KING'S College LONDON
University of London

Adaptive Designs : An Overview

Professor Andy Grieve

Department of Public Health Sciences

King's College, London

Andy.Grieve@kcl.ac.uk

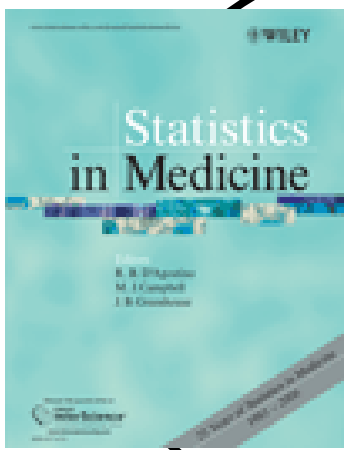
KING'S College LONDON
University of London

Recent Special Issues Adaptive Designs Current Interest - High

Journal of Biopharmaceutical Statistics, 16: 275–283, 2006

ISSN: 1054-3406 print/1520-5711 online
DOI: 10.1080/10543400600614742

Drug Information Journal
OFFICIAL PUBLICATION OF THE DRUG INFORMATION ASSOCIATION

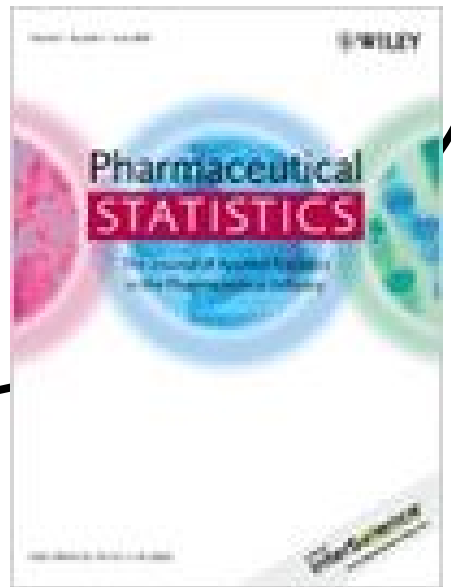


2006

of the International Biometric Society

Biometrical Journal

Methods · Applications · Case Studies · Regulatory Aspects · Reviews



KING'S College LONDON
University of London

Outline

- An Old Idea Resurrected (1930s)
- Examples from the literature
  - Up-and-Down Designs (1940s)
  - Play-The-Winner Designs (1960s)
  - Randomised Play-The-Winner Designs (1970s)
  - Clinical CRM Designs (1990s)
  - Adaptive Interim Analyses (1990s)
- An Old Idea Full-circle – Bayesian Adaptive Randomisation
- Taxonomy of adaptive designs
- Why does the debate continue as to their worth?

Adaptive Ideas Are Not New

ON THE LIKELIHOOD THAT ONE UNKNOWN PROBABILITY EXCEEDS ANOTHER IN VIEW OF THE EVIDENCE OF TWO SAMPLES.

By WILLIAM R. THOMPSON. From the Department of Pathology, Yale University.

Biometrika, 1933

Thus, if, in this sense, P is the probability estimate that one *treatment* of a certain class of individuals is *better* than a second, as judged by data at present available, then we might take some monotone increasing function of P, say $f_{(P)}$, to fix the fraction of such individuals to be treated in the *first manner*, until more evidence may be utilised, where $0 \leqslant f_{(P)} \leqslant 1$; the remaining fraction of such individuals $(1 - f_{(P)})$ to be treated in the *second manner*; or we may establish a probability of treatment by the two methods of $f_{(P)}$ and $1 - f_{(P)}$, respectively. If

- Ethical Design – concentrating on delivering the best treatment to the most patients
- Forerunner of the Randomised Play-the-winner Design
- Catalyst for the development of Bandit designs

KING'S College LONDON
University of London

Up-and-Down Designs

KING'S College LONDON
University of London

Dose Response
An Old – Non-linear - Design problem

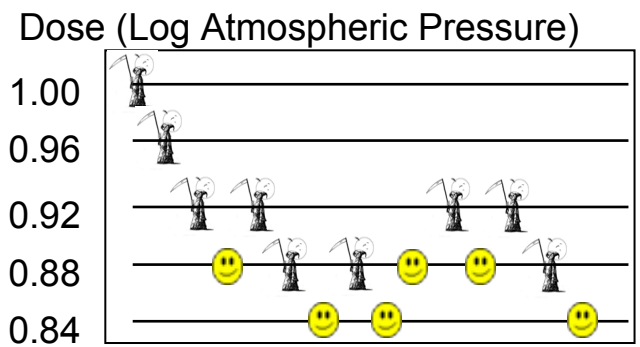
- Non-linear response function
  - Optimal design available if we know the function
  - We don't know the function
- Solution :
  - Do some experiments
  - Learn a bit
  - Optimise
  - Learn a bit more
  - Optimise
  - etc

