AptivSolutions℠
Accelerating the Possibilities

Dose Selection in Drug Development: What Can Go Wrong? Can we put it Right?

Professor Andy Grieve
SVP Clinical Trials Methodology
Innovation Centre, Aptiv Solutions GmbH
Cologne, Germany

Outline

- Regulatory attitude to dose-response information
- Phase 2b Dose-Response/Finding/Selection Designs
  - Number of Doses and Dose-Interval
  - Adaptive strategies
- Emax Models
  - Issues in fitting
  - Non-monotonic models
- Smoothing approaches
- Slowness of translation

Regulatory Environment - Background

- There are two distinct approaches (objectives) to dose-response studies
  - Estimating the dose-response relationship
  - Determining which doses are “signficantly different” from placebo
- Regulatory guidance is somewhat contradictory in terms of its attitudes to the appropriateness of these approaches

Regulatory Environment

- ICH E9, Section 3.3.3 : **Trials to Show Dose-response Relationship**
  - *…the application of procedures to estimate the relationship between dose and response, including the construction of confidence intervals and the use of graphical methods, is as important as the use of statistical tests.*
- ICH E10, Section 2.3.1 : **Dose-response Concurrent Control (See Section 1.3.3) - *Description***
-
  - *A dose-response study is one in which subjects are randomly assigned to two or more dosage groups, with or without a placebo group. Dose-response studies are carried out to **establish the relation between dose and efficacy** and adverse effects and/or to **demonstrate efficacy**. The first use is considered in ICH E4; the use to demonstrate efficacy is the subject of this guidance. Evidence of efficacy could be based on **significant differences in pair-wise comparisons** between dosage groups or between dosage groups and placebo, or on evidence of a **significant positive trend with increasing dose**, even if no two groups are significantly different. In the latter case, however, further study may be needed to assess the effectiveness of the low doses. As noted in ICH E9, the particular approach for the primary efficacy analysis should be prespecified.*

Dose-Response Estimation

- ICH E4 Section IV:
  - *Several dose levels are needed, at least two in addition to placebo, but in general, study of more than the minimum number of doses is desirable. A single dose level of drug versus placebo allows a test of the null hypothesis of no difference between drug and placebo, but cannot define the dose- response relationship. Similarly, although a linear relationship can be derived from the response to two active doses (without placebo), this approximation is usually not sufficiently informative.* ***Study designs usually should emphasize elucidation of the dose-response function, not individual pair-wise comparisons.*** *If a particular point on the curve, e. g., whether a certain low dose is useful, becomes an issue, it should be studied separately.*

Requirement of statistical significance

- ICH E4 Section III:
  - *In principle, being able to detect a statistically significant difference in pair-wise comparisons between doses is not necessary if a statistically significant trend (upward slope) across doses can be established using all the data. It should be demonstrated, however, that the lowest dose(s) tested, if it is to be recommended,* ***has a statistically significant and clinically meaningful effect.***
  - The last phrase MAY mean that additional studies are required to establish it (cf previous 2 slides)

Requirement of statistical significance

- Points to Consider on Multiplicity Issues in Clinical Trials Section 2.5.3:
-
  - *For therapeutic dose response studies that aim at identifying one or several doses of an investigational drug for its recommended use in a specific patient population, the control of the family-wise type I error in the strong sense is mandatory. …*
  - *Sometimes a study is not powered sufficiently for the aim to identify and recommend a single effective and safe dose (or a dose range) but is successful only at demonstrating an overall positive correlation of the clinical effect with increasing dose. This is already a valuable achievement. Estimates and confidence intervals from pairwise comparisons of single doses are then used in an exploratory manner for the planning of future studies. In this case, an adjustment of the type I error is not necessary.*

Primary estimands in dose response studies

- *E4 Introduction*
  - *Historically, drugs have often been initially marketed at what were later recognized as excessive doses (i. e., doses well onto the plateau of the dose- response curve for the desired effect), sometimes with adverse consequences (e. g., hypokalemia and other metabolic disturbances with thiazide- type diuretics in hypertension).* ***This situation has been improved by attempts to find the smallest dose with a discernible useful effect or a maximum dose beyond which no further beneficial effect is seen, but practical study designs do not exist to allow for precise determination of these doses.***
  - This latter point may be related to a distinction between finding the effect at a given dose and finding the doses that deliver a given effect size (cf Hemmings, section 2.5.2)

Learn vs Confirm in Phase 2

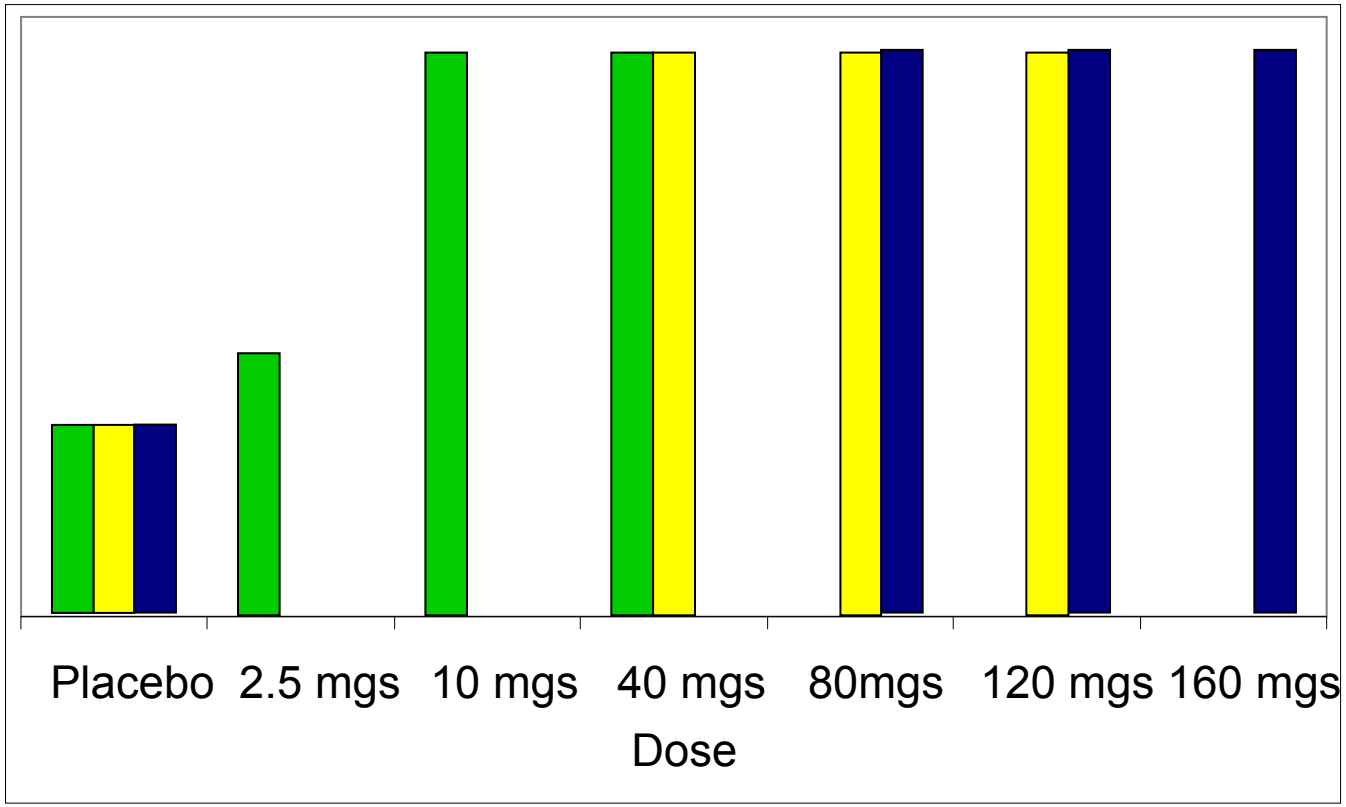
- Sheiner (1997)
  - *Current dose- ranging studies typically assign patients to one of a small number of possible dose magnitudes or to placebo. Outcome is observed after some time. Analysis is by intention to treat, and if the hypothesis that the outcomes differ among arms is rejected, then the minimum effective dose is taken as the minimum magnitude studied dose that demonstrates a significant outcome difference from placebo. This approach has at least the following faults: …..*
  - *All of these faults are directly attributable to the fact that confirming is being done when learning is needed. A first step toward learning is to take advantage of the fact that the dose effect relationship within an individual has a fairly predictable shape: it starts at zero and smoothly increases to an asymptote.*

Phase 2b Dose-Response/Finding/Selection Designs

Phase 2b Dose Response/Finding/Selection Designs
1st Example

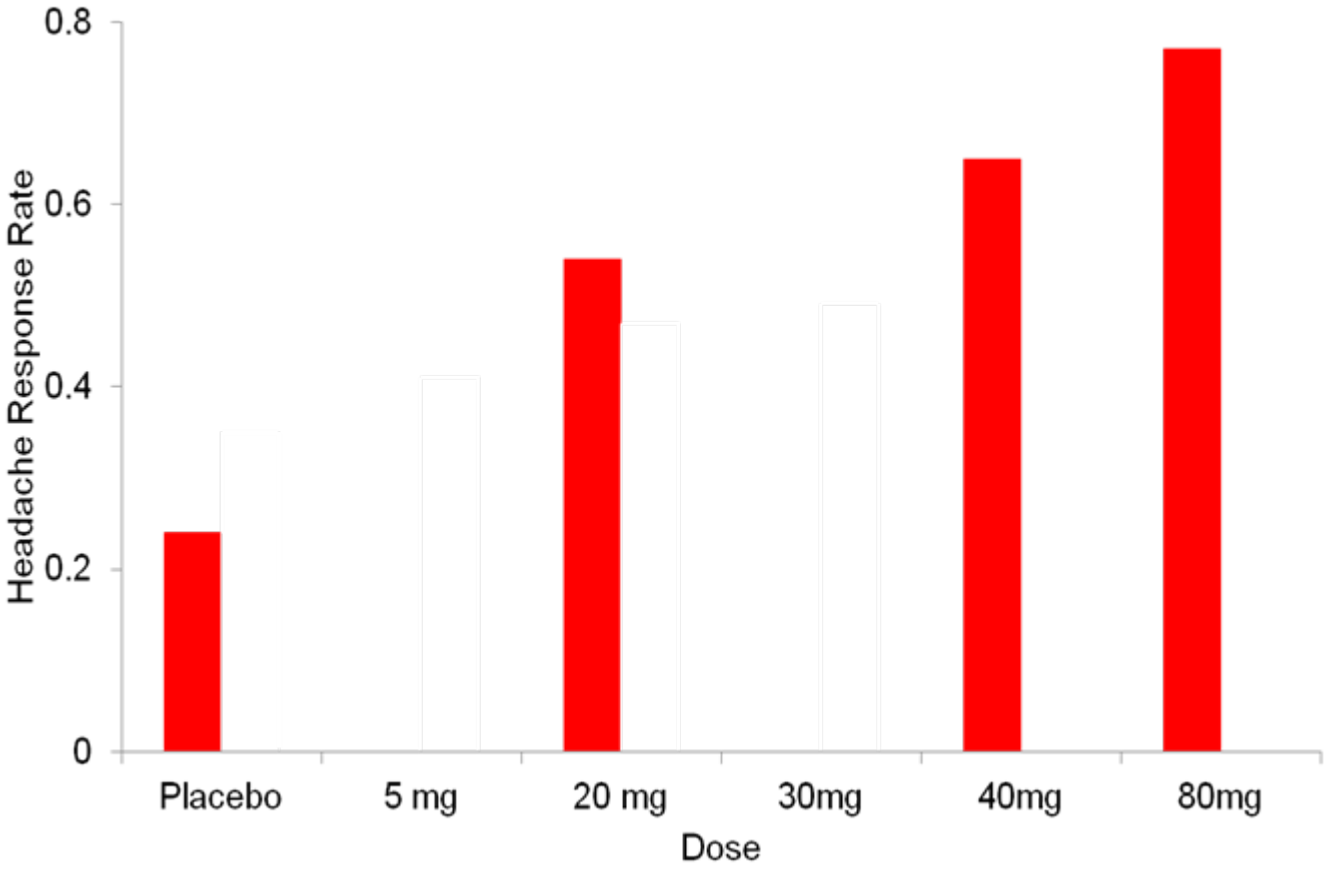
- Development in osteoarthritis
- 1st Cycle - pla, 80 mg, 120 mg, 160 mg (x2)
  - All 3 doses better than placebo, no differences between them
  - Doses based on pre-clinical data
- 2nd cycle - pla, 40mg, 80 mg, 120 mg (x4)
  - All 3 doses better than placebo, no differences between them
- 3rd Cycle – pla, 2.5 mg, 10mg, 40mg (x 64)
  - 2.5mg not different from placebo

Phase 2b Dose Selection Design Circa 1993



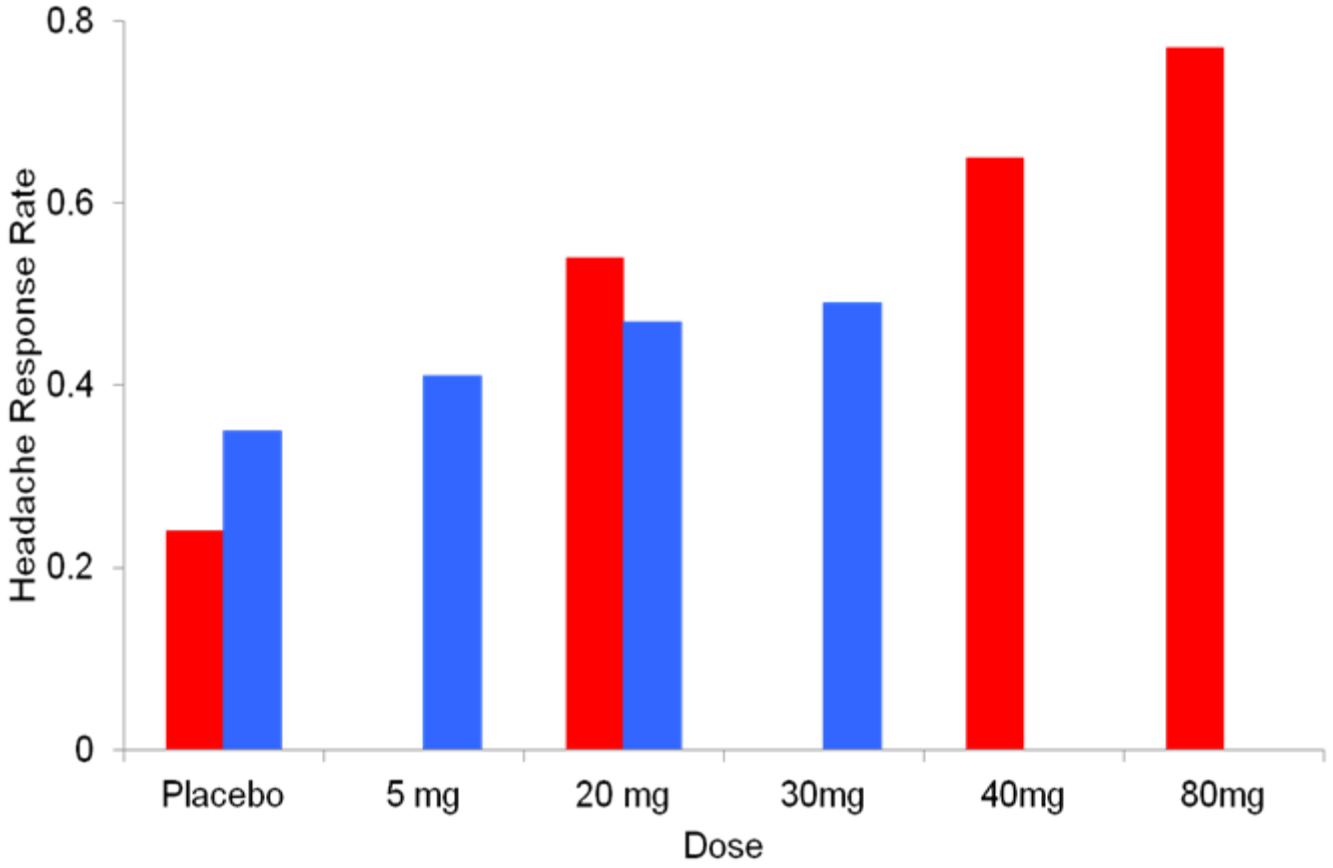
- More Efficient
  - wide range of doses, smaller numbers of patients per group
  - followed by one large parallel group study focusing on the doses showing promise in exploratory study.

