 a	97.7a	27.5 b	27.5 bc	30.0 a	70.0 b	100.0 a	60.7 ab
15. Std 2 & 3	3.0 kg/ha + 0.5 l/ha	81.4 c	57.3 a	90.0 a	95.0 a	91.4 a	66.3 ab	27.5 a	95.0 b	100.0 a	85.7 a
LSD (P= .05)		3.14	9.44	4.16	10.23	26.58	34.60	10.33	3.32	0.0	47.4
Standard deviation		2.20	6.16	2.91	7.16	18.60	24.21	7.23	2.32	0.0	33.2
CV		43.8	39.80	13.35	30.04	25.52	38.48	30.12	81.02	0.0	49.2

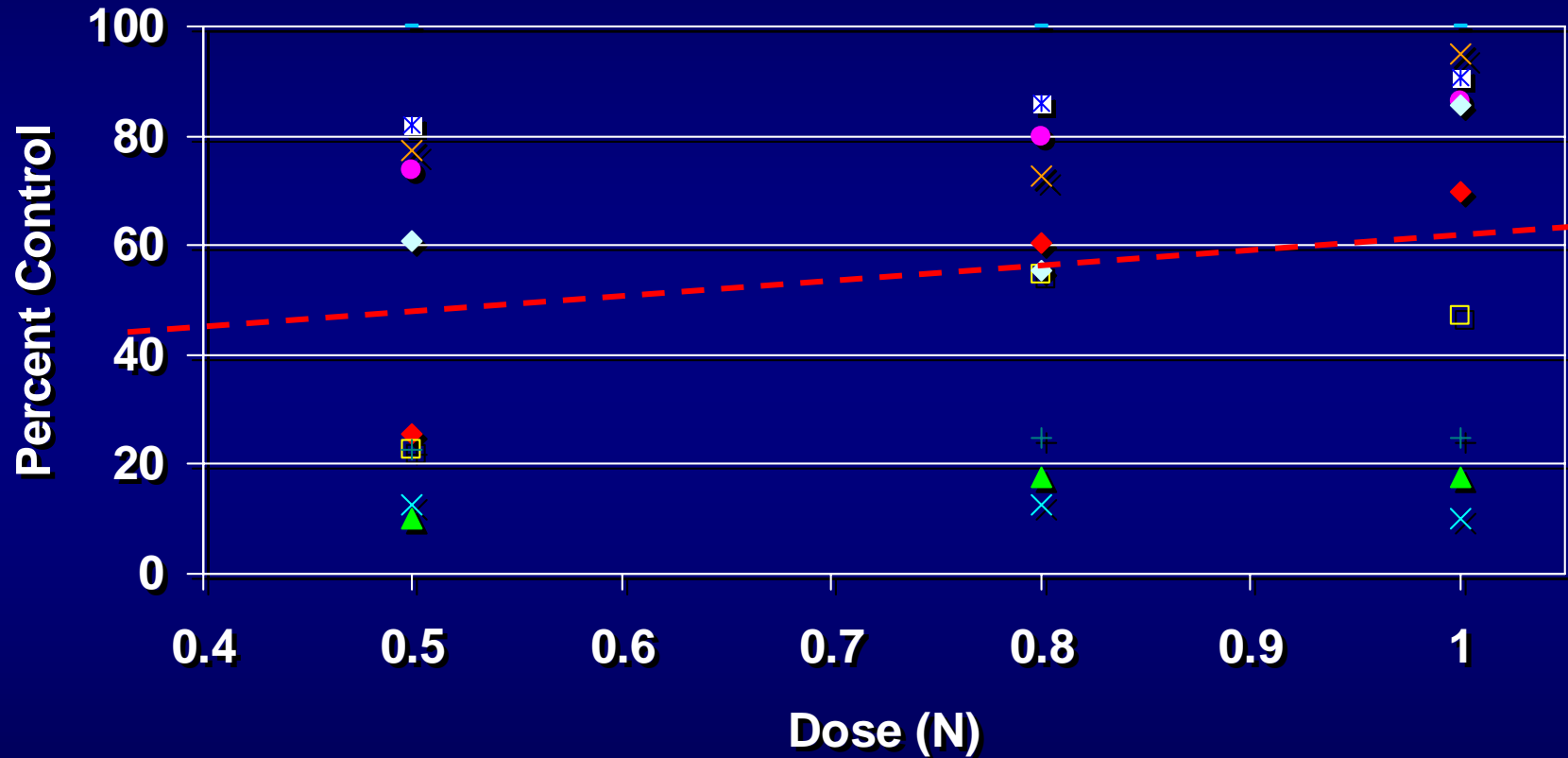
Dose justification

- Often presented as multiple comparison tests
 - including results from standards
- May show no significant difference
 - But lower doses consistently give lower control !