- Up-and-Down Design
  - Allocates patients to dosing groups (usually unequally)
  - Dose finding process
  - Nth patient gets allocated to dose depending on response of (N-1)th patient
    - Success : lower dose
    - Failure : increase dose

Up-And-Down Experiment
Dixon & Mood(J Am Statis Ass, 1948)

- Developed to estimate the ED_{50}
- ASIDE -similar ideas were developed in sensitivity testing, psychophysics predating Dixon and Mood – Georg v Bekesy (Acta Oto-laryngol.1947)
- Estimate : Determine the number of doses giving successes and the number giving failures. Take the smaller total. Take the average dose for the smaller total - A. Then the estimated ED_{50} =A+Δ/2 for successes and ED_{50} =A-Δ/2 for failures. (s.e. can also be determined)

Impact of Mechanical Head Trauma - Choi (Biometrics, 1990)



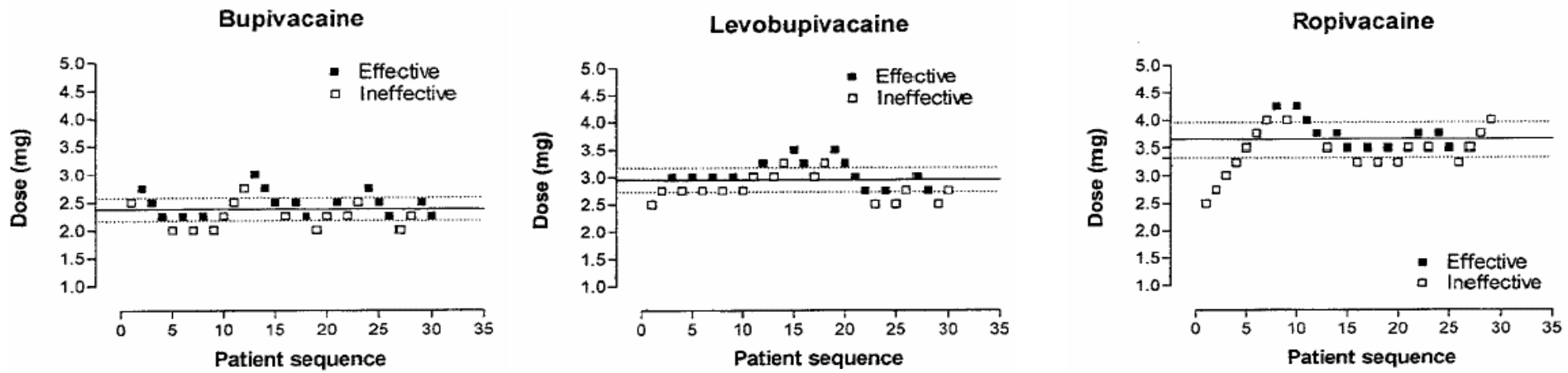
9 failures , 6 successes

A=0.86

Ed50= 0.86+0.04/2=0.88

Examples of Up-and-Down Designs Many in Anaesthetics

- Camorcia et al(Anesthesiology, 2004)
  - Ropivacaine, Levobupivacaine & bupivacaine in intrathecal labor



Analgesic Potency Ratio	Dixon and Massey Method	Probit Regression	*P* Value
Bupivacaine: levobupivacaine	0.81 (0.69–0.94)	0.79 (0.70–0.88)	< 0.01
Bupivacaine: ropivacaine	0.65 (0.56–0.76)	0.62 (0.55–0.69)	< 0.001
Levobupivacaine: ropivacaine	0.80 (0.70–0.92)	0.79 (0.70–0.88)	< 0.01

One-way analysis of variance with Bonferroni correction.

?

KING'S College LONDON
University of London

Play-the-Winner Designs

Play-the the-Winner Rule Zelen (J Am Statis Ass, 1969)

- Treatment assignment depends on the outcome of previous patients - Response adaptive assignment
- When response is determined quickly
- 1st subject: toss a coin, H = Trt A, T = Trt B
- On subsequent patients
  - assign previous treatment if it was successful
  - Otherwise, switch treatment assignment
- Advantage: Potentially more patients receive the bettertreatment
- Disadvantage: Investigator knows the next assignment

- TRT A : S S F S S S F
- TRT B : S F
- Patient 1 2 3 4 5 6 7 8 9 .....
  - Analysis based on sequences
- Literature Examples - Postoperative Venous Thromboembolism
  - Larsen et al (Pharm, Med, 1994)
  - Reirtsen et al (Scand J Gastroenterol, 1993)
  - Mowinckel et al (Eur J Surg, 1995)
  - Bjerkeset et al (World J. Surg.,1997)
  - Reirtsen et al (Scand. J. Lab. Invest,1998)

KING'S College LONDON
University of London

Randomised Play-the-Winner Designs

Randomise Play the Winner (RPW) Design Wei LJ, Durham SD (J. Am Statis Ass, 1978

