US ARMY DUGWAY PROVING GROUND

# Test Design Adequacy for Logistic Regression

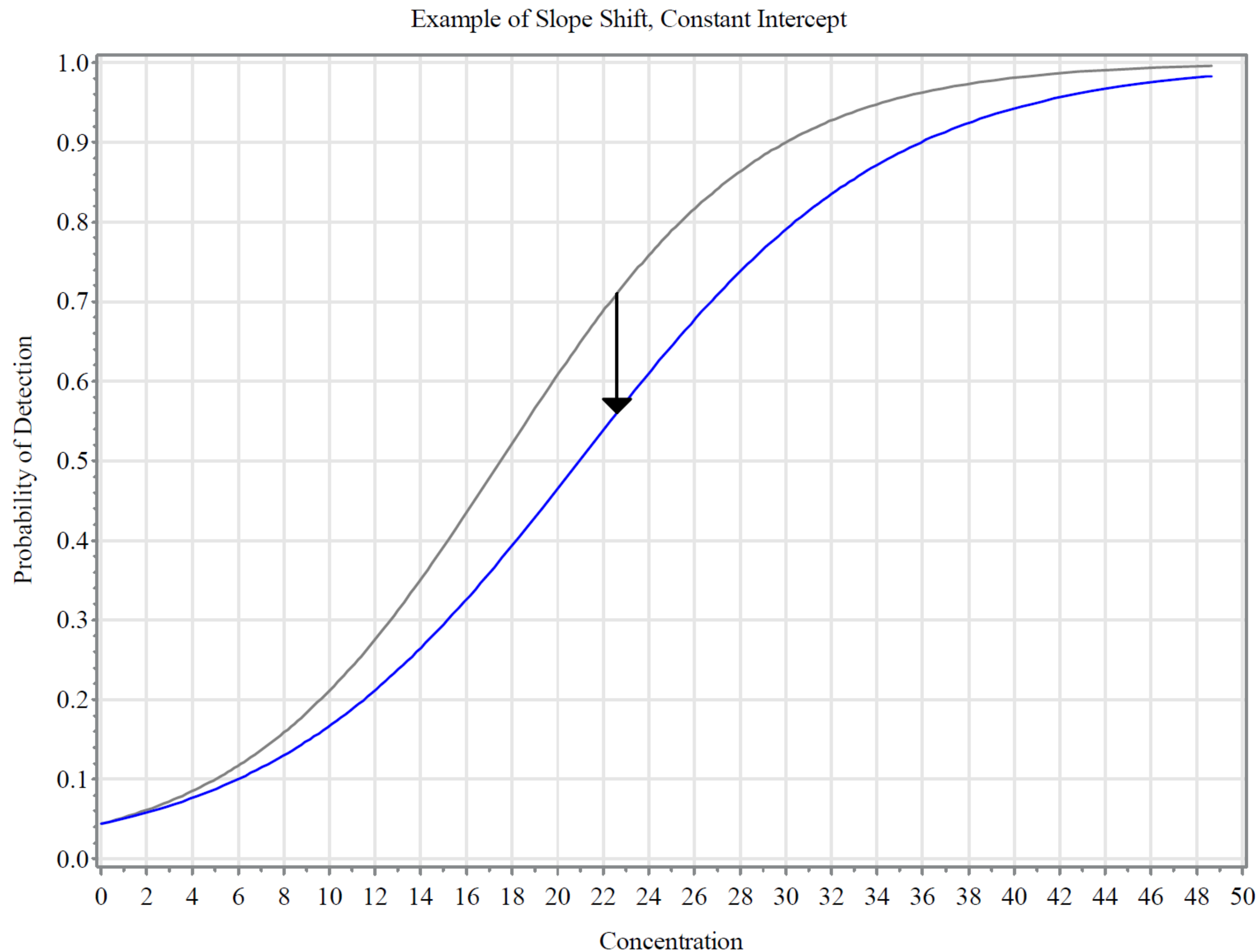
Prepared by Kendal Ferguson and Scott Hunter