Phase 2b Dose Response/Finding/Selection Designs
2nd Example -1st Cycle



All doses significantly different from Placebo

Phase 2b Dose Response/Finding/Selection Designs
2nd Example - 2nd Cycle



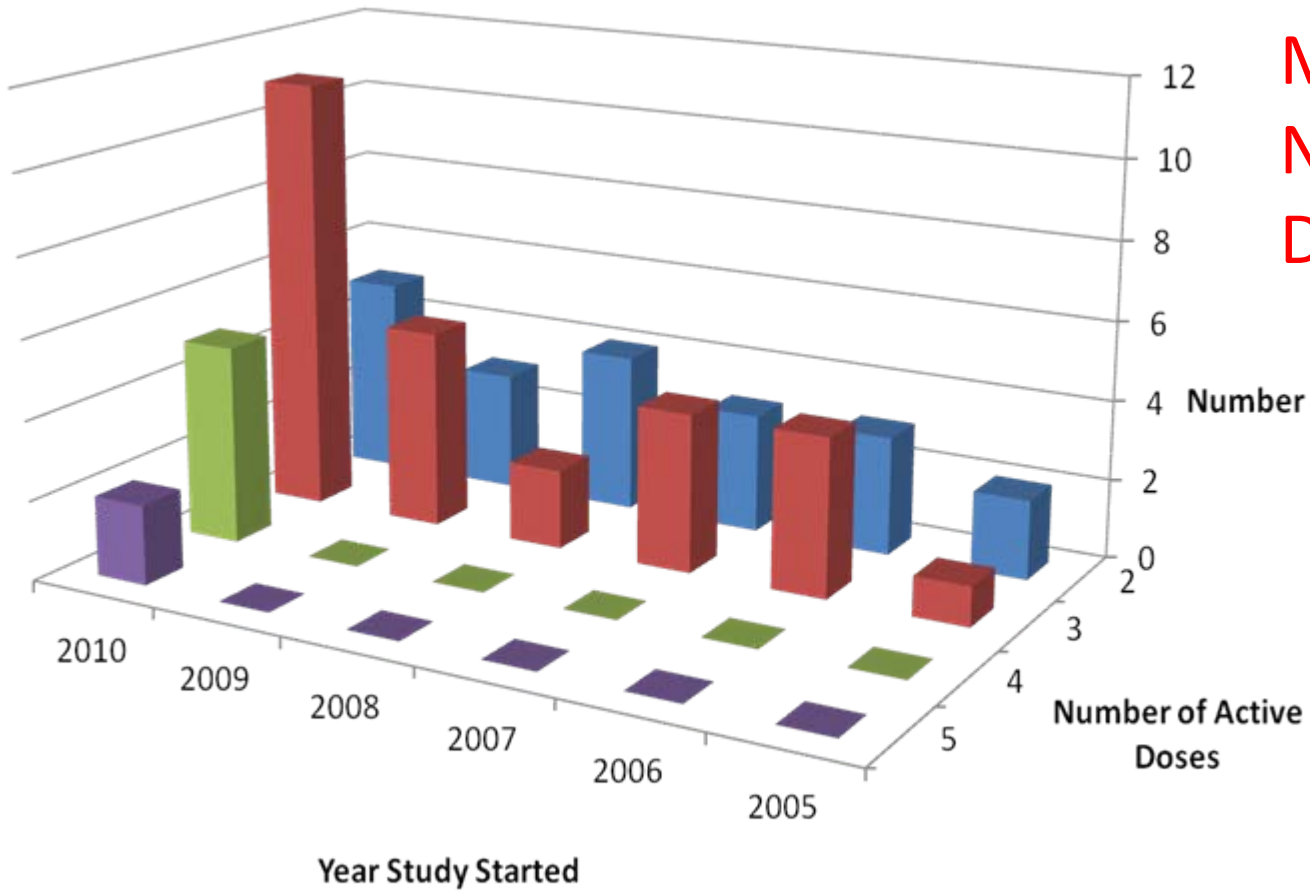
20mg, 30mg significantly different from Placebo,
5 mg not significant – 16x increase in dose

2

Comparison Between Successful and Unsuccessful Phase II Programs

Initial Dose Finding Unsuccessful - More Studies Required			Initial Dose Finding Successful	
Study	Initial Dose Range	Total Dose Range Examined	Study	Dose Range Examined
1	4	64	1	40
2	1	4	2	8
3	6	16	3	4
4	4	8	4	10
			5	4
Median	4	12	Median	8

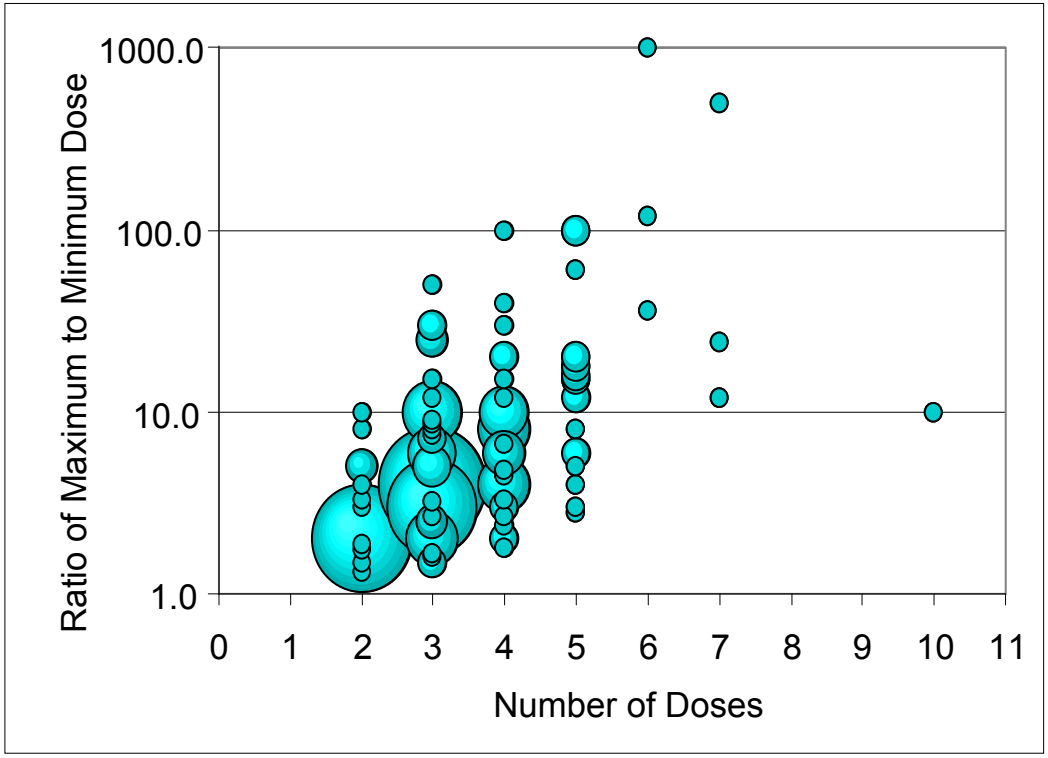
Phase 2b Dose Response/Finding/Selection Designs