Dose Justification

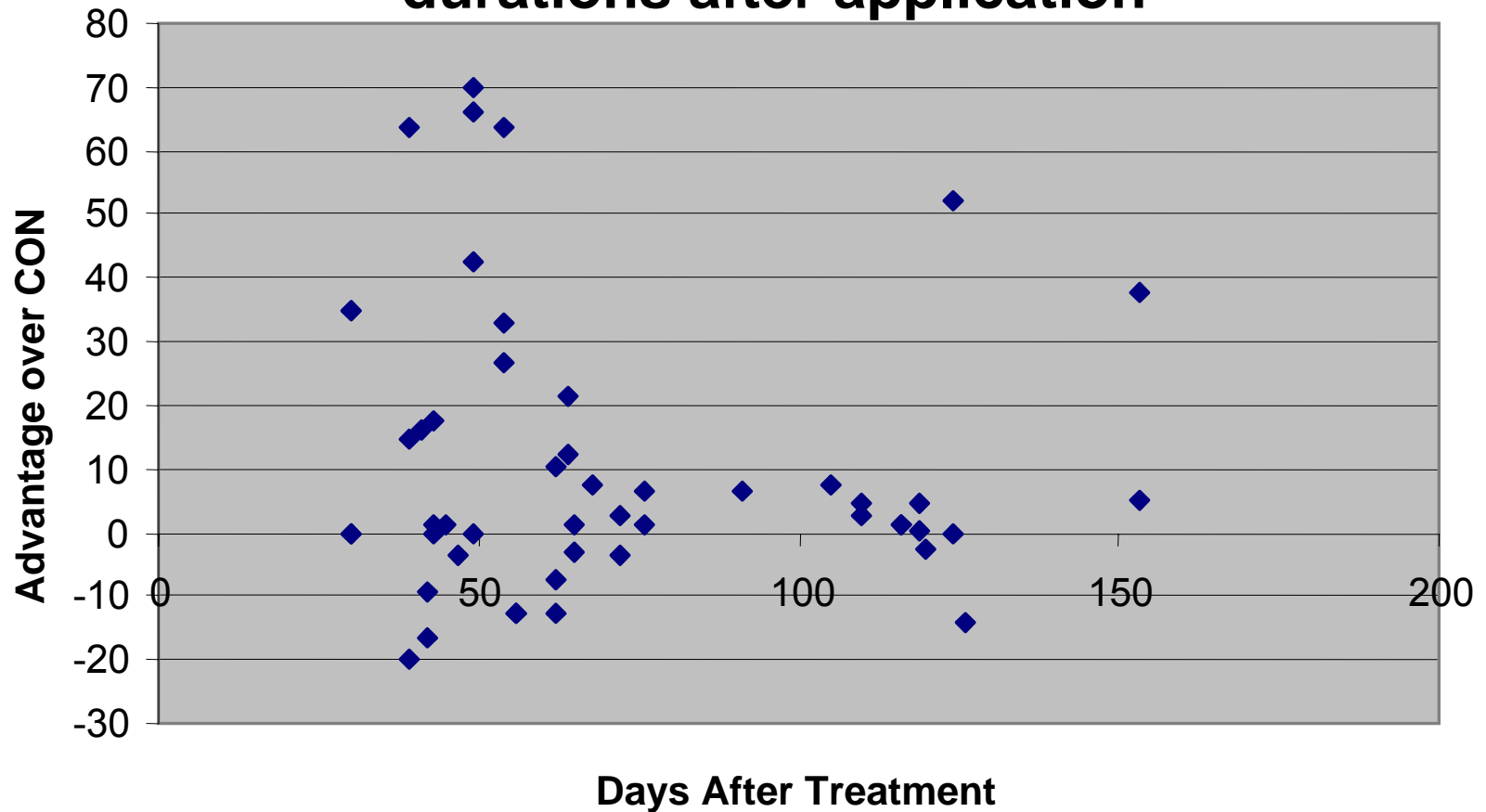
1. Untreated	-	0a	0a	0.0 d	0.0 c	0.0c	0.0 c	0.0 b	0 a	0.0 b	0.0 b
2. New	N	69.8 bc	47.3 a	17.5 b	10.0 bc	90.8 a	86.3 ab	25.0 a	95.0 b	100.0 a	85.7 a
3. New	0.8 N	60.5 ab	54.5 a	17.5 b	12.5 bc	85.8 a	80.0 ab	25.0 a	72.5 b	100.0 a	55.4 ab
4. New	0.5 N	25.6 bc	22.7 a	10.0 c	12.5 bc	82.0 a	73.8 ab	22.5 a	77.5 b	100.0 a	60.7 ab
5. New WDG	N	67.4 bc	36.4 a	12.5 bc	12.5 bc	90.8 a	88.8 a	28.8 a	82.5 b	100.0 a	73.2 ab
6. New WDG	0.8 N	72.1 bc	57.3 a	10.0 c	15.0 bc	82.0 a	71.3 ab	31.3 a	67.5 b	100.0 a	58.9 ab
7. New WDG	0.5 N	39.5 c	36.4 a	10.0 c	12.5 bc	73.3 a	40.0 abc	26.3 a	55.0 b	100.0 a	85.7 a
8. Std 1	N	79.1 bc	54.5 a	17.5 b	22.5 b	94.5 a	80.0 ab	23.8 a	87.5 b	100.0 a	100.0 a
9. Std 1	0.8 N	55.8 ab	57.3 a	10.0 c	17.5 bc	68.8a	63.8 ab	26.2 a	80.4 b	100.0 a	87.5 a