For Conference on Applied Statistics in Defense
23 October 2016

# Motivation

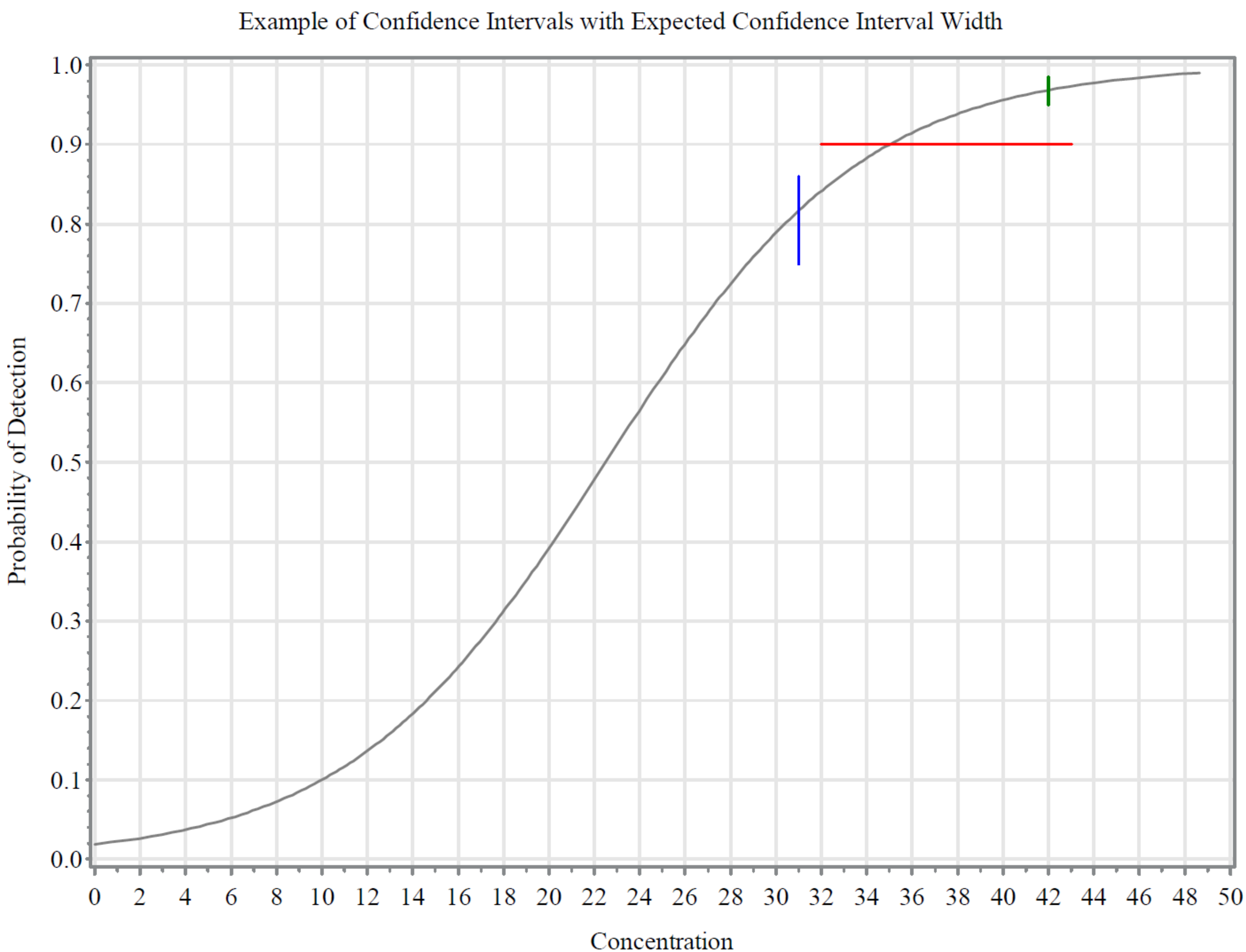
- When creating test designs, power analysis can be a useful tool in helping to decide design adequacy but shouldn't necessarily be THE answer
- Software is generally useful in calculating power. However, things become more complicated when dealing with a binary response
- We wanted to develop a methodology to help users and decision makers visualize and understand the impact of their assumptions and make assessment of test design adequacy more meaningful NOT just provide a power number

# Scenario



Example of Slope Shift, Constant Intercept
Probability of Detection
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50
Concentration

# Other Measures to Consider



# DesignExpert

- S... and

Evaluation Power Options

**Degrees of Freedom for Evaluation**

| | |
|---|---|
| Model | 2 |
| Residuals | 77 |
| *Lack of Fit* | *7* |
| *Pure Error* | *70* |
| Corr Total | 79 |

A recommendation is a minimum of 3 lack of fit df and 4 df for pure error.

This ensures a valid lack of fit test.

Fewer df will lead to a test that may not detect lack of fit.