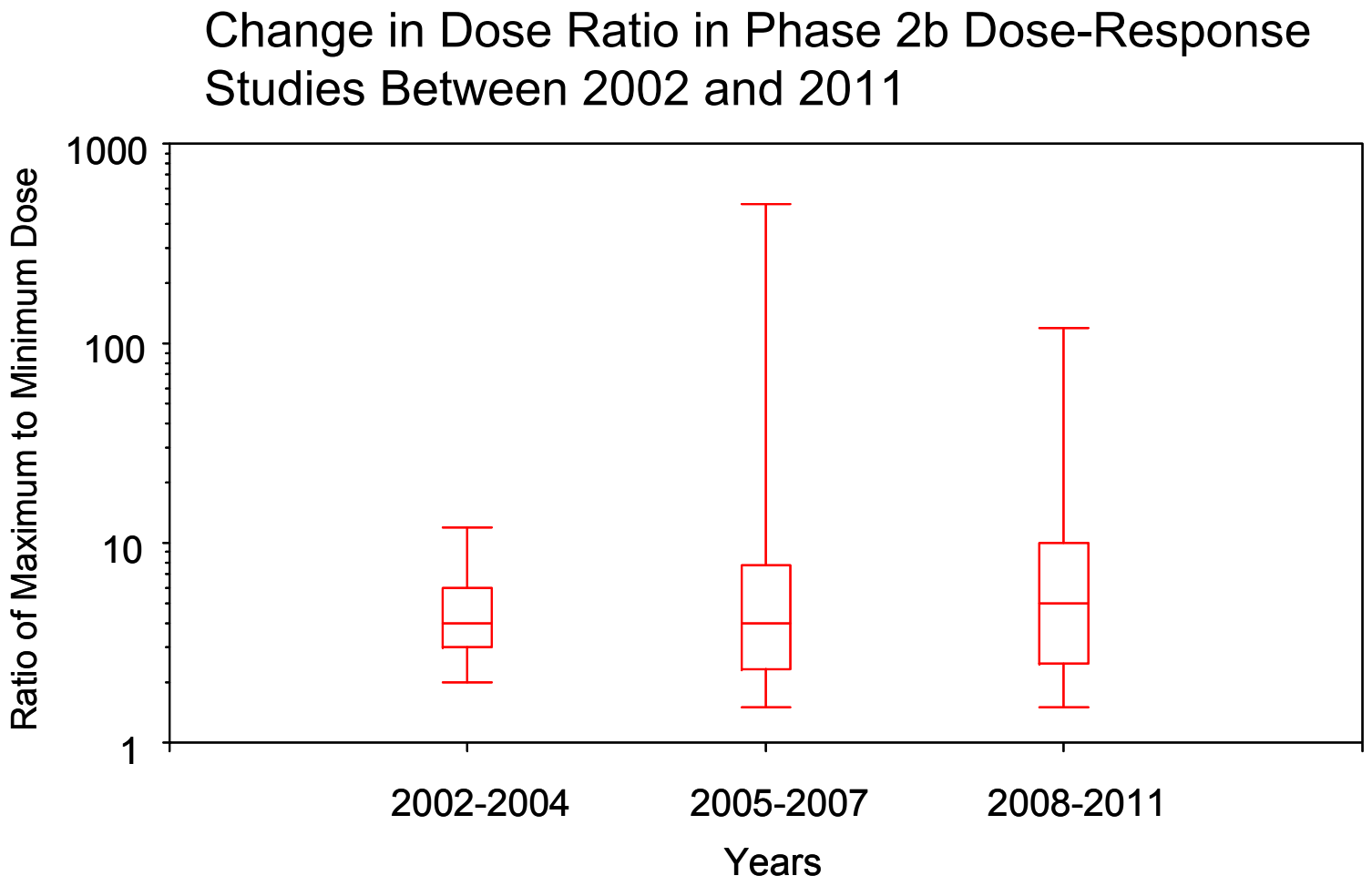
Median Number of Doses = 3

Phase 2 Adult Dose Response/Finding/Selection/Ranging Designs

Relationship Between Dose Ratio & Number of Doses

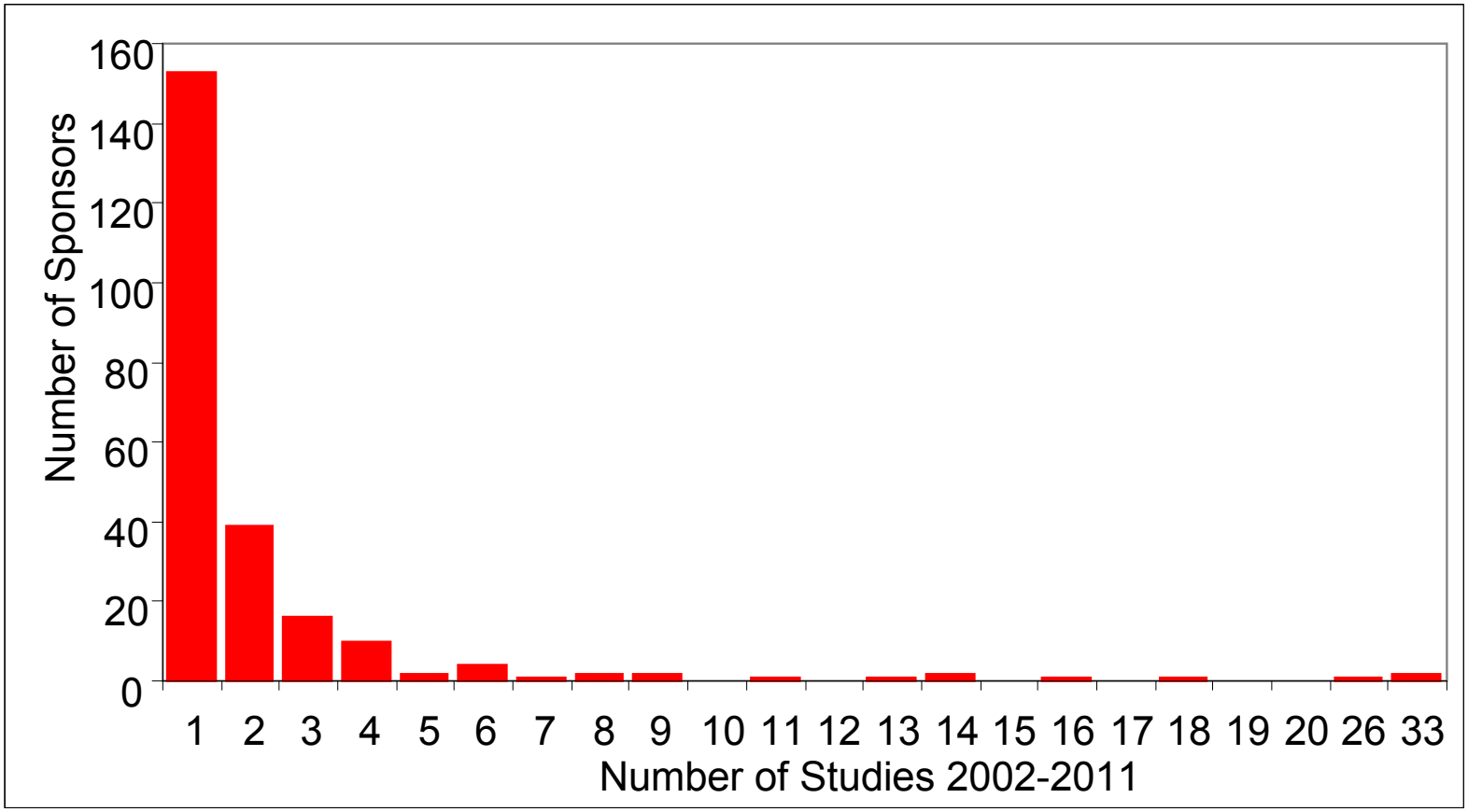


Phase 2 Adult Dose Response/Finding/Selection/Ranging Designs



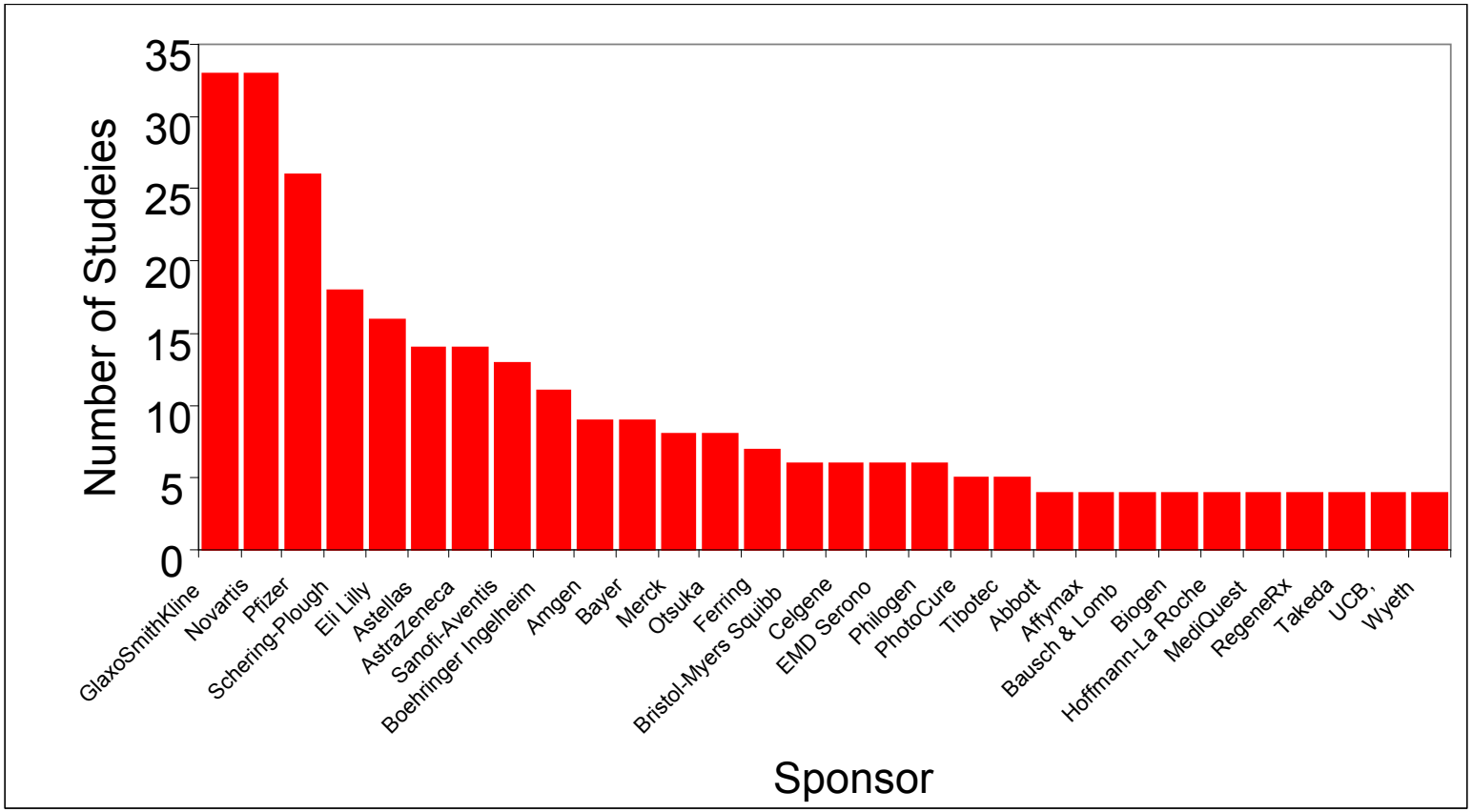
Dose-Response Studies 2002-2011

Numbers of Sponsors Sponsoring multiple studies
Between 2002 and 2011



Dose-Response Studies 2002-2011

Numbers of Phase 2b Studies Sponsored by the Top 20 Sponsors Between 2002 and 2011

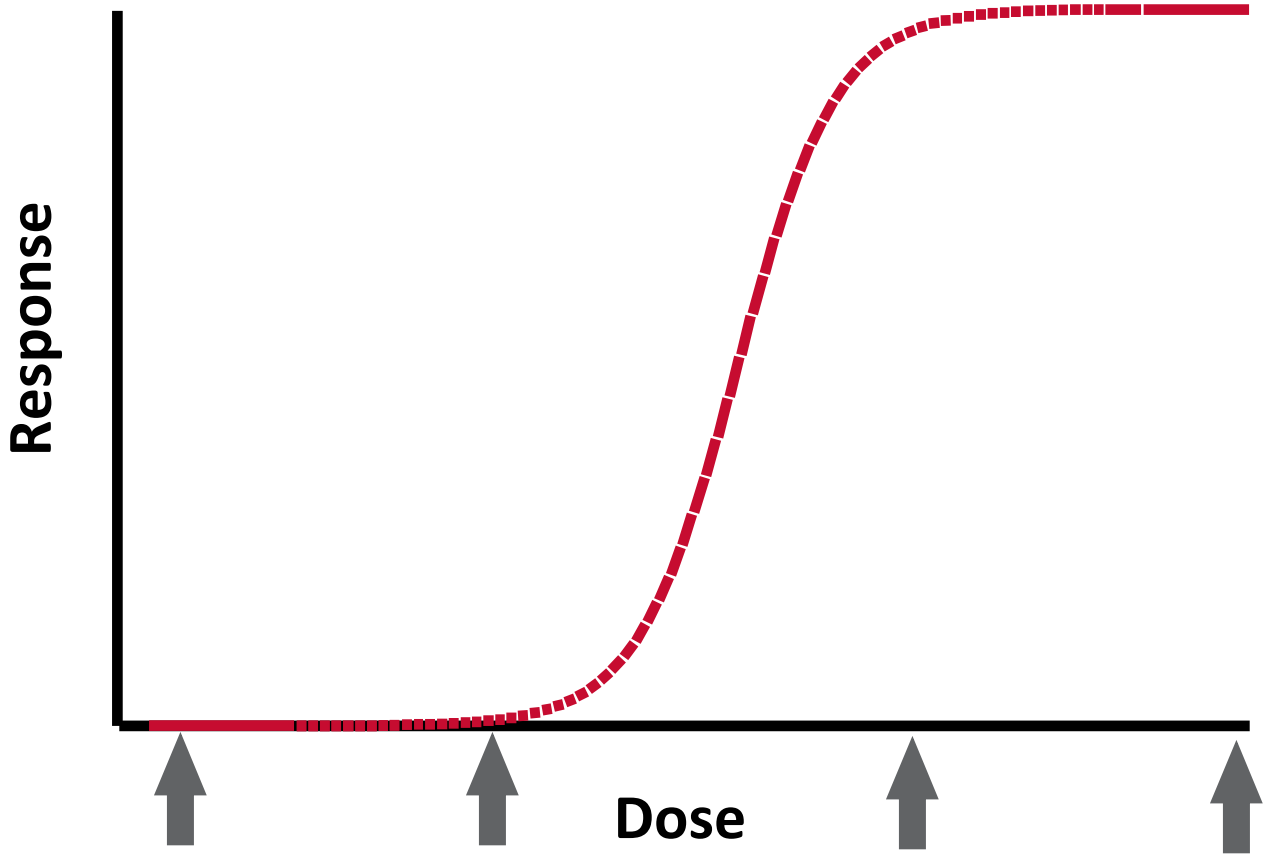


Deficiencies in Clinicaltrials.gov

- Incompleteness
  - Missing Information
    - Number of arms
    - Individual dose level
- Accessibility
  - Limited ability to download information
  - 20 fields are available
  - Does not include information on treatment arms and doses

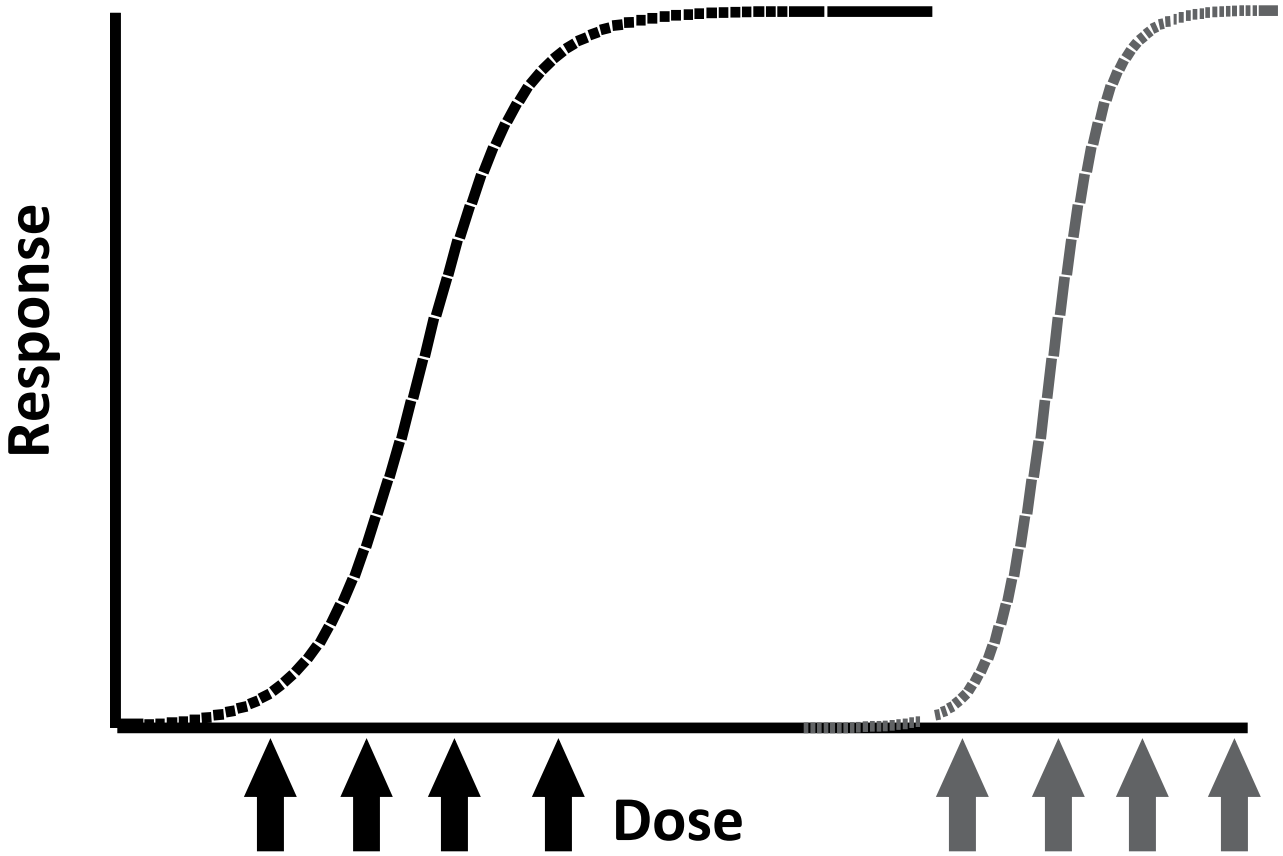
Phase 2b Dose Response/Finding/Selection Designs

Standard Design



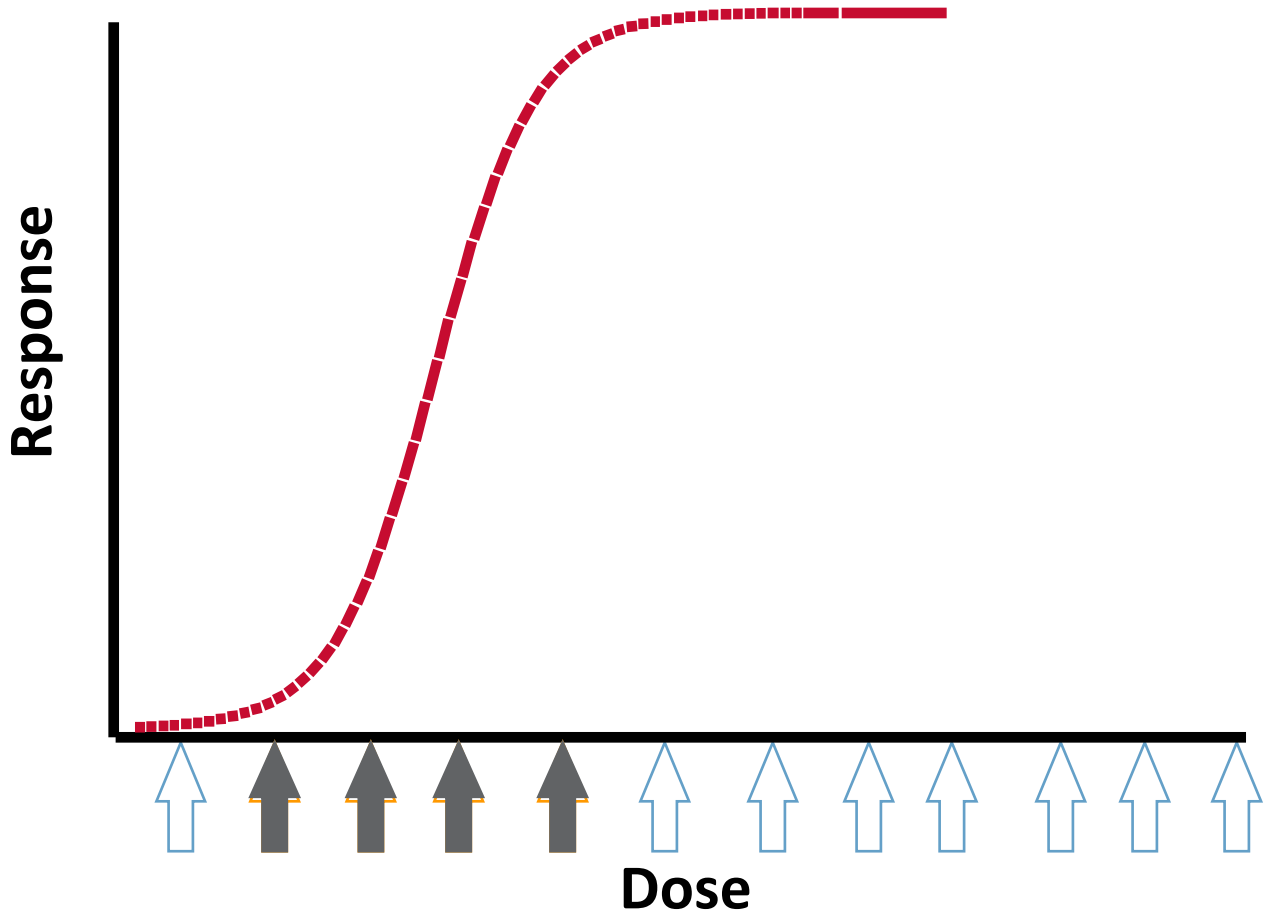
Phase 2b Dose Response/Finding/Selection Designs

Placebo + 4 doses available where to put them ?



Phase 2b Dose Response/Finding/Selection Designs

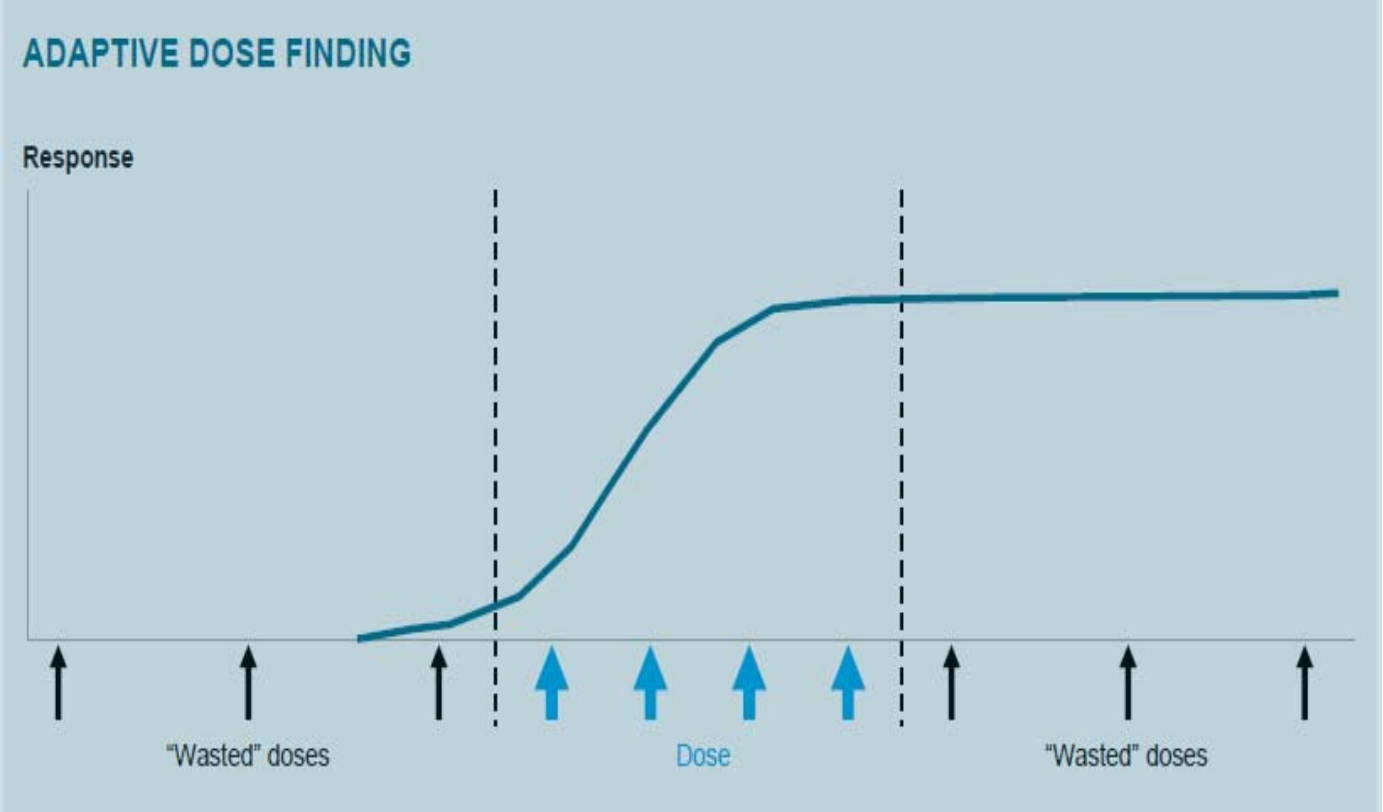
Choose Many Doses & Adapt



- Increase # of doses
- Adapt to steep part of dose response curve
- Concentrate on estimation rather than comparing individual doses to placebo
- Use of Bayesian Methods

