Analysis of a T cell frequency assay

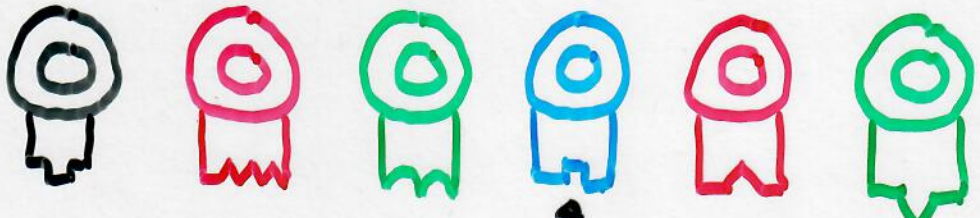
Karl Broman
University of California, Berkeley

joint work with :

Terry Speed, U. of Cal., Berk.

Michael Tigges, Chiron Corp.

A bit of immunology:



a variety of T-cells

antigen

antigen-presenting cell

Proliferation of activated T-cell clones

differentiation

memory cells (respond more quickly next time)

effector cells (destroy antigen)

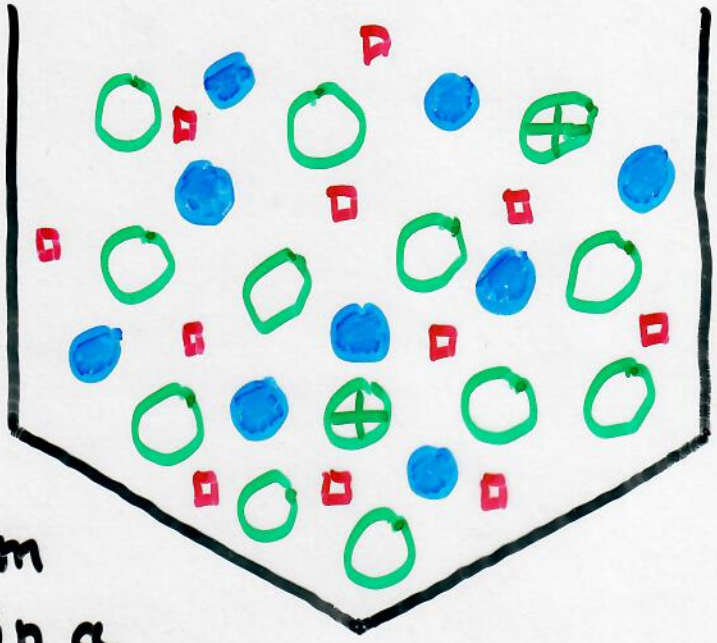
The assay :

Goal :

estimate the frequency of T-cells in a blood sample which respond to two test antigens

The real goal :

in the context of a vaccine trial, determine whether the vaccine causes an increase in the frequency of responding T-cells



V-bottom well in a microtiter plate

- antigen
- responding T-cell
- antigen-presenting cell
- other (non-responding) T-cells

combine
- diluted blood cells + growth medium
- antigen
- ^3H-thymidine

Replicating cells take up ^3H-thymidine

We extract the DNA and measure the amount of incorporated ^3H-thymidine using a scintillation counter

The usual setups:

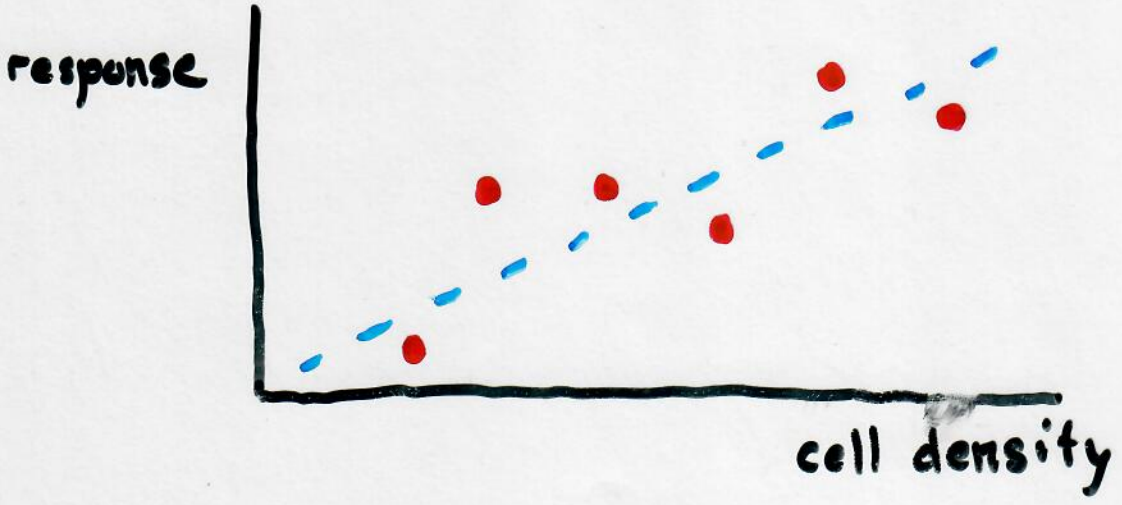
1. Use 3 wells with antigen and 3 wells without antigen