**Power calculations are performed using response type "Proportion".**

**Delta=0.1, Proportion=0.9, Run Variation (as % grand mean proportion)=10%, Samples per Run=1**

**Note: Currently FDS is unavailable for proportion responses.**

| | | | | Power at 20 % alpha level to detect signal/noise ratio of |
|---|---|---|---|---|
| **Term** | **StdErr1** | **VIF** | **Ri-Squared** | **0.33346 Std. Dev.** |
| A | 0.32 | 1.00 | 0.0000 | 26.0 % |
| B | 0.11 | 1.00 | 0.0000 | 58.3 % |

☐ Arcsine Square Root

☑ Logit

OK | Cancel | Apply | Help

# SAS Simulation Input

- Reads in design from .csv
- Inputs

| | |
|---|---|
| Number of Simulations | 10000 |
| Factor Name | Detector |
| Number of Levels | 2 |
| Covariate | Concentration |
| Covariate Unit | $mg/m^3$ |
| Slope Sign (1-pos, 2-neg, 3 +/-) | 1 |
| Alpha (Hypothesis Test) | 20 |
| Alpha (Confidence Interval) | 20 |
| Simulation Quantile | 80 |

# SAS Simulation Input

| | |
|---|---|
| Low Value (x1) | 1.6 |
| Probability at Low Value (x1) | 0.5 |
| Std Dev for Low Value (x1) | 0.3 |
| High Value (x2) | 3.2 |
| Probability at High Value (x2) | 0.9 |
| Std Dev for High Value (x2) | 0.5 |
| Correlation (x1,x2) | 0.4 |