Phase 2b Dose Response/Finding/Selection Designs

Choose Many Doses & Adapt



Invention Reinvented, McKinsey Perspectives on Pharmaceutical R&D 2010

Adaptive Dose-Response Pfizer Example

- Setting: Venous thrombo-embolism (VTE) prophylaxis in patients undergoing an elective total knee replacement
- PD 0348292: an oral direct factor Xa inhibitor
- Dose selection critical for an anticoagulant
  - Under-dosing: increased risk of thrombosis
  - Over-dosing: increased risk of bleeding
- Objective of Phase 2b dose-ranging trial
  - Find a dose equivalent to the current standard of care, enoxaparin 60 mg/day

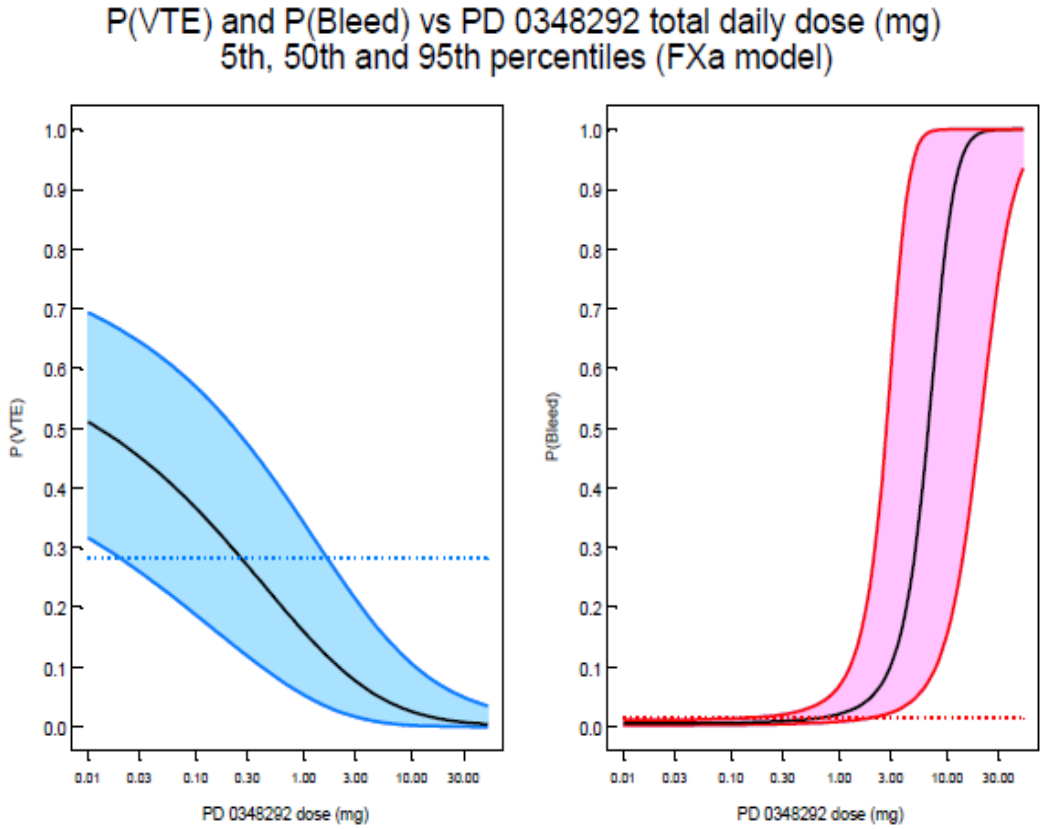
Richard Lalonde, Clinical Pharmacology, Pfizer Inc
Pharmaceutical Sciences World Congress, New Orleans, November 2010

Adaptive Dose-Response Pfizer Example

- Biomarker:
  - Inhibition of thrombin generation
- Literature Data:
  - Clinical outcome (incidence of VTE and major bleeding [MB]) for comparator anticoagulants
- Model:
  - Linked biomarker response and clinical outcome for comparators with an integrated PK-PD model
- Estimated Dose:
  - Predicted VTE and MB dose-response for PD 0348292 based on its biomarker response

Richard Lalonde, Clinical Pharmacology, Pfizer Inc
Pharmaceutical Sciences World Congress, New Orleans, November 2010

Predicted PD 0348292 Dose-Response Relationships for VTE & MB



Richard Lalonde, Clinical Pharmacology, Pfizer Inc
Pharmaceutical Sciences World Congress, New Orleans, November 2010

Adaptive Dose-Response Pfizer Example

- 6-arm randomized, parallel group study with adaptive dose range based on interim dose decision analyses of VTE and M
  - Start with 5 doses of PD 0348292 (0.1 to 2.5 mg QD)
  - Eliminate PD 0348292 doses based on excessive VTE or MB
  - Add higher PD 0348292 doses (4 and 10 mg QD) if we eliminate lower doses and MB rate acceptable
  - Enoxaparin 30 mg BID as control
- Dose decision interim analyses (dose-response regression model) after every 147 evaluable patients
- Total sample size of 1250 patients

Richard Lalonde, Clinical Pharmacology, Pfizer Inc
Pharmaceutical Sciences World Congress, New Orleans, November 2010

Adaptive Dose-Response Pfizer Example

- Study designed using M&S was approved by senior management and conducted successfully
- Study met key objective
  - Identified the dose equivalent to enoxaparin with good precision
- Safely explored a 100-fold dose range to allow characterization of dose-response relationship for efficacy (vs ~ 4-fold dose range for competitors
  - ~1/3 sample size of traditional parallel group study
  - Savings of 2750 patients
  - Savings >$20M in trial costs
  - Shortened development time by I year

Richard Lalonde, Clinical Pharmacology, Pfizer Inc
Pharmaceutical Sciences World Congress, New Orleans, November 2010

Emax Models

Emax Dose Response Models

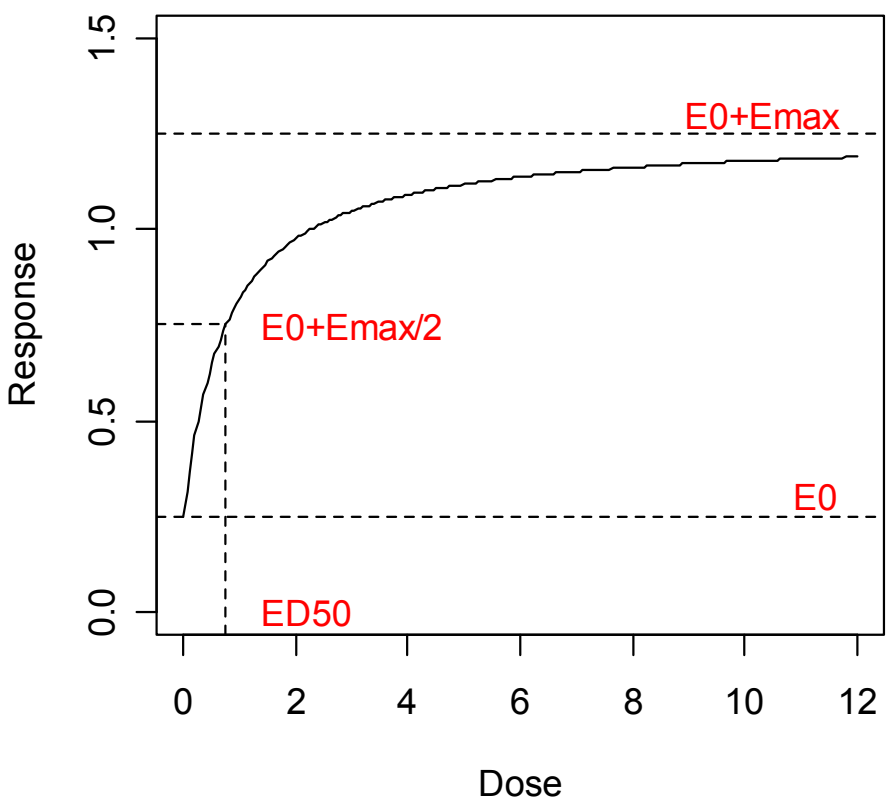
- Hyperbolic Emax model

$$E(Y|D) = E_0 + \frac{E_{max} D}{ED_{50} + D}$$

- Sigmoid Emax model

$$E(Y|D) = E_0 + \frac{E_{max} D^{\lambda}}{ED_{50}^{\lambda} + D^{\lambda}}$$

Hyperbolic Emax model (3 Parameters)



E_0 = baseline (response in absence of drug)

Emax = difference between maximum achievable response (at infinite dose) and baseline; also known as “efficacy”

ED50 = concentration that produces half-maximal effect; also known as “potency” ; > 0

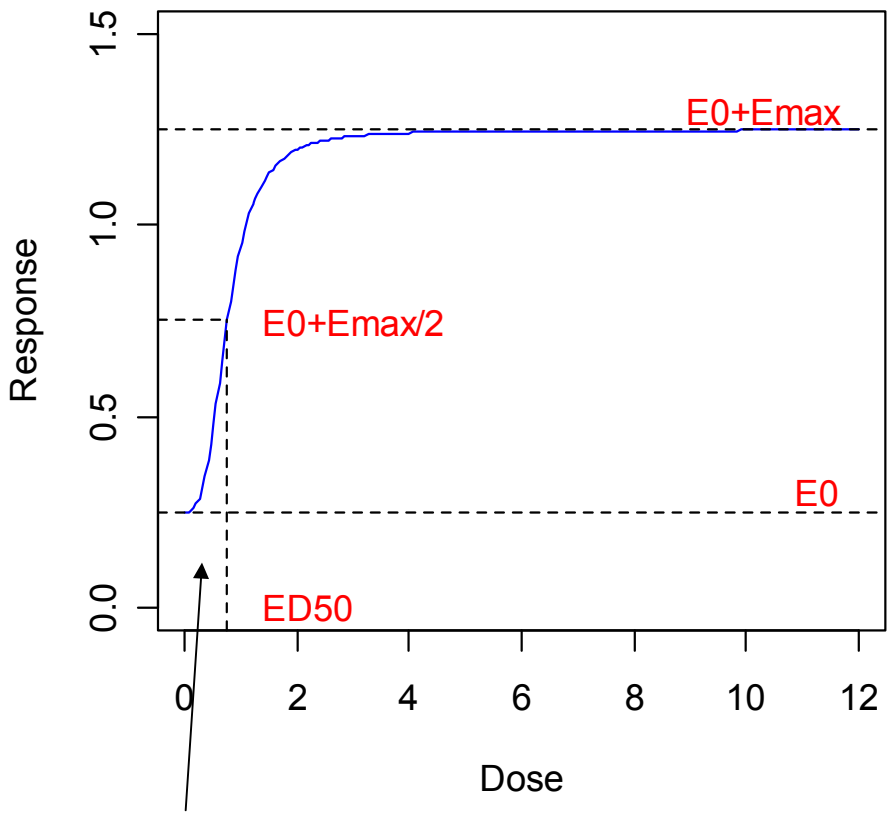
Can represent decreasing response with increasing dose with a negative Emax parameter

Other Hyperbolic Models

- Emax model for dose is an empirical analogue of Emax concentration model
- Concentration version of Emax model has a theoretical basis connected to chemical law of mass action and rates of saturable enzyme reactions in receptor models
  - With low exposure variability and proportional PK, the Emax is also implied for dose response
- The hyperbolic model also used for many other clinical pharmacological applications
  - inhibition, clearance, Michaelis Menton enzyme kinetics, protein binding, etc.

Four Parameter Emax Model
Hill Parameter = 3

$$E(Y|D) = E_0 + \frac{E_{max} D^{\lambda}}{ED_{50}^{\lambda} + D^{\lambda}}$$

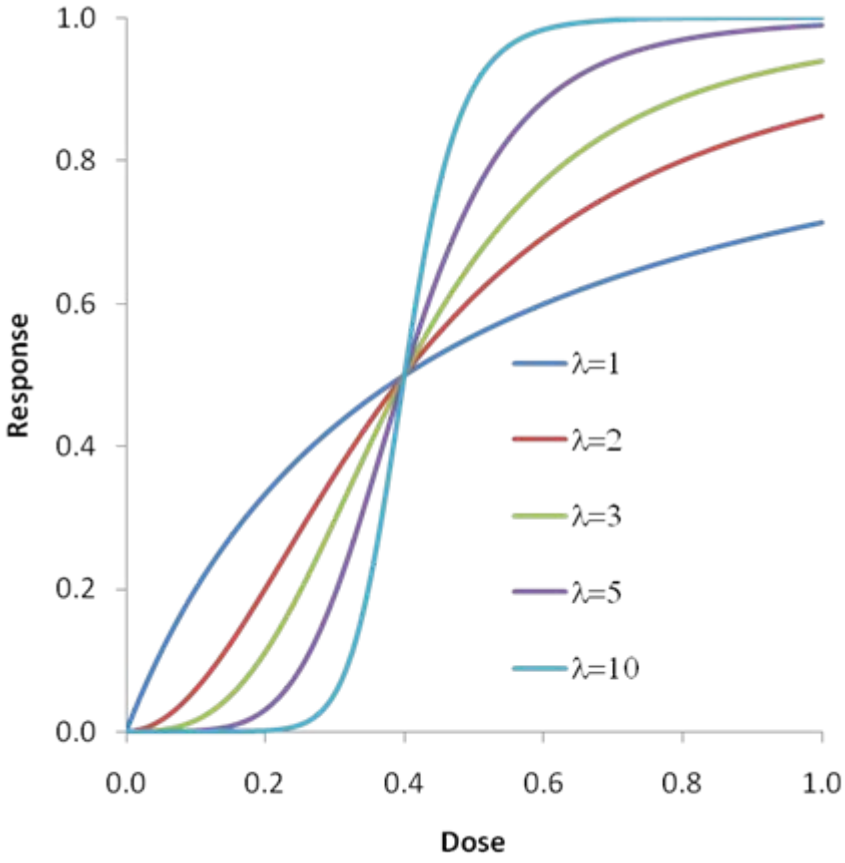


Threshold

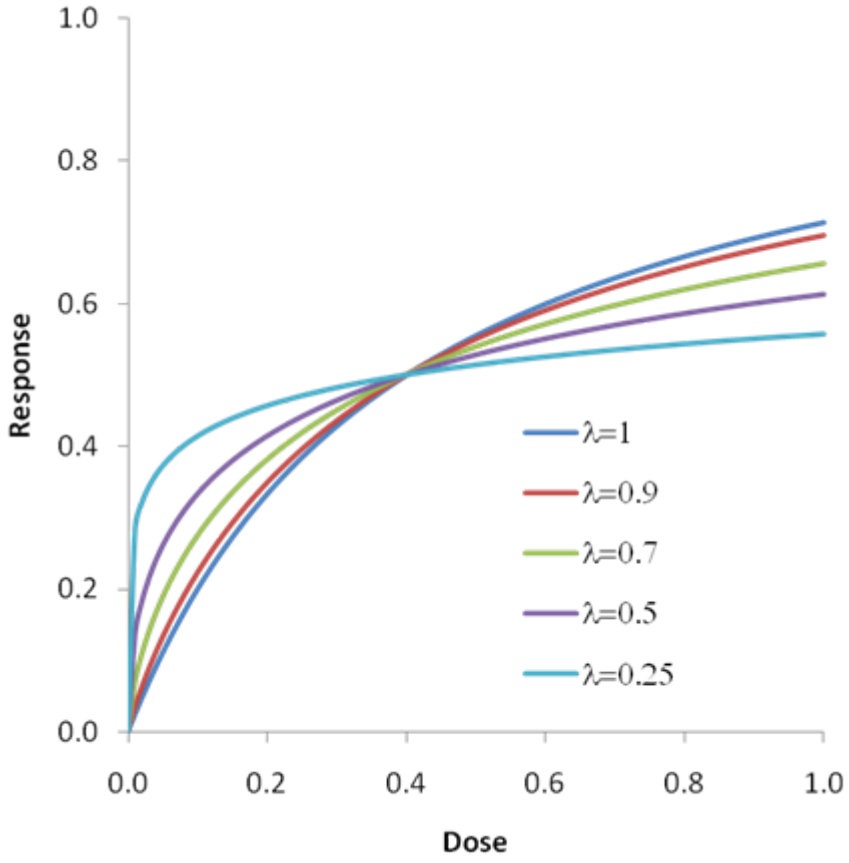
- The power parameter is called the "Hill" parameter
- The model can be re-expressed as a logistic model
- Addition of "sigmoidisity (Hill) parameter" Lambda
  - Varies the steepness of ascent to Emax
  - Induces "threshold" (when lambda > 1) which can be pronounced