Stimulation index =

$$\frac{\text{ave of 3 counts with antigen}}{\text{ave of 3 counts without antigen}}$$

2. Limiting dilution assay

- several dilutions of cells
- many wells at each dilution



Our assay:

- stimulation index is too imprecise
- LDA requires too much blood

So, we study a single plate or pair of plates at a single, well-chosen dilution

8 × 12 microtiter plate

LDA 713, plates 1 and 2
11,400 cells per well

cells alone			gD2			gB2			Tetox		PHA
179	249	460	2133	2528	2700	2171	1663	6200	761	9864	12842
346	1540	306	8299	1886	3245	1699	2042	3374	183	7748	10331
117	249	1568	1174	4293	979	1222	1536	2406	6497	2492	6188
184	414	308	2801	2438	1776	2193	3211	1936	2492	5134	927
797	233	461	1076	1527	2866	2205	2278	2215	3725	3706	4050
305	348	480	3475	902	3654	2046	1285	1187	9899	5891	3646
1090	159	89	1472	90	3639	657	2393	1814	3330	4174	2389
280	571	329	4448	3643	881	3462	2118	1013	8793	4313	672

178	111	630	4699	5546	5182	3982	3104	2496	4275	2831	9727
244	593	259	5622	560	1073	1479	2978	4362	5017	5074	10706
261	964	167	2991	3390	3986	2321	2157	3278	8216	3579	3538
221	544	299	1838	4368	322	1022	1554	2980	2732	6177	5212
533	228	615	1938	4046	333	3253	5091	2843	200	1110	5063
818	98	160	1032	3269	4918	1778	3810	2372	6355	1869	2695
234	472	243	4143	3351	1118	530	1174	1881	3447	4491	2945
169	481	478	3237	1565	2211	2460	2715	4793	3029	6225	4679

Traditional method of analysis

- split wells into positives and negatives using a cutoff (e.g. mean + 3 SD of "cells alone" wells)

positive : one or more responding cells

negative : no responding cells

- imagine that the number of responding cells in a well is Poisson(λ_i) i = antigen group

$$\Pr(\text{no resp. cells}) = e^{-\lambda_i}$$

$$\hat{\lambda}_i = -\log\left[\frac{\#\text{ negative wells}}{\#\text{ wells}}\right]$$

cutoff: mean + 3SD of cells alone = 1401

LDA 713, plates 1 and 2
11,400 cells per well

cells alone			gD2			gB2			Tetox		PHA
179	249	460	2133	2528	2700	2171	1663	6200	761	9864	12842
346	1540	306	8299	1886	3245	1699	2042	3374	183	7748	10331
117	249	1568	1174	4293	979	1222	1536	2406	6497	2492	6188
184	414	308	2801	2438	1776	2193	3211	1936	2492	5134	927
797	233	461	1076	1527	2866	2205	2278	2215	3725	3706	4050
305	348	480	3475	902	3654	2046	1285	1187	9899	5891	3646
1090	159	89	1472	90	3639	657	2393	1814	3330	4174	2389
280	571	329	4448	3643	881	3462	2118	1013	8793	4313	672

178	111	630	4699	5546	5182	3982	3104	2496	4275	2831	9727
244	593	259	5622	560	1073	1479	2978	4362	5017	5074	10706
261	964	167	2991	3390	3986	2321	2157	3278	8216	3579	3538
221	544	299	1838	4368	322	1022	1554	2980	2732	6177	5212
533	228	615	1938	4046	333	3253	5091	2843	200	1110	5063
818	98	160	1032	3269	4918	1778	3810	2372	6355	1869	2695
234	472	243	4143	3351	1118	530	1174	1881	3447	4491	2945
169	481	478	3237	1565	2211	2460	2715	4793	3029	6225	4679