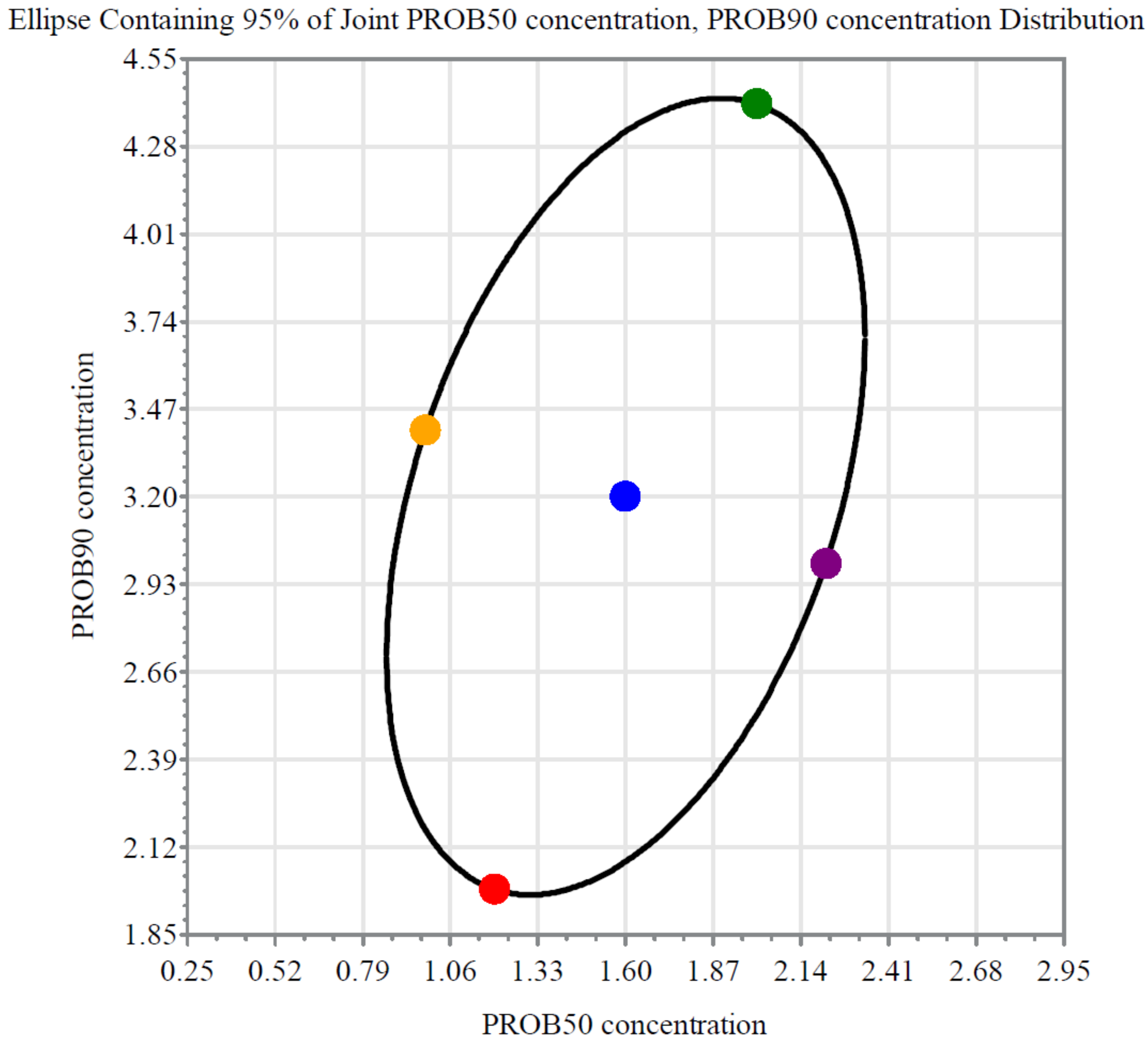
# SAS Simulation Input

| | |
|---|---|
| Variant2 Change Int=1 Slope=2 | 2 |
| Variant2 Shift (Intercept) | |
| Variant2 Shift (Slope) | -0.1 |
| | |
| Confidence Interval on X for Desired Probability Level | 90 |
| | |
| Confidence Interval for a Desired X | 4 |
| Confidence Interval for a Desired X | 5 |