Flexibility of the 4-parameter Emax Model

Lamda >1



Lamda <1

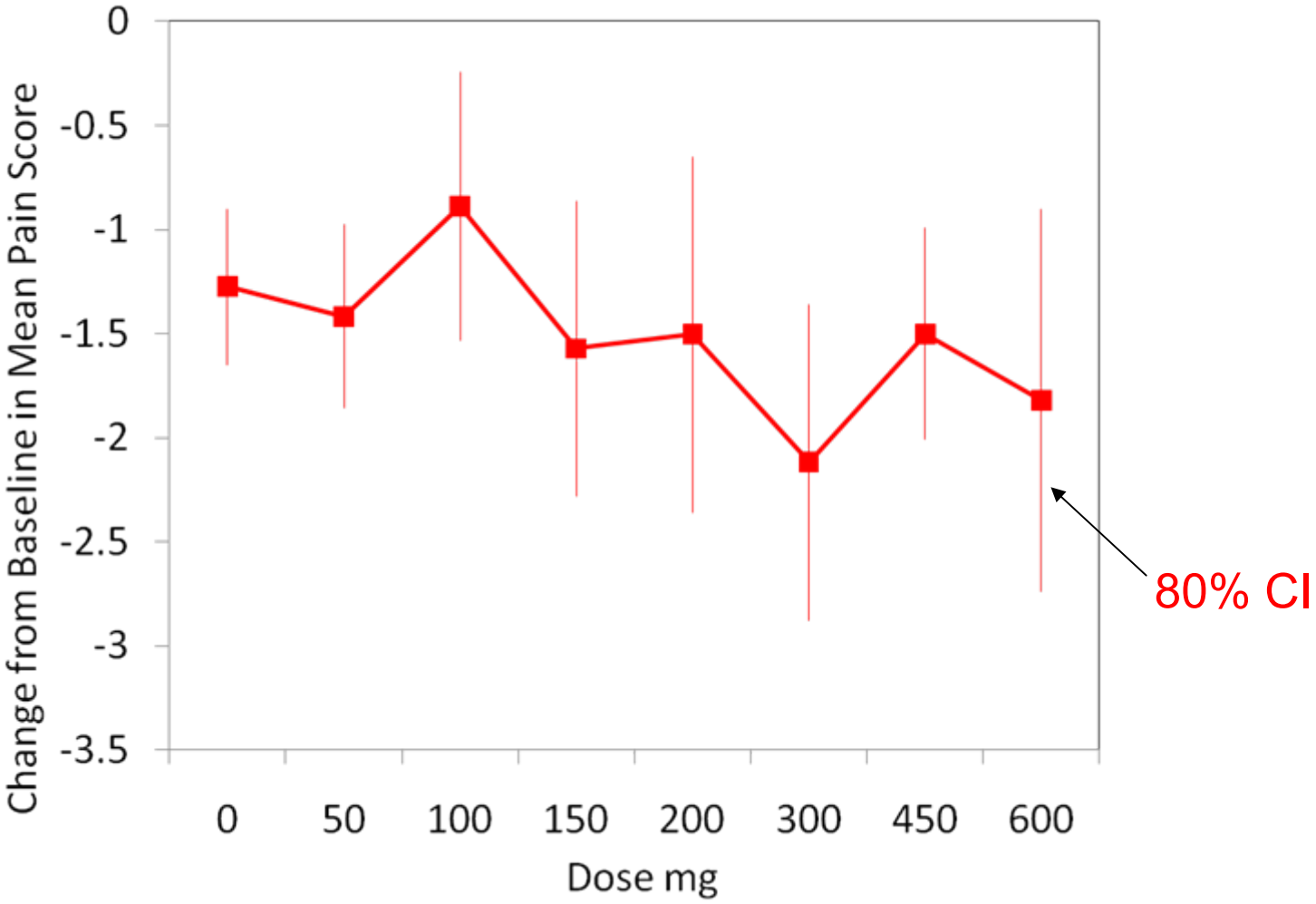


Problems in Fitting Emax Models

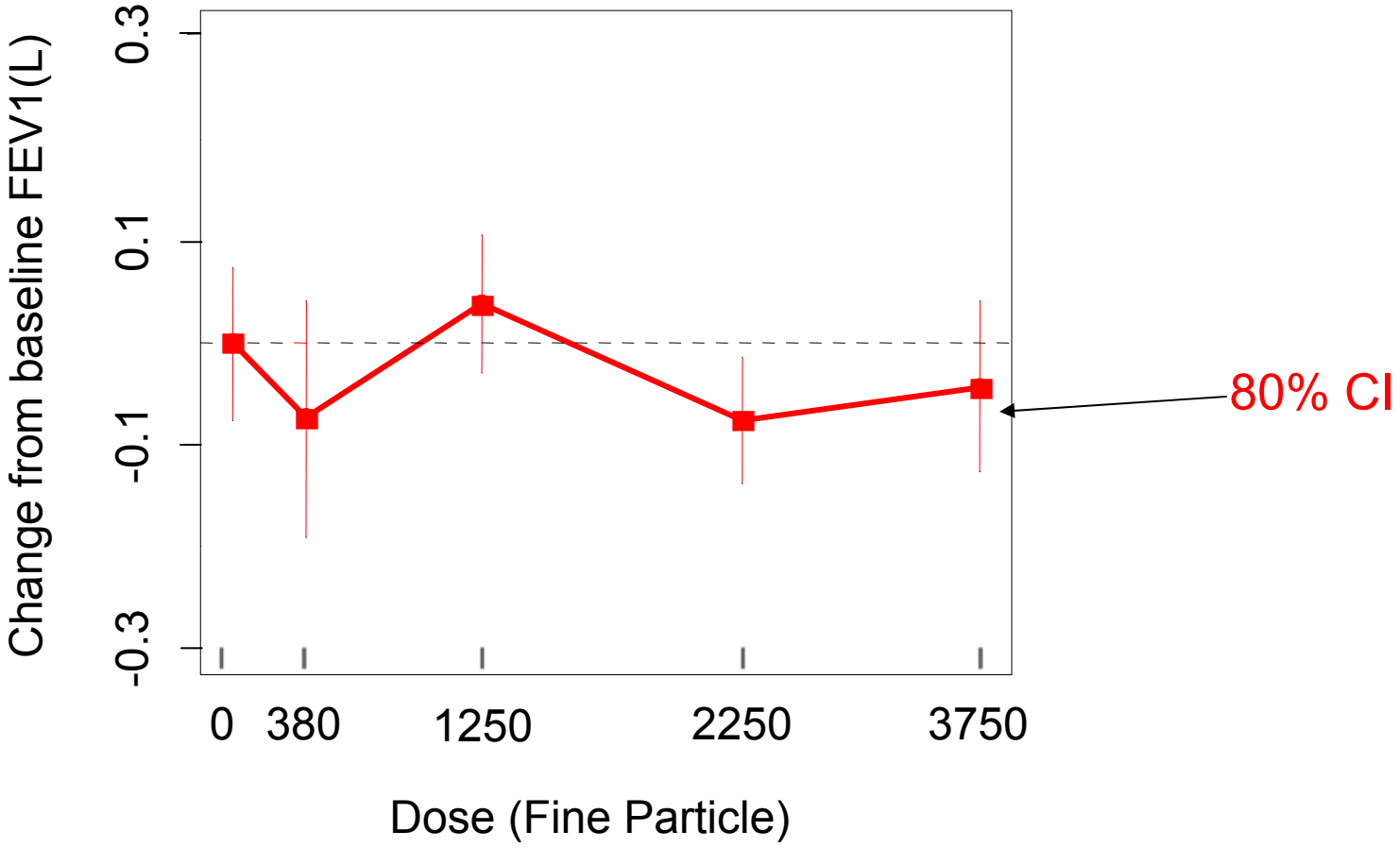
Results from 3 Dose Response Studies

- Study 1 - Pain in Post-Herpetic Neuralgia
  - 0, 50, 100, 150, 200, 300, 450 and 600 mg
  - Endpoint: Change from baseline in average weekly pain score
- Study 2 - ASTHMA (Inhaled compound ,Disease-modifying (suppress inflammation)
  - 0, 380, 1250, 2250, 3750 micrograms fine particle dose to lungs
  - Endpoint: Change from baseline at 6 weeks in FEV1
- Study 3 - Acute Neuroprotection following Stroke
  - 0, 10, 16, 22, 26, 33, 38, 45, 52, 59, 67, 76, 84, 96,108, 120 mg
  - Endpoint: Change from Baseline at 13 weeks in Scandinavian Stroke Score (Neurological)