$\hat{q}$: 46/48 12/48 8/48 6/44

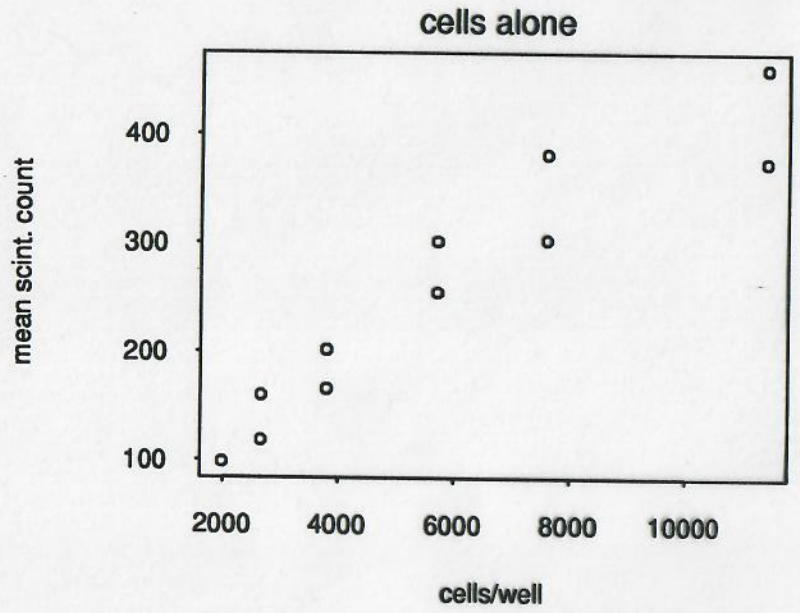
$\hat{\lambda} = -\log \hat{q}$: 0.04 1.39 1.79 1.99

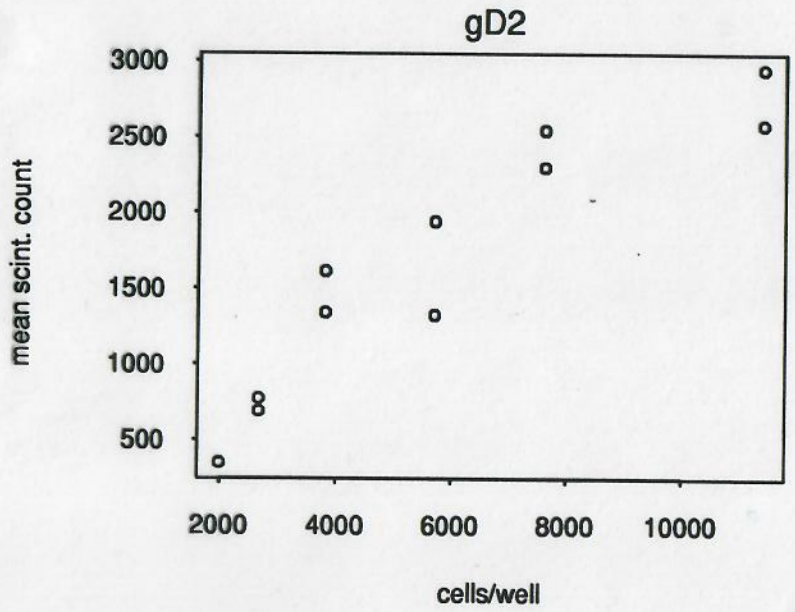
$\frac{\hat{\lambda}_{adj} \times 10^6}{11,400}$: — 118 153 171