10. Std 1	0.5 N	27.9 bc	13.6 a	10.0 c	12.5 bc	80.0 a	81.3 ab	21.2 a	57.1 b	100.0 a	75.0 ab
11. Std 2	N	74.4 bc	36.4 a	12.5 bc	12.5 bc	67.5 a	63.8 ab	23.8 a	92.5 b	100.0 a	60.7 ab
12. Std 2	0.8 N	48.8 bc	36.4 a	10.0 c	15.0 bc	84.5 a	83.8 ab	25.0 a	65.0 b	100.0 a	73.2 ab
13. Std 2	0.5 N	41.9 bc	36.4 a	10.0 c	10.0 bc	74.7 a	37.5 abc	23.8 a	72.5 b	100.0 a	48.2 ab
14. Std 3	1/3 N	44.2 c	48.2 a	90.0 a	97.7a	27.5 b	27.5 bc	30.0 a	70.0 b	100.0 a	60.7 ab
15. Std 2 & 3	N + 1/3 N	81.4 c	57.3 a	90.0 a	95.0 a	91.4 a	66.3 ab	27.5 a	95.0 b	100.0 a	85.7 a
LSD (P= .05)		3.14	9.44	4.16	10.23	26.58	34.60	10.33	3.32	0.0	47.4
Standard deviation		2.20	6.16	2.91	7.16	18.60	24.21	7.23	2.32	0.0	33.2
CV		43.8	39.80	13.35	30.04	25.52	38.48	30.12	81.02	0.0	49.2