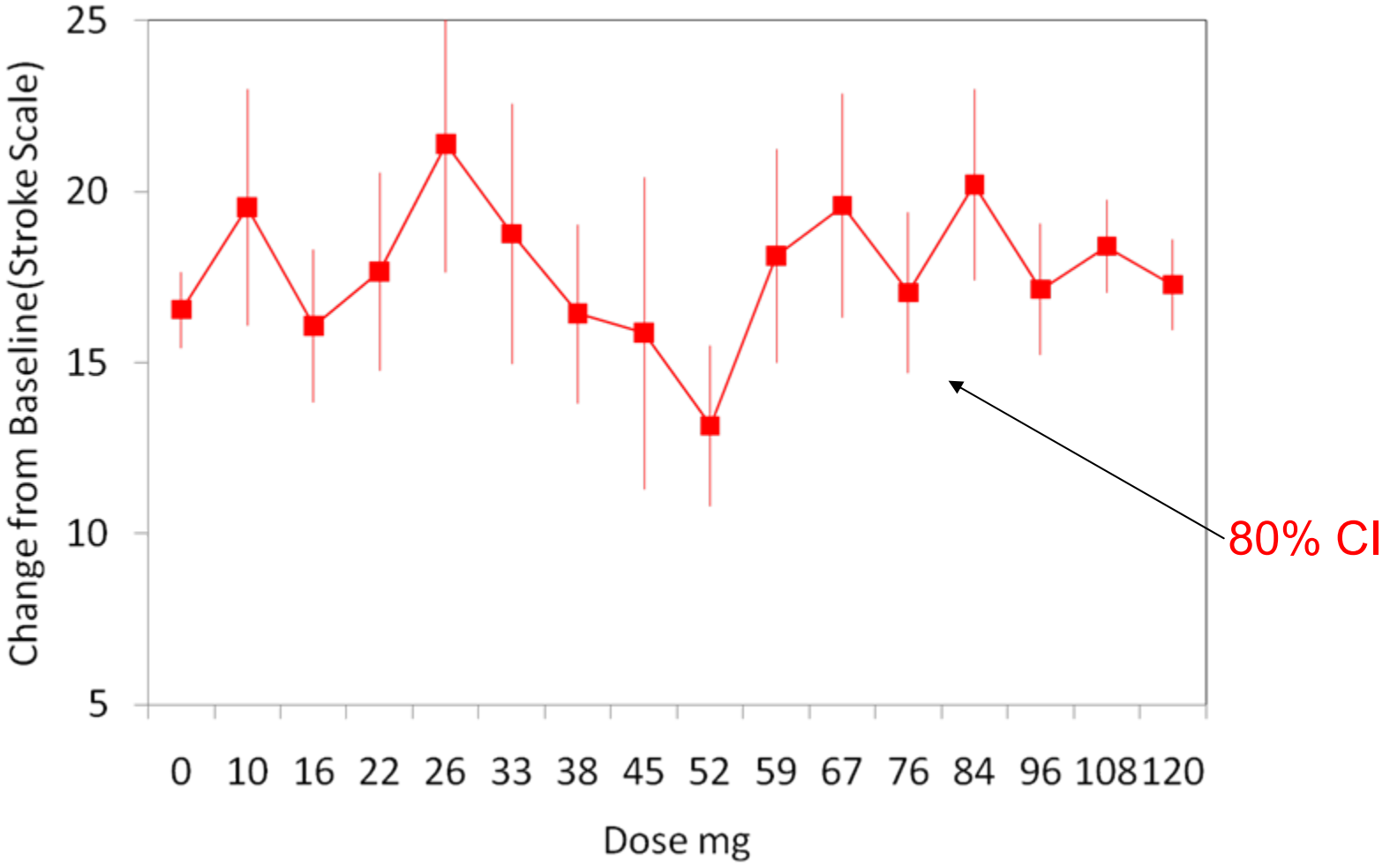
Pain Study



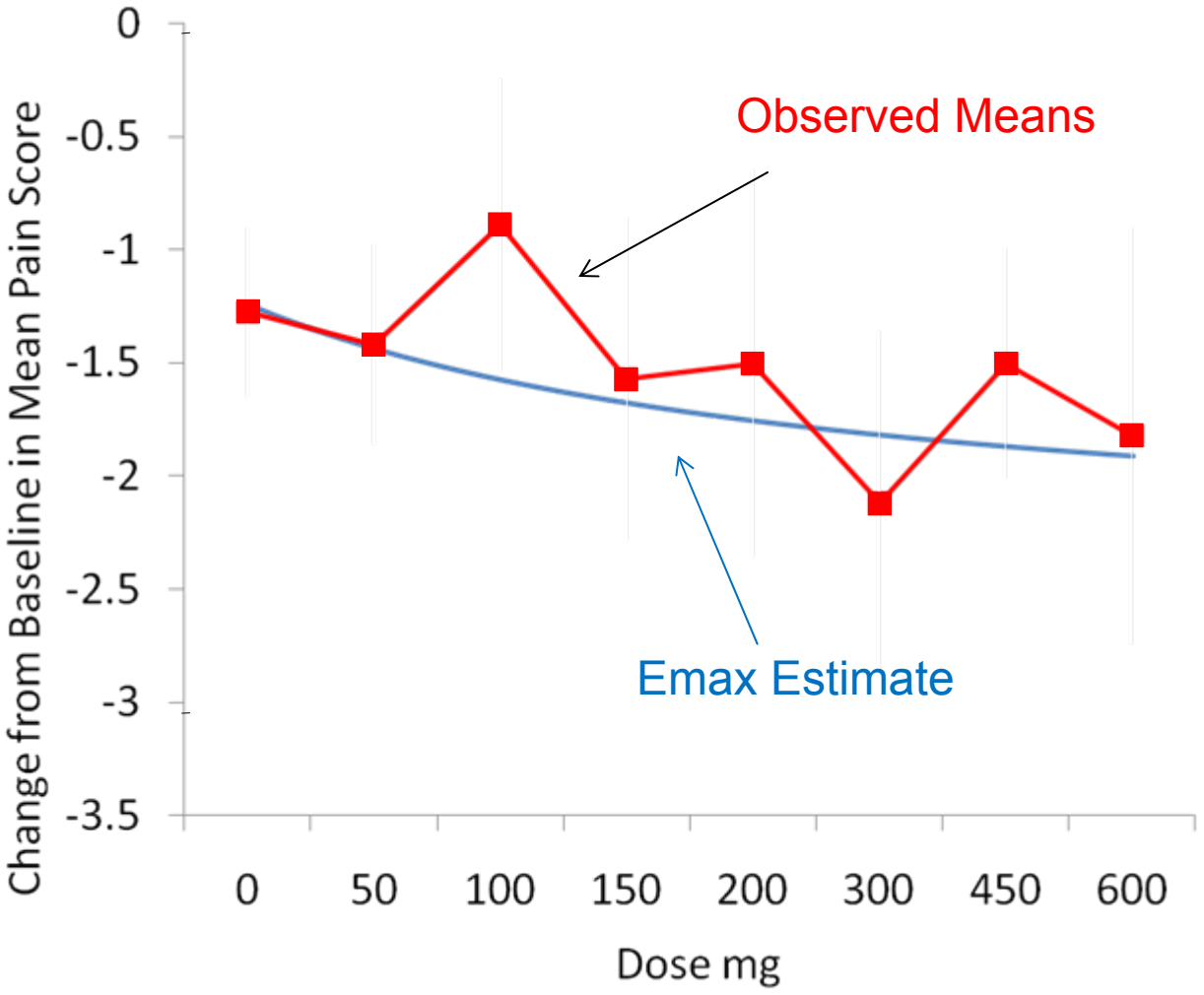
Asthma Study



Stroke Study



Pain Study
Fit of 3 Parameter Emax

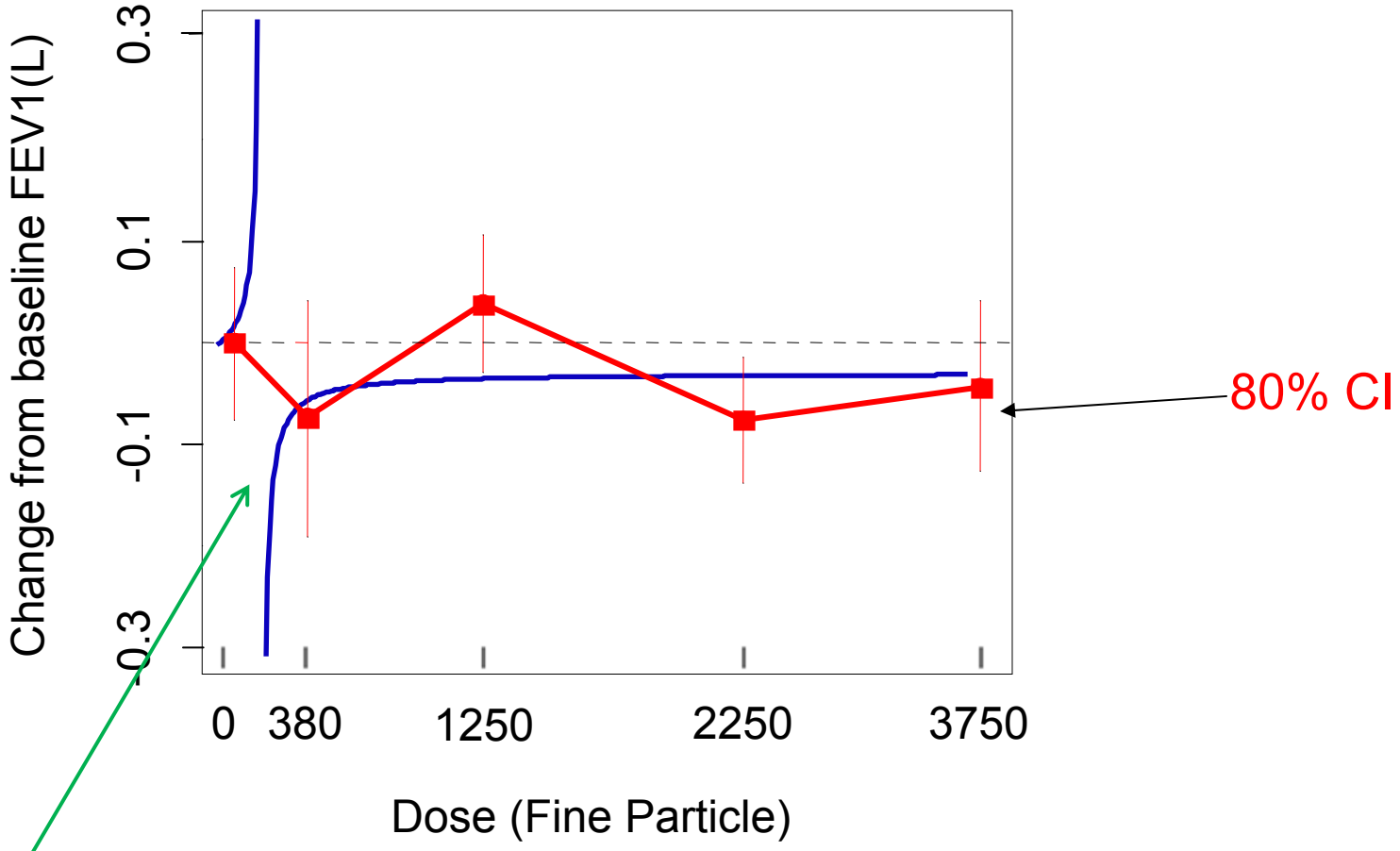


Asthma and Stroke Studies
Fitting Emax Models

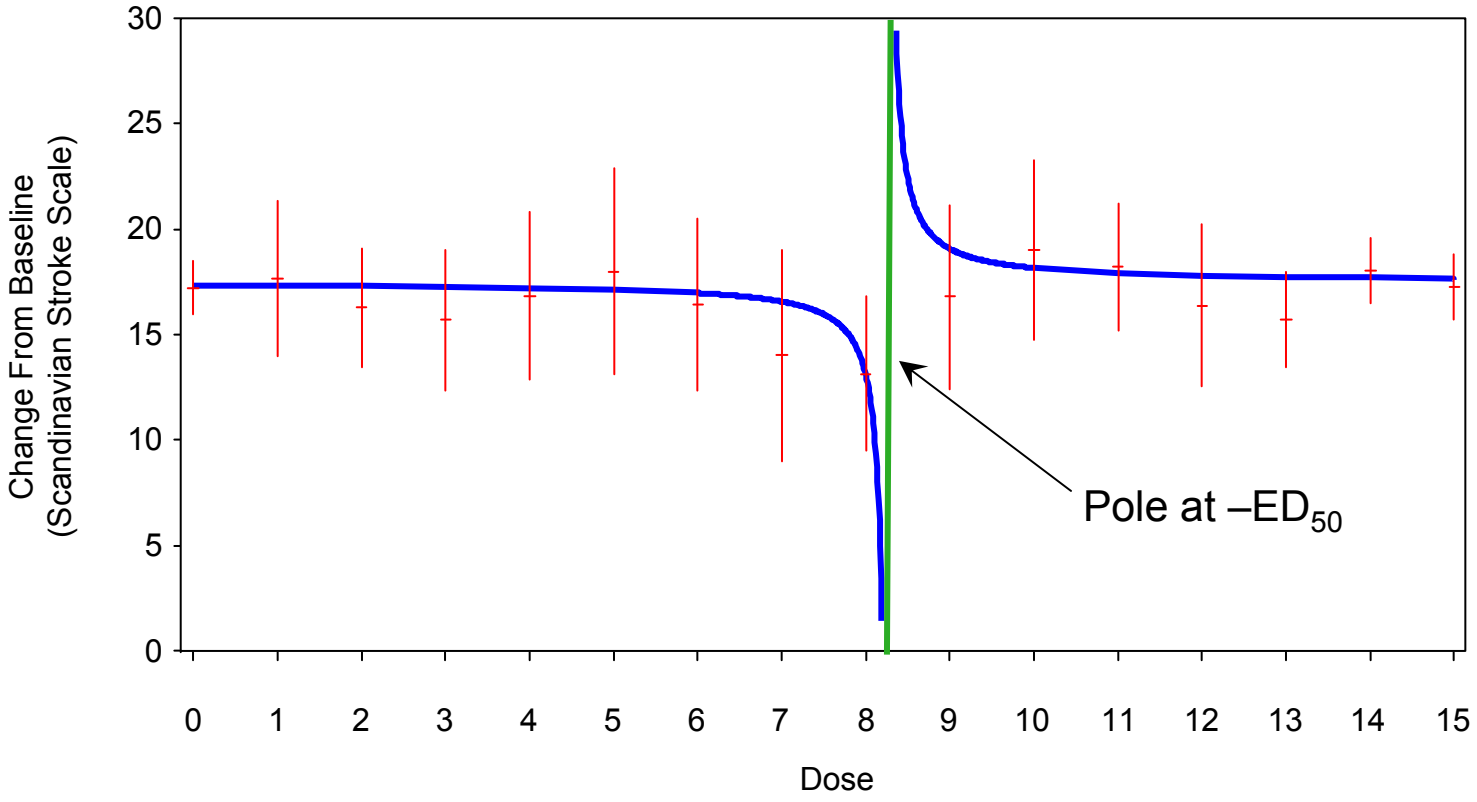
- Asthma Study
  - 4-parameter Emax model fails to converge with numerical errors in derivatives
  - 3-parameter model converges, but the ED50 is negative

E0	ed50	emax
-0.003	-222.779	-0.028

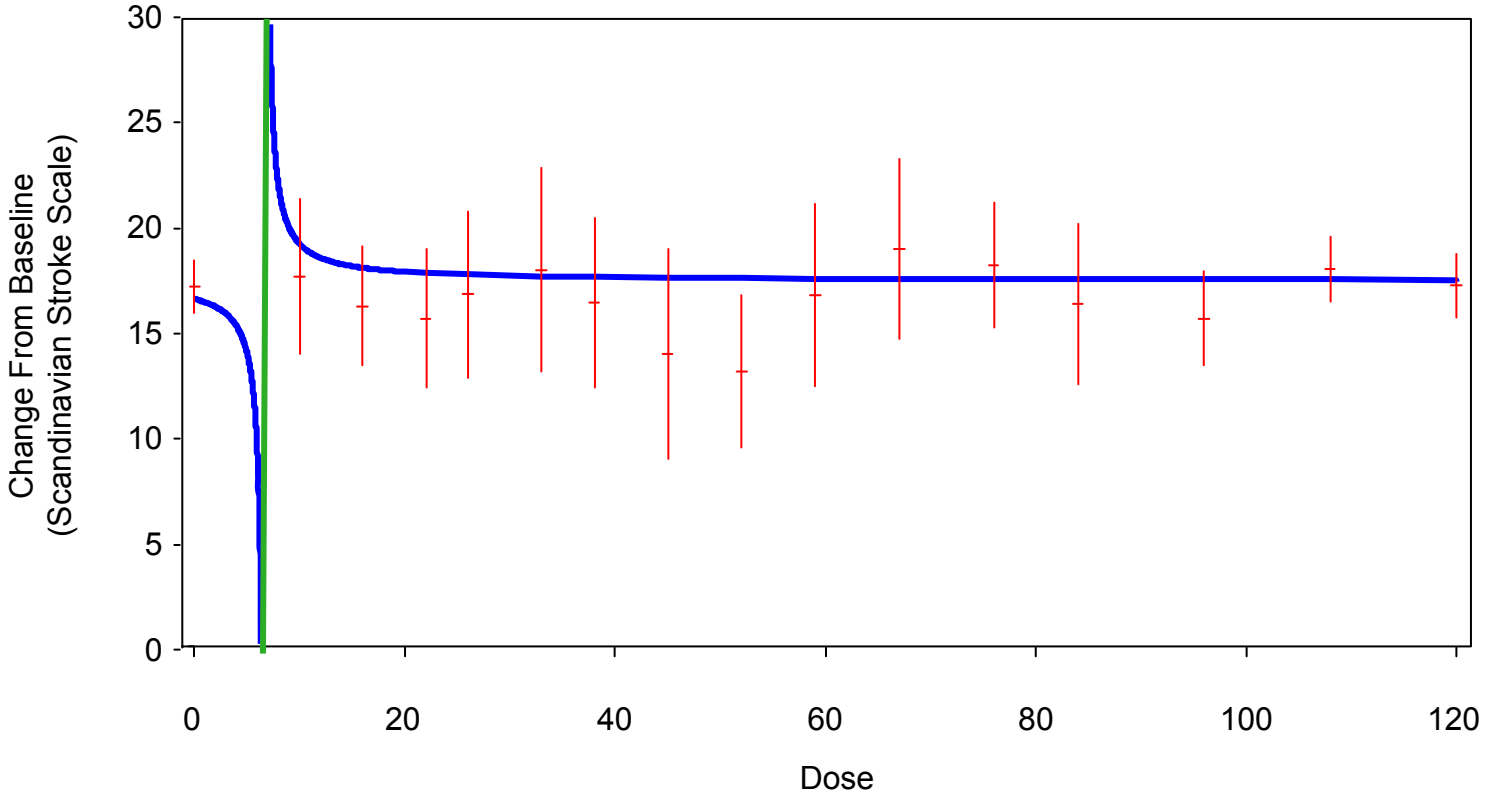
Asthma Study
Fitted 3 Parameter Emax Model