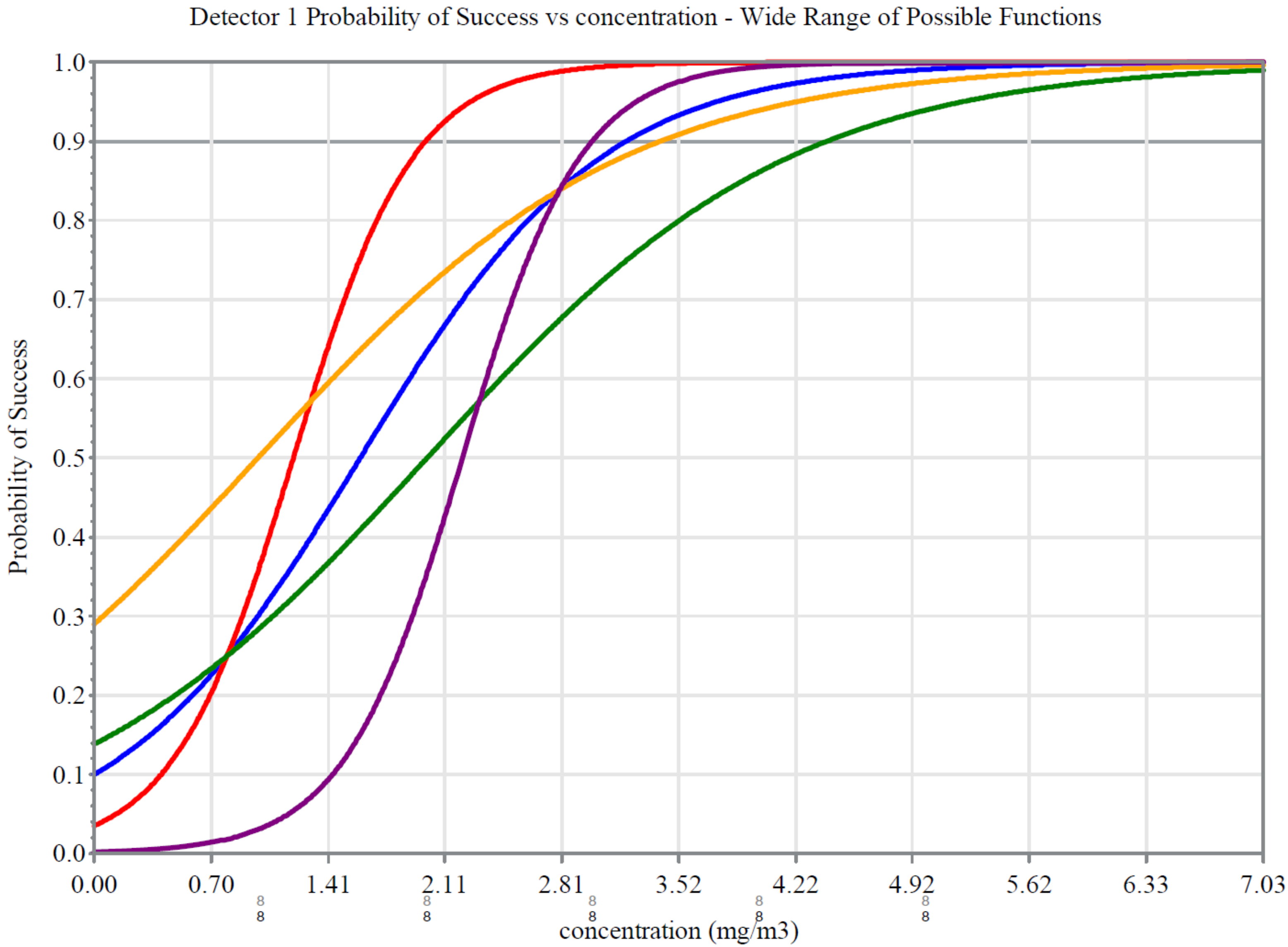
# Behind the Scenes Work

- Generate the baseline and shifted curves using inputs to create baseline and Newton Raphson method to create the shifted curve
- Generate random data from these curves at test design points
- Obtain coefficient estimates for logistic models and calculate confidence intervals at points of interest
- Obtain distributions of interval widths and report width at desired quantile
- Estimate power using proportion of simulations which indicate shift is significant

# Output


Ellipse Containing 95% of Joint PROB50 concentration, PROB90 concentration Distribution
4.55
4.28
4.01
3.74
3.47
3.20
2.93
2.66
2.39
2.12
1.85
0.25
0.52
0.79
1.06
1.33
1.60
1.87
2.14
2.41
2.68
2.95
PROB90 concentration
PROB50 concentration

# Output




Number of detection opportunities at amounts shown in gray (Detector 1) and black ( Detector 2)

# Output

```
Asymptotic Requirements
Percent of 80% Confidence Intervals Containing True b: 79.9
Percent of 80% Confidence Intervals Containing True c: 82.4
Percent of detector 1 80% Confidence Intervals Containing True PROB90 Amt: 83
Percent of detector 2 80% Confidence Intervals Containing True PROB90 Amt: 85.5
Percent of detector 1 80% Confidence Intervals Containing True 4 mg/m3 Probability of Success: 79.8
Percent of detector 2 80% Confidence Intervals Containing True 4 mg/m3 Probability of Success: 79.3
Percent of detector 1 80% Confidence Intervals Containing True 5 mg/m3 Probability of Success: 79.4
Percent of detector 2 80% Confidence Intervals Containing True 5 mg/m3 Probability of Success: 79.2

Power(Alpha= 20%)
Percent of Simulations with Significantly Positive b: 100
Percent of Simulations with Significantly Negative c: 44.5

Two-Sided Confidence Interval Widths(Alpha= 20%, Percentile=80%)
Width of 80th Percentile Interval on detector 1 PROB90 Amt: 1.73 mg/m3
Width of 80th Percentile Interval on detector 2 PROB90 Amt: 2.65 mg/m3
Width of 80th Percentile Interval on detector 1 4 mg/m3 Probability of Success: 15.4%
Width of 80th Percentile Interval on detector 2 4 mg/m3 Probability of Success: 21.4%
Width of 80th Percentile Interval on detector 1 5 mg/m3 Probability of Success: 9.7%
Width of 80th Percentile Interval on detector 2 5 mg/m3 Probability of Success: 17.7%
```

# Conclusion

- Statistical software give power numbers, but do they make sense?
- The simulation approach provides a more intuitive way to understand if testing is adequate
- Other indicators can be evaluated besides power
- Outputs clearly illustrate assumptions and provide discussion points for decision makers

# Future Work

- Add more factors
- More comparisons with JMP simulations
- Web-based app so those without SAS can use the method

# References

Institute for Defense Analyses, Alexandria, Virginia, IDA Document Number D-5205, Power Analysis Tutorial for Experimental Design Software, November 2014.

Stat-Ease Inc, Minneapolis, Minnesota, Version 10.01, Handbook for Experimenters, http://www.statease.com/pubs/handbk_for_exp_sv.pdf, accessed on 25 October 2016.

Mathematics Department, University of British Columbia, Vancouver, BC Canada, The Newton-Raphson Method, https://www.math.ubc.ca/~anstee/math104/104newtonmethod.pdf, accessed on 8 August 2016.

Stanford University, Stanford, California, Taylor Approximation and the Delta Method, http://web.stanford.edu/class/cme308/OldWebsite/notes/TaylorAppDeltaMethod.pdf, accessed on 25 October 2016.