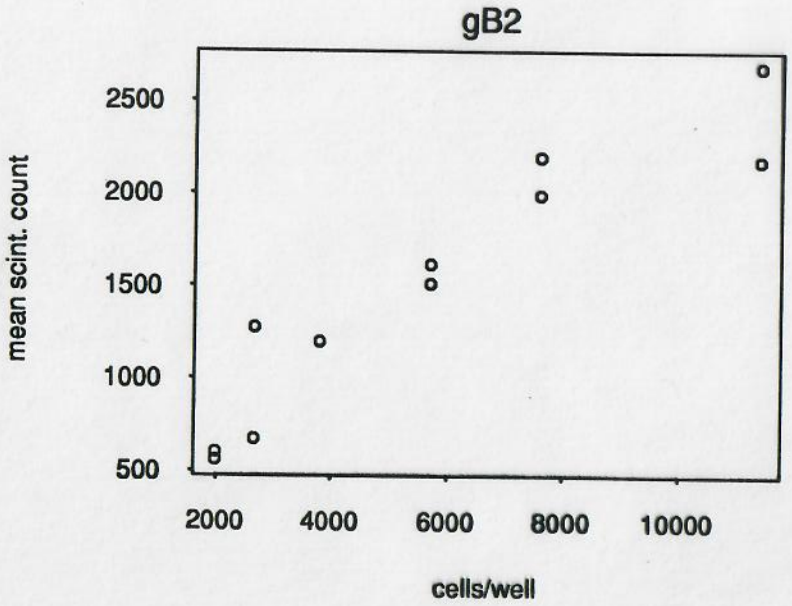
Problems with the cutoff method

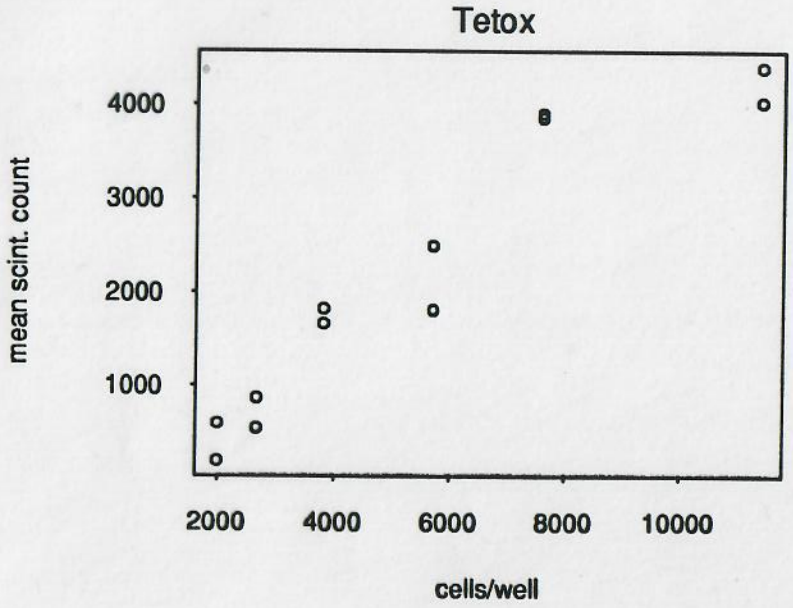
- it can be difficult to choose a cutoff
- need to avoid all (+) or all (-)
- using only a single dilution – can we get more information from the data?

LDA #713


cells alone
mean scint. count
400
300
200
100
2000
4000
6000
8000
10000
cells/well


gD2
mean scint. count
3000
2500
2000
1500
1000
500
2000
4000
6000
8000
10000
cells/well


gB2
mean scint. count
2500
2000
1500
1000
500
2000
4000
6000
8000
10000
cells/well