Stoke Study
Nominal Doses



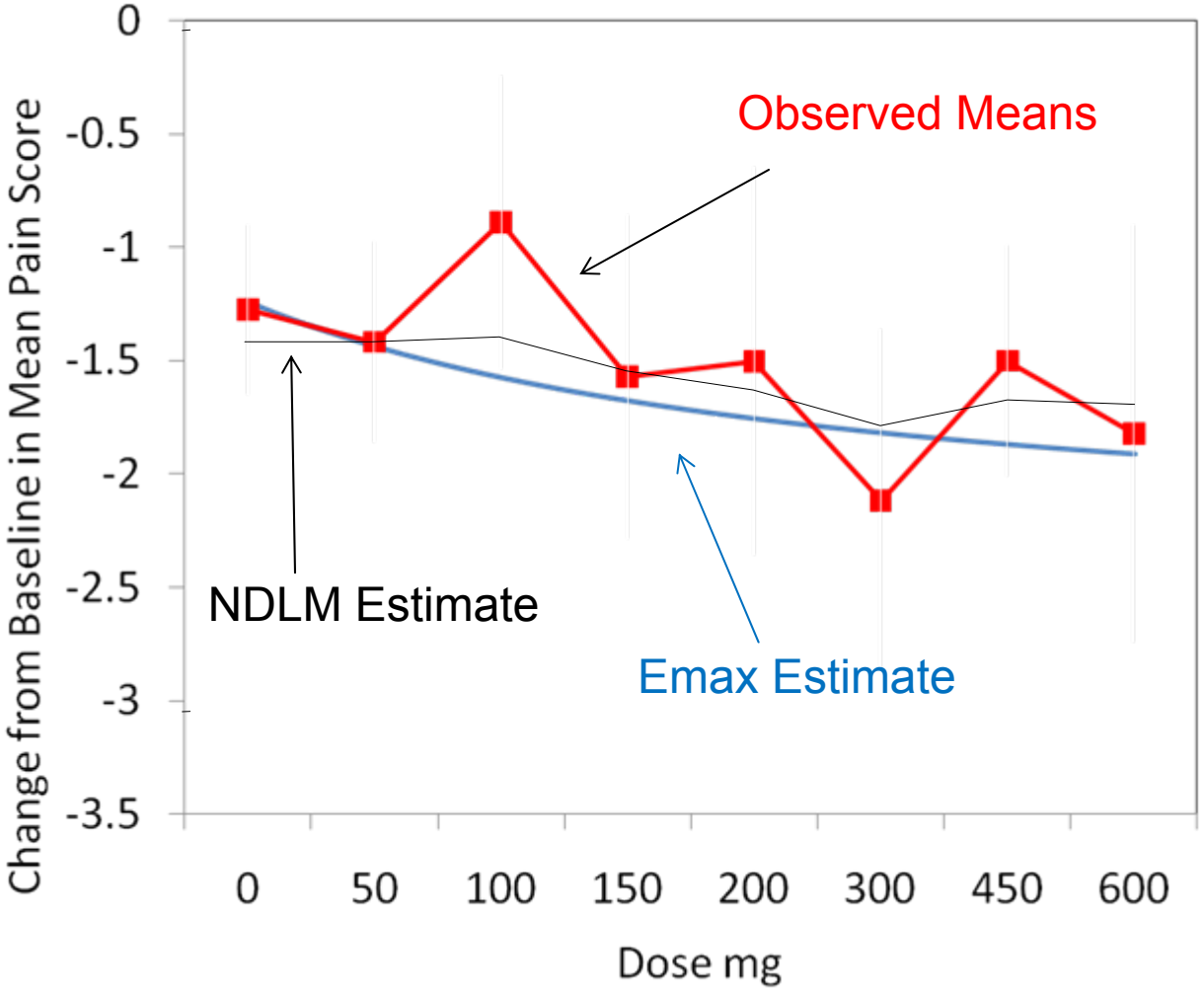
Stroke Study
Actual Doses



Unstable ML Estimation Is Common Under the Null Hypothesis

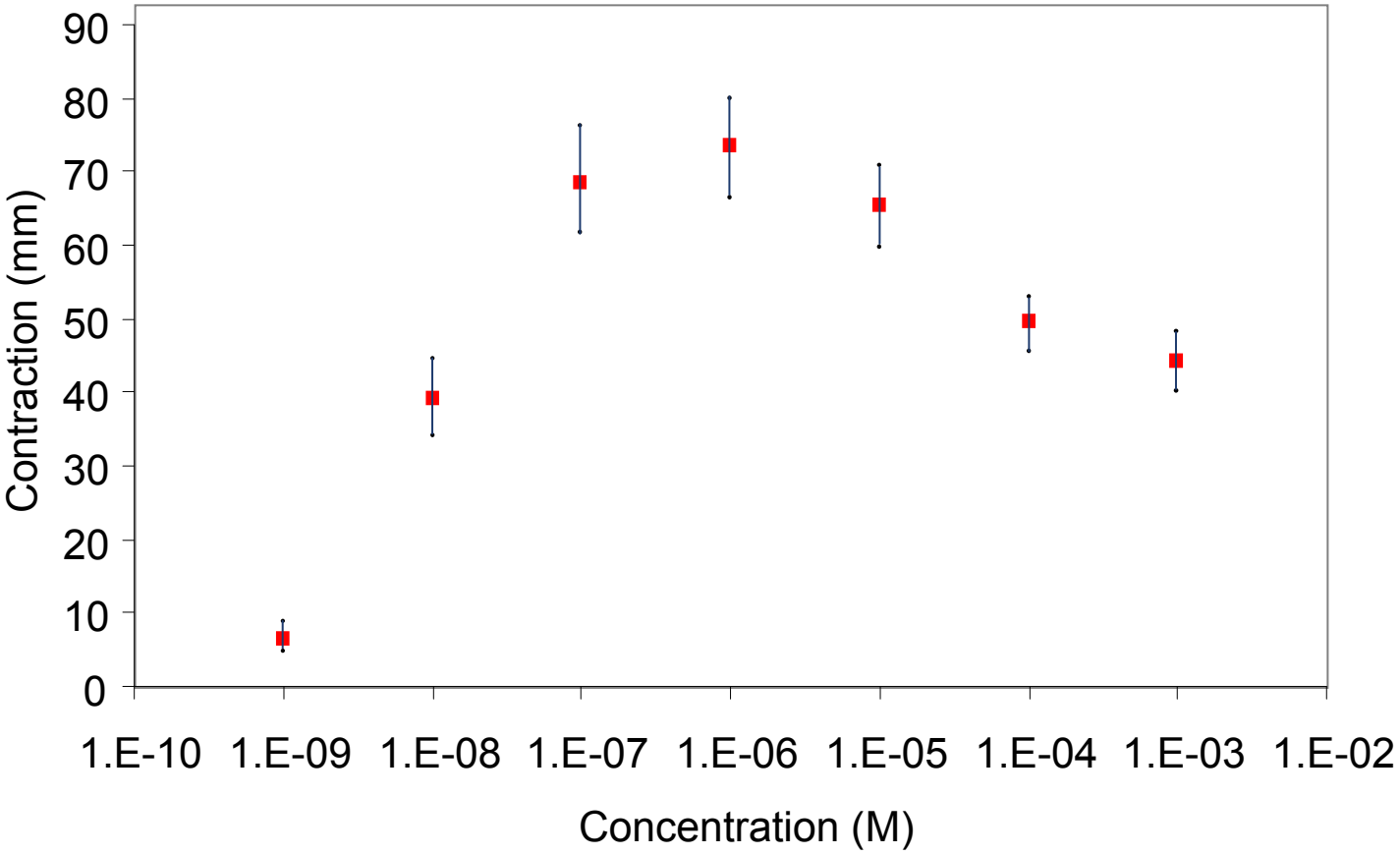
- Emax models are not identified
  - Emax parameter equal to zero for null effect
  - ED50 and Lambda are arbitrary
- Similar problems when the trend is weak
  - Signal to noise ratio is a key determinant of ML dose response estimation properties
- Simulated performance under Null Hypothesis
  - Non-convergence 0.41
  - Converged (Large ED50) 0.23
  - Negative ED50 0.14
  - 3 Parameter Emax 0.22
  - Alpha level (nominal 0.05): 0.067 (sim err 0.008) - slightly anti-conservative, depends on starting values and stopping criteria

Pain Study

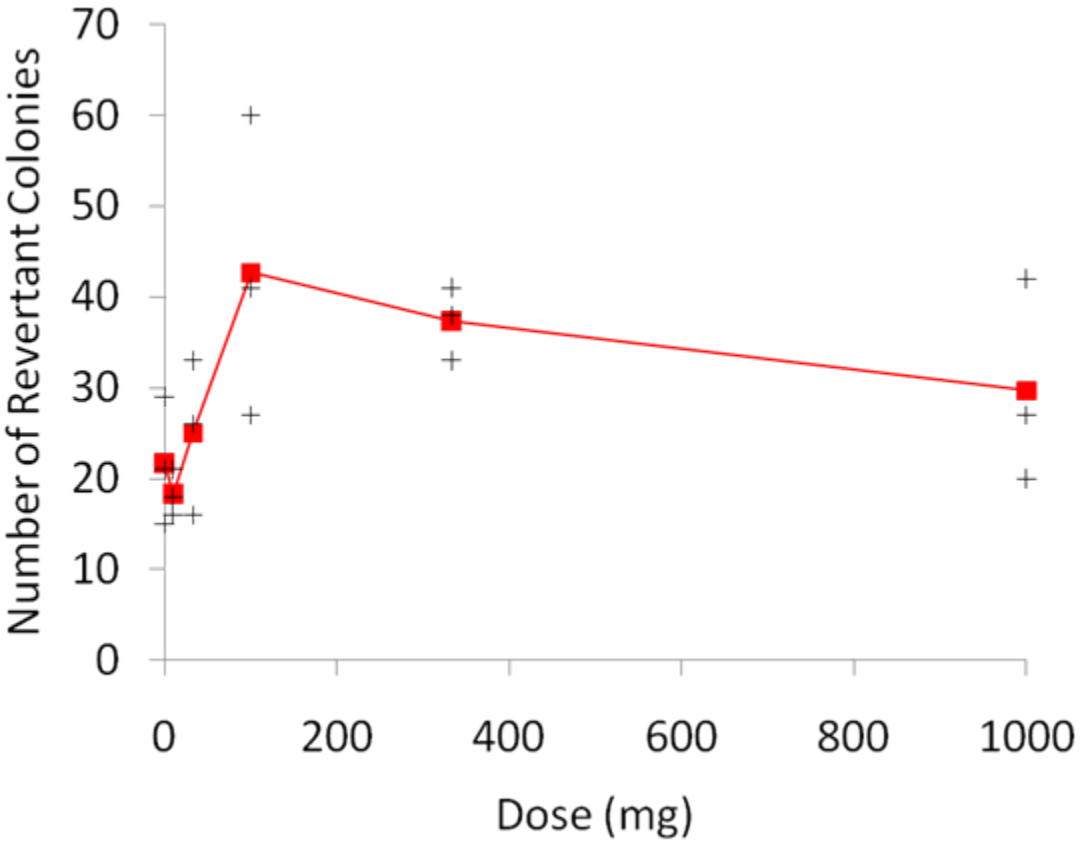


Non-Monotonic Models

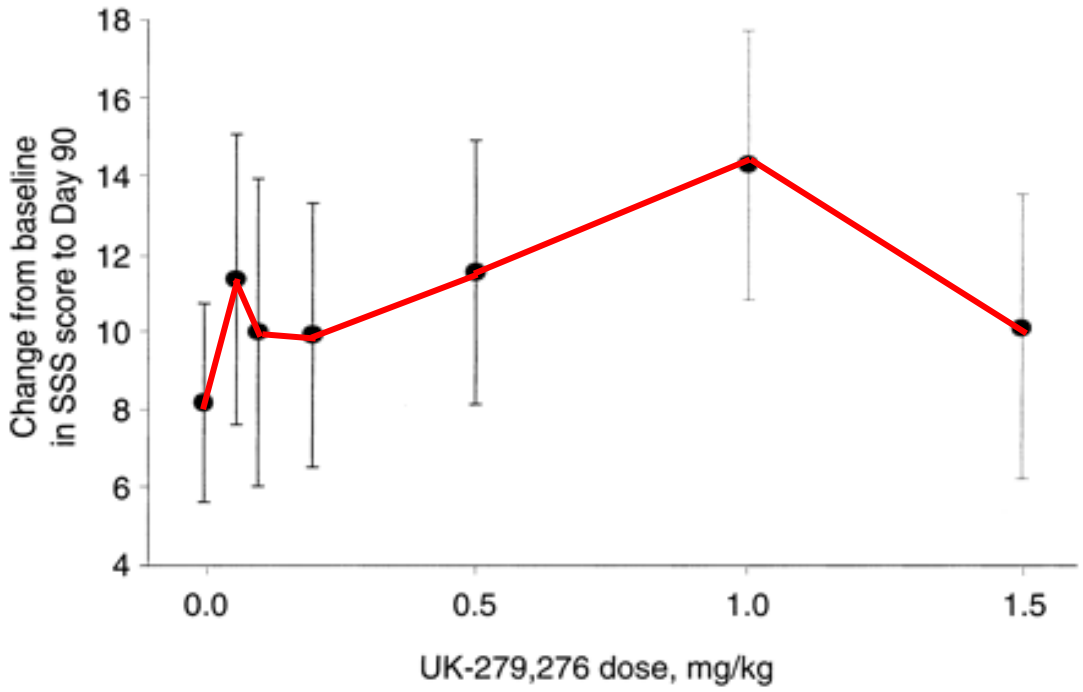
Pre-Clinical Study
Effect of Acetycholine on Isolated Guinea-Pig Ileum



Mutagenicity Testing
Ames Test Data



Stroke Clinical Trial Patient Safety Study



Extending the Emax Model

- Suppose drug combines with 2 receptors
- Both drug/receptor complexes produce the same kind of effect Y_1 and Y_2

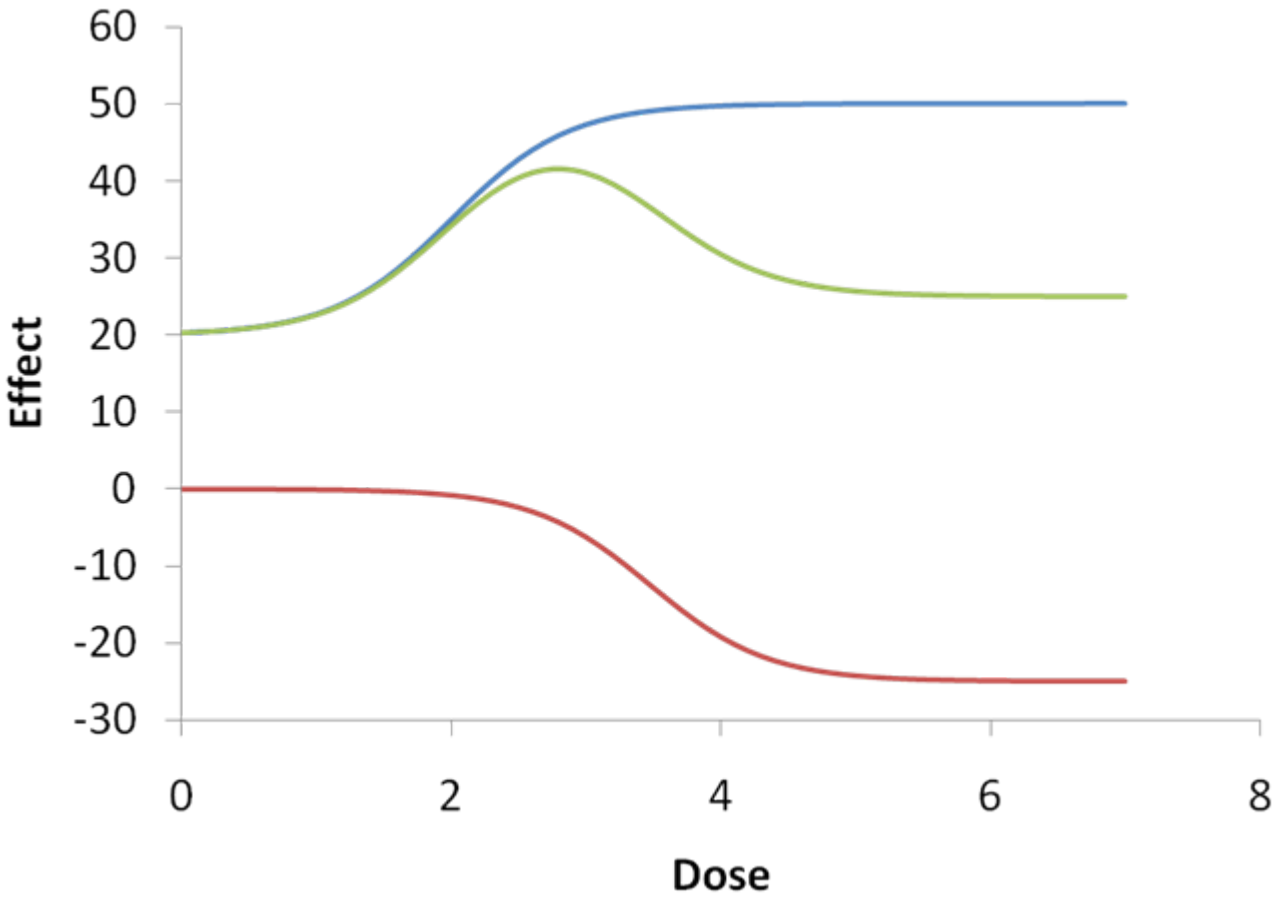
$$E(Y_1|D) = \frac{E^1_{max} D}{ED^1_{50} + D} \qquad E(Y_2|D) = \frac{E^2_{max} D}{ED^2_{50} + D}$$

- Overall Effect

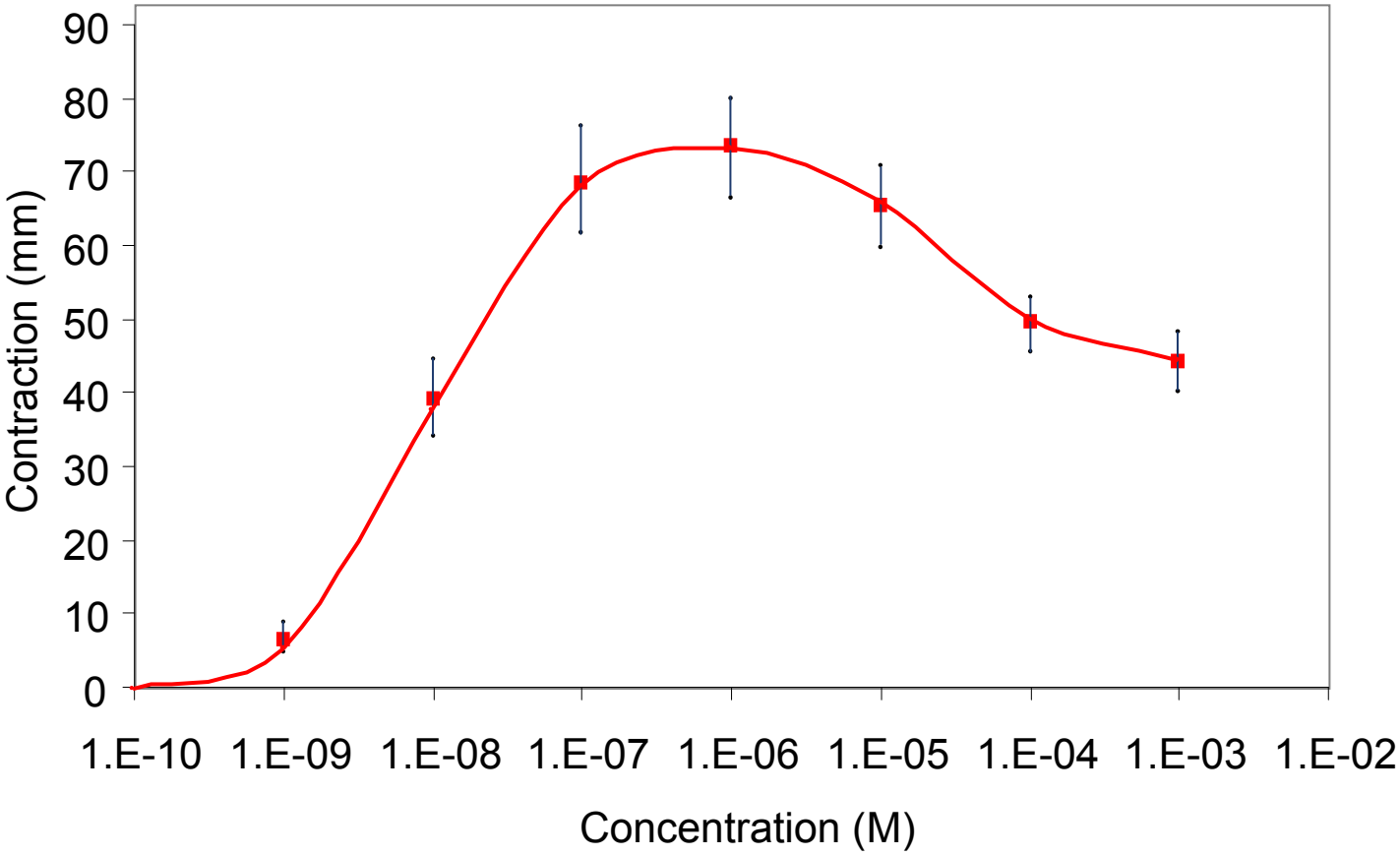
$$E(Y|D) = \frac{E^1_{max} D}{ED^1_{50} + D} + \frac{E^2_{max} D}{ED^2_{50} + D}$$