Dose Justification



PSD evaluators summary

Summary of weed control at different durations after application



PSD evaluators summary

- Justification required for the approval of the product
- Does the product add to the level of control of the standard?
- Statistics required for effectiveness
- Acceptable under COPR but may not be under PPPR

Test products; Crop Safety

Product	RES (g/ha)	CON (g/ha)	Form ⁿ	Doses
New	2000	50	EC	N, 2N
New WDG	2000	50	WDG	N, 2N
Std 1	2000		WP	N, 2N
Std 2	2000		WDG	N, 2N
Std 2&3	2000	50		N, 2N

Phytotoxicity assessments

Site reference		OSR S2				OSR S5			
Crop/variety		Caracas				Winner			
Application Date		17.11.05				16.1.06			
Crop GS (BBCH) at application		15				14-16			
Assessment date		30.11.05	19.12.05	21.4.06	30.5.06	30.1.06	27.2.06	3.4.06	10.5.06
Crop GS (BBCH) at assessment		15	16	33	65	16	16	30	65
Treatment – evaluation interval (days)		13 DAT	32 DAT	155DAT	194 DAT	14 DAT	42 DAT	77 DAT	114 DAT
Treatment	Rate of use								
1. Untreated	-	0.0c	0.0 d	0.0 e	0.0 a	0.0b	0.0 b	0.0 a	0.0 a
2. NEW	7.0 l/ha	1.3 bc	0.0 d	0.0 e	0.0 a	0.0b	0.0 b	0.0 a	0.0 a
3. NEW	14.0 l/ha	4.0 abc	3.0 c	1.3 de	0.3 a	0.0b	0.0 b	0.0 a	0.0 a
4. New WDG	3.5 kg/ha	4.0 abc	5.5 bc	4.0 b	2.0 a	0.0b	0.0 b	0.0 a	0.0 a
5. New WDG	7.0 kg/ha	5.8 a	7.3 b	3.8 bc	0.5 a	0.0b	0.0 b	0.0 a	0.0 a
6. Std 1	3.5 kg/ha	4.3 ab	3.0 c	1.8 b-e	0.3 a	0.0b	0.0 b	0.0 a	0.0 a
7. Std 1	7.0 kg/ha	6.0 a	5.8 b	3.5 bcd	1.0 a	2.8a	2.8 a	0.0 a	0.0 a
8. Std 2	3.0 kg/ha	3.3 abc	3.0 c	1.5 cde	0.3 a	0.5b	0.5 b	0.0 a	0.0 a
9. Std 2	6.0 kg/ha	3.3 abc	6.0 b	3.3 bcd	0.5 a	0.0b	0.0 b	0.0 a	0.0 a
10. Std 2 + Std 3	3.0 kg/ha + 0.5 l/ha	3.3 abc	6.8 b	4.0 b	1.0 a	0.0b	0.0 b	0.0 a	0.0 a
11. Std 2 + Std 3	6.0 kg/ha + 1.0 l/ha	6.8	11.0 a	6.8 a	1.8 a	0.0b	0.0 b	0.0 a	0.0 a
LSD (P= .05)		2.5	1.7	1.4	1.0	0.3	0.3	0.0	0.0
Standard deviation		66.4	36.6	52.1	149.3	112.6	112.6	0.0	0.0
CV		3.6	2.5	2.0	1.5	0.5	0.5	0.0	0.0

Vigour assessments

Site reference		OSR S2			
Crop/variety		Caracas			
Application Date		17.11.05			
Crop GS (BBCH) at application		15			
Assessment date		30.11.05	19.12.05	21.4.06	30.5.06
Crop GS (BBCH) at assessment		15	16	33	65
Treatment – evaluation interval (days)		13 DAT A	32 DAT A	155 DAT A	194 DAT A
Treatment	Rate of use				
1. Untreated	-	100.0 a	100.0 a	100.0 a	100.0 a
2. NEW	7.0 l/ha	100.0 a	100.0 a	100.0 a	100.0 a
3. NEW	14.0 l/ha	100.0 a	100.0 a	98.5 a	100.0 a
4. New WDG	3.5 kg/ha	100.0 a	100.0 a	96.8 abc	99.5 a
5. New WDG	7.0 kg/ha	100.0 a	100.0 a	92.3 de	96.3 b
6. Std 1	3.5 kg/ha	100.0 a	100.0 a	97.8 ab	100.0 a
7. Std 1	7.0 kg/ha	100.0 a	100.0 a	93.8 cd	98.3 a
8. Std 2	3.0 kg/ha	100.0 a	100.0 a	97.3 abc	99.5 a
9. Std 2	6.0 kg/ha	100.0 a	100.0 a	94.8 bcd	98.5 a
10. Std 2 + Std 3	3.0 kg/ha + 0.5 l/ha	100.0 a	100.0 a	94.8 bcd	99.0 a
11. Std 2 + Std 3	6.0 kg/ha + 1.0 l/ha	100.0 a	100.0 a	90.0 e	96.3 b
LSD (P= .05)		0.0	0.0	2.2	1.3
Standard deviation		0.0	0.0	2.3	1.3
CV		0.0	0.0	3.2	1.8

More important than a statistical test

- Valid and consistent methodologies
 - Same assessment timing and methods
 - Same standards
- Sufficient pest population
- Tests conducted over a range of conditions to fully cover the label recommendations and expected agronomic conditions
- Explain poor results

Regulatory submissions

Frequently

- Show little if any thought given to statistics
- Poor choice of tests
- Include inappropriate data
 - Trials software automatically produces ANOVA across all the data

What does the regulator want?

- Where a statistical analysis would help
- Identify what you are trying to demonstrate
 - AKA null hypothesis!
- Choose an appropriate test for the data and hypothesis
- Use only the data necessary to test the hypothesis

Also important

- Design (and conduct) the trial properly
- Try and ensure consistency across the trial series
- Look at the raw data!
- Analyse across trials
- Explain anomalies
- Would a graphical summary help?