Tetox
mean scint. count
4000
3000
2000
1000
2000
4000
6000
8000
10000
cells/well

Our model

(unobserved) k_{ij} = # of responders in well j of group i

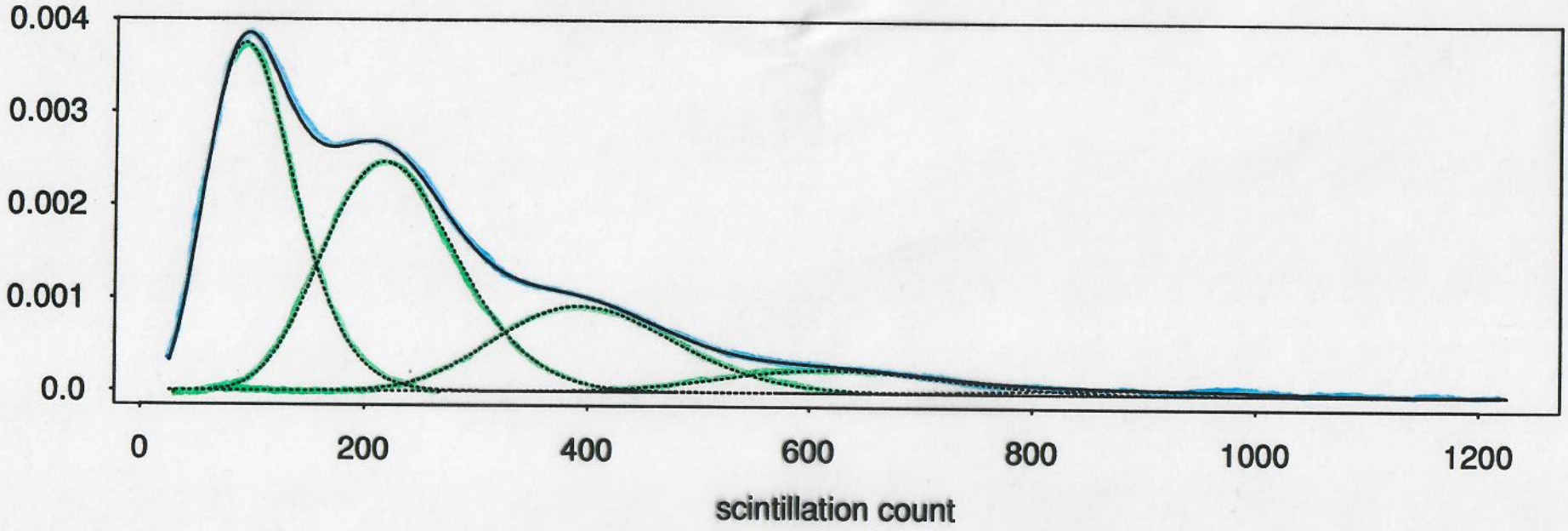
(our data) y_{ij} = transformed scint. count for well j of group i

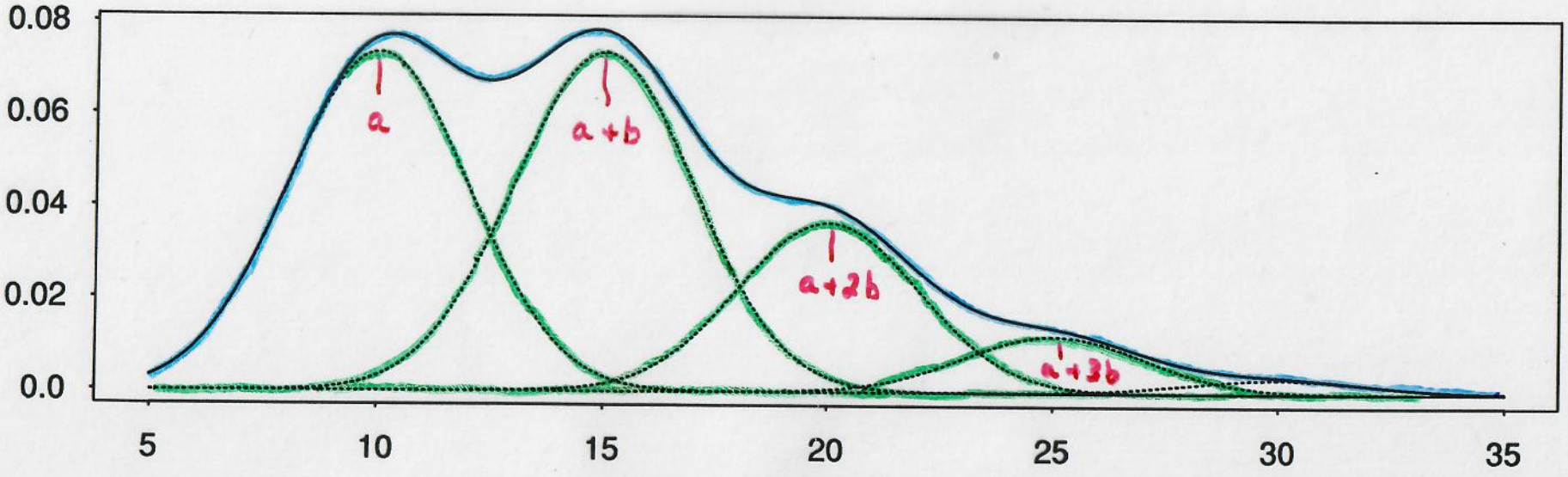
Assume

$$k_{ij} \sim \text{Poisson}(\lambda_i)$$

$$y_{ij} \mid k_{ij} \sim \text{normal}(a + bk_{ij}, \sigma^2)$$

(y_{ij}, k_{ij}) mutually independent





$e^{-\lambda}$ $\lambda e^{-\lambda}$ $\lambda^2 e^{-\lambda}/2$ $\lambda^3 e^{-\lambda}/6$

Parameters: $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$

(a, b, σ) for each plate

Estimation: ML via the EM alg.