- If the two Emax values are of opposite sign the combination has a maximum

Superposition of Two Emax Curves

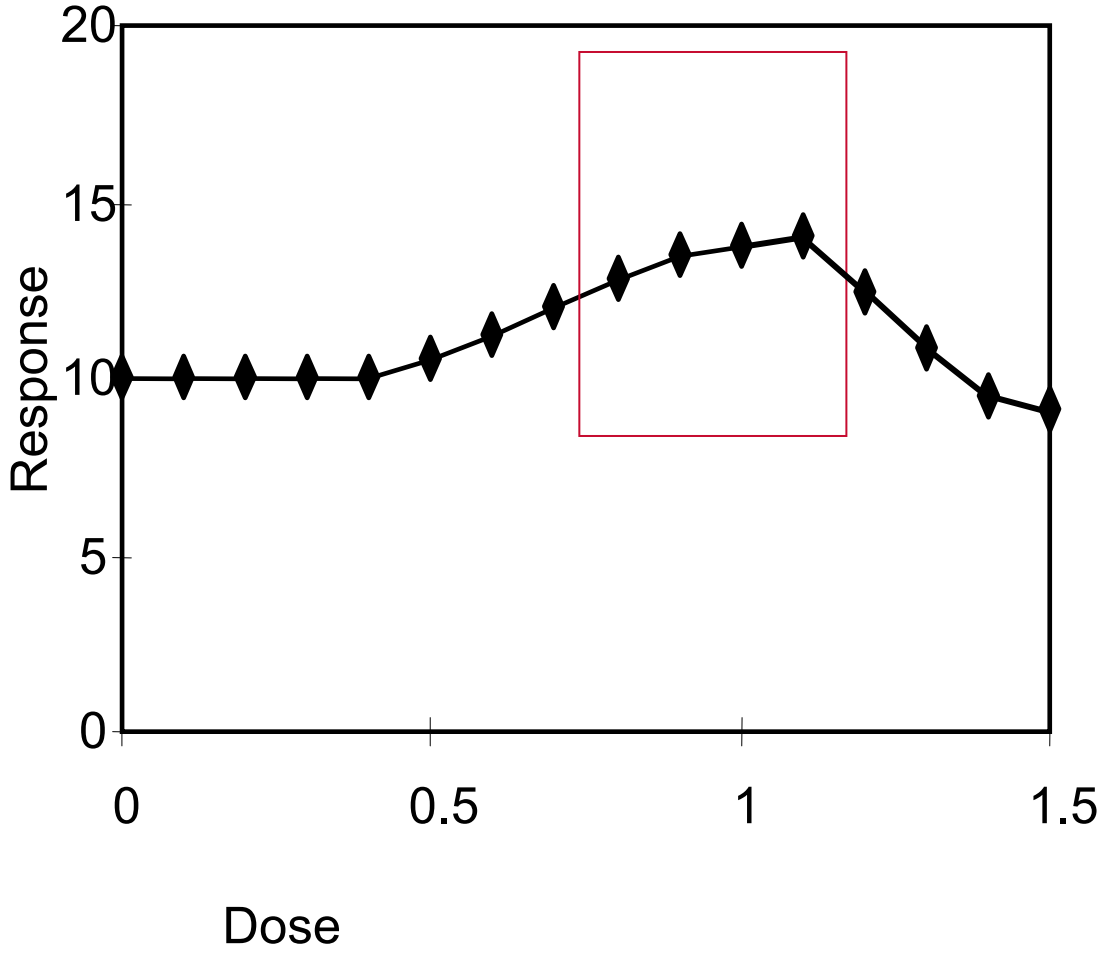


Guinea-Pig Contraction Data Fitted Using Superimposed Emax Models



Modelling Dose Response

- We model f (z , θ) as a 2nd order polynomial NDLM (West and Harrison 1997)



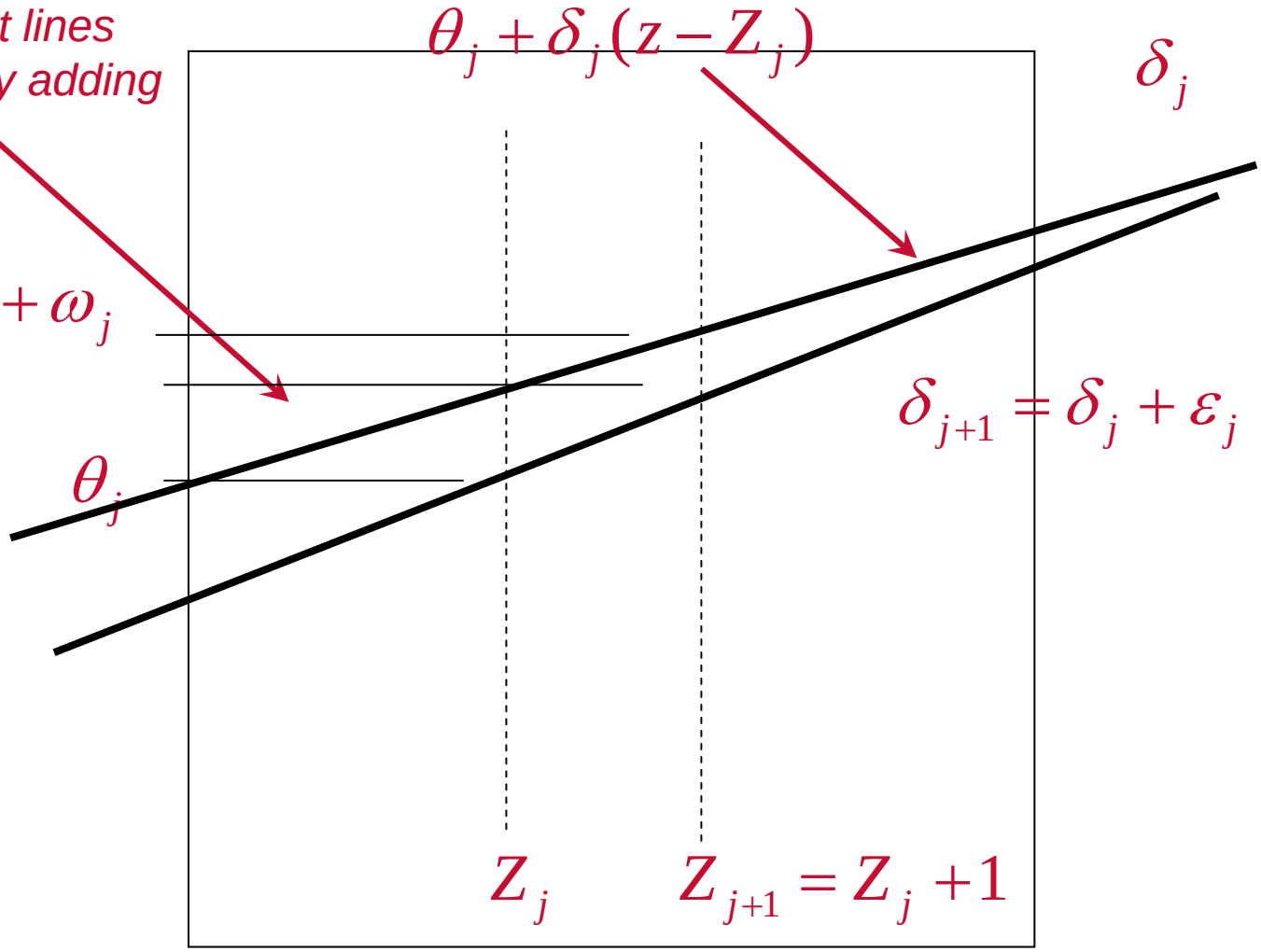
Alternative:
A Class of Semi-Parametric Smoothing Models

2nd Order Polynomial NDLM

Locally around z = Z_j a straight line with level θ_j and slope δ_j

Parameters of the straight lines change between doses by adding a (small) evolution noise:

$$\theta_{j+1} = \theta_j + \delta_j + \omega_j$$
$$\theta_{j+1} = \theta_j + \delta_j$$



Evolution Variance = Smoother

Normal Dynamic Linear Model

- NDLM – 2nd oder polynomial
- Observation Equation: $Y_{jk} = m_j + \nu_{jk}$, $\nu_{jk} \sim N(0, V\sigma^2)$
- System Equation: $\mu_j = \mu_{j-1} + \delta_{j-1} + \omega_j$, $\omega_j \sim N(0, W_j\sigma^2)$

  $\delta_j = \delta_{j-1} + \varepsilon_j$, $\varepsilon_j \sim N(0, W_j\sigma^2)$
- Issues:
  - Choice of W_j - can be fixed
    - can learn about it
  - Covariates can be included by making the expected responses depend linearly on the covariates
  - $E(y_{jk} \mid z = Z_{jk}, x_k) = \theta_j + b \times x_k$
  - The NDLM is then applied to these θ_j's

Other approaches

- Montonicity
  - Restrict NDLM to monotone function
  - Isotonic regression
- Splines
- Gaussian process regression (GPR)
  - NDLM, splines & GPR all impose a multivariate structure on the data
- Mixture of models
  - Bayesian model averaging

Translation of Methodology Into Practice
Why is it so slow?

The Slow Pace of Translation of Methodological Development into Practice

- Despite these advantages and recommendations, Bayesian adaptive designs have not been widely adopted in practice
- Only 20 (1.6%) of 1,235 phase I cancer trials have been reported to follow an innovative Bayes design by Rogatko et al. (J Clin Oncol 2007)
- This could be in agreement with the previous report from Altman and Goodman (JAMA 1994)
  - « *Newer technical innovations still typically take 4 to 6 years before they achieve 25 citations in the medical literature.* »

Special Issue Paper

Statistics in Medicine

Received 18 October 2010, Accepted 11 July 2011 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.4363

Bayesian adaptive clinical trials: a dream for statisticians only?

Sylvie Chevret[a,b,c*†]

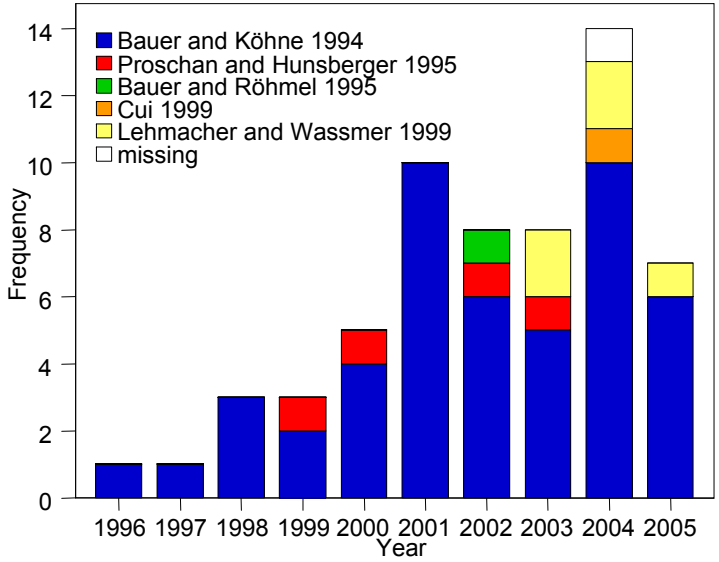
The Slow Pace of Translation of Methodological Development into Practice

- Bayesian clinical trials have been recommended for the last two decades
  - From the early phase trials up to the phase III
  - However, they have been reported poorly used in practice
  - Possibly due to the usual time lag of the technical innovation spread - This was confirmed in this study with only 3% of biostatistical papers reaching 25 citations after publication, as compared to 15% of reviews and 32% of clinical trial reports

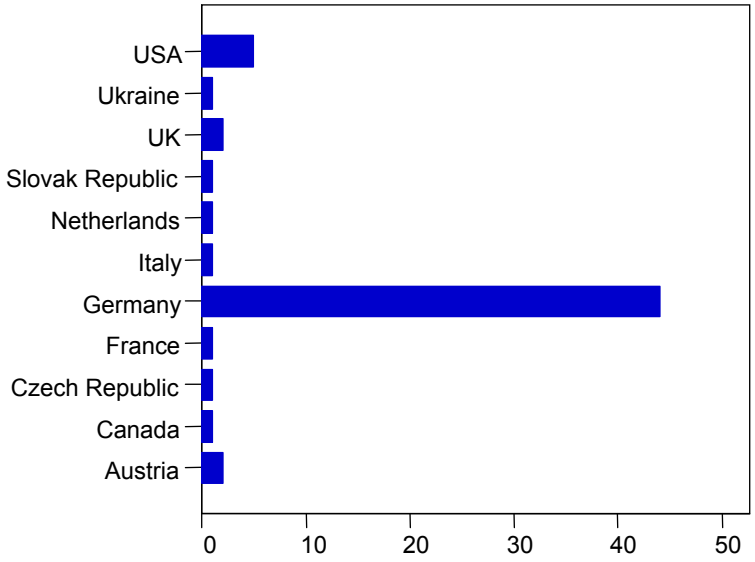
Sylvie Chevret (SIM, 2011)

Review of Adaptive Interims Bauer and Einfalt (Biometrical J, 2006)

- Identified 75 papers dealing with adaptive designs : 1989-2004
  - combination tests
  - conditional error function
  - did not consider Bayesian approaches
- Searched for “applied papers” in SCI, SSCI, IAHCI referring to at least one of the 75 papers
- Identified 60 applied medical papers



By: year and adaptive methodology



By: Country of corresponding author

Adaptive Confirmatory Interim Designs

- Adaptive interims not widely used
- Methods used mainly in Germany
- Adaptations in practice are limited to sample size re-assessment
- Sophistications – dropping treatment arms, modifying endpoints etc have not entered medical literature
- Standard of presenting statistical methods poor
  - pressures on space ?
- Mid-trial changes may impact negatively on the “persuasiveness” of the results

Altman and Goodman (JAMA 1994)

Newer Statistical Methods That may be Seen More Often In the Coming Years

Method	Description	Purpose
Bootstrap (also called resampling; related to the jackknife)[17]	Multiple new data sets are generated by random sampling "with replacement" from the original data	To calculate SEs or assess the stability of a statistical model, often when standard assumptions are unreliable or the sampling distribution is unknown
Gibbs sampling[18,34]	Random sampling from conditional distributions within a complex structure	Bayesian estimation of complex models
Generalized additive models[35]	Nonparametric smoothing of explanatory variables in regression	To replace regression when assumptions are not tenable
Classification and regression trees[19,36] (also known as recursive partitioning)	Division of a set of subjects by combinations of characteristics, to minimize the differences within groups and to maximize the differences between groups	To find combinations of variables of predictive importance
Models for longitudinal data ("general estimating equations")[20]	Modeling repeated measurements of an outcome variable while allowing for covariates	Regression for multiple assessments of outcome
Models for hierarchical data (also called multilevel models)[37]	Fitting mixed linear models to hierarchical data using iterative generalized least squares	Modeling data with more than one level of variation (eg, within and between patients)
Neural networks[38]	Nonparametric modeling of complex data	To provide nonlinear approximations to multivariable functions or for classification

Conclusions

- Choosing 2,3 or 4 doses in a phase II dose-response design is potentially wasteful and counter productive
- Consideration should be given to increasing the number of doses, the range of doses and the analytic methods.
- Efforts should be taken to encourage the speedy translation of innovative methodology into practice