Choice of Transformation: Box-Cox analysis

Estimated SEs: SEM algorithm (Meng & Rubin, 1991)

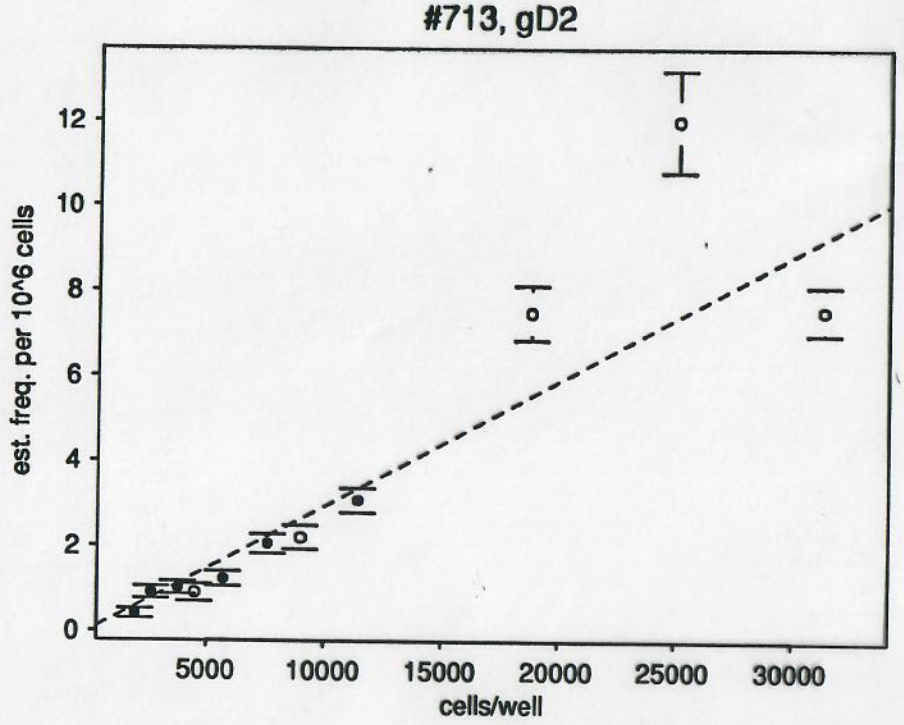
LDA 713, plates 1 and 2
11,400 cells per well

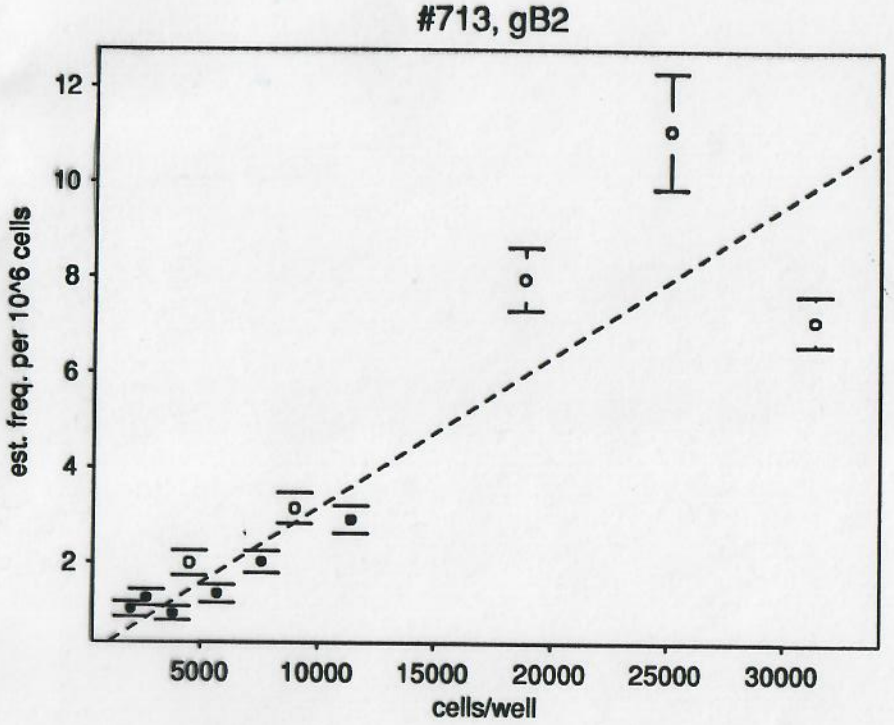
Estimated frequencies of responding cells per well

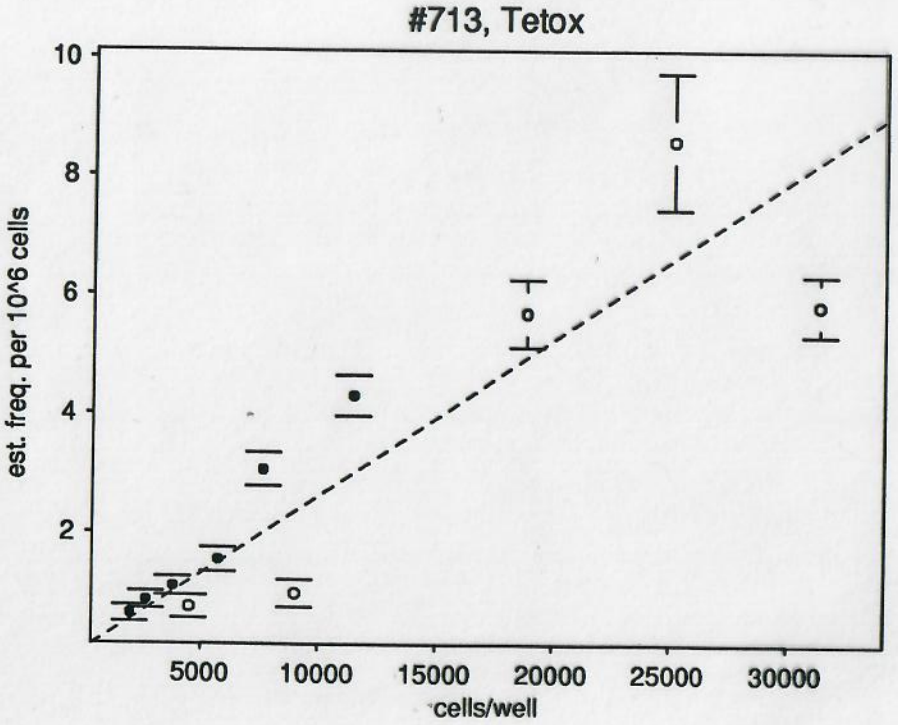
	cells alone	gD2	gB2	Tetox
joint	0.4 (0.1)	3.5 (0.3)	3.3 (0.3)	4.7 (0.3)
separate:				
plate 1	0.3 (0.1)	3.0 (0.4)	2.8 (0.4)	4.4 (0.5)
plate 2	0.5 (0.1)	3.9 (0.4)	3.9 (0.4)	5.0 (0.5)

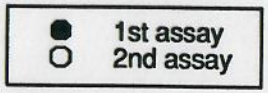
Estimates of other parameters

	a	b	σ
joint:			
plate 1	16.4 (0.9)	10.3 (0.3)	3.6 (0.5)
plate 2	14.8 (0.8)	9.4 (0.2)	2.9 (0.4)
separate:			
plate 1	16.7 (0.9)	10.3 (0.3)	3.5 (0.4)
plate 2	14.5 (0.7)	9.3 (0.2)	2.8 (0.3)


#713, gD2
est. freq. per 10^6 cells
cells/well
0
2
4
6
8
10
12
5000
10000
15000
20000
25000
30000


#713, gB2
est. freq. per 10^6 cells
cells/well
2
4
6
8
10
12
5000
10000
15000
20000
25000
30000


#713, Tetox
est. freq. per 10^6 cells
cells/well
2
4
6
8
10
5000
10000
15000
20000
25000
30000


1st assay
2nd assay

Problems

- I don't know anything
- Jargon
- Maintaining confidence
- Avoiding frustration
- What is the question?
- Things get complicated; what can we ignore?
- Will the fancy analysis really improve on the crude one?

Benefits

- learn science
- learn statistics
- help people
- meet people

Strategies

1. First meeting

- Ask as many questions as possible (even "obvious" ones)
  - the question
  - the data
  - possible difficulties

- Gauge mathematical sophistication (of client, of field)

- Ask about the usual methods of analysis (get references)

2. In between

- Talk to many people

- Formulate ideas & more questions

3. Later

- Tailor methods to the science
- Keep things simple
- Describe things simply
- Always ask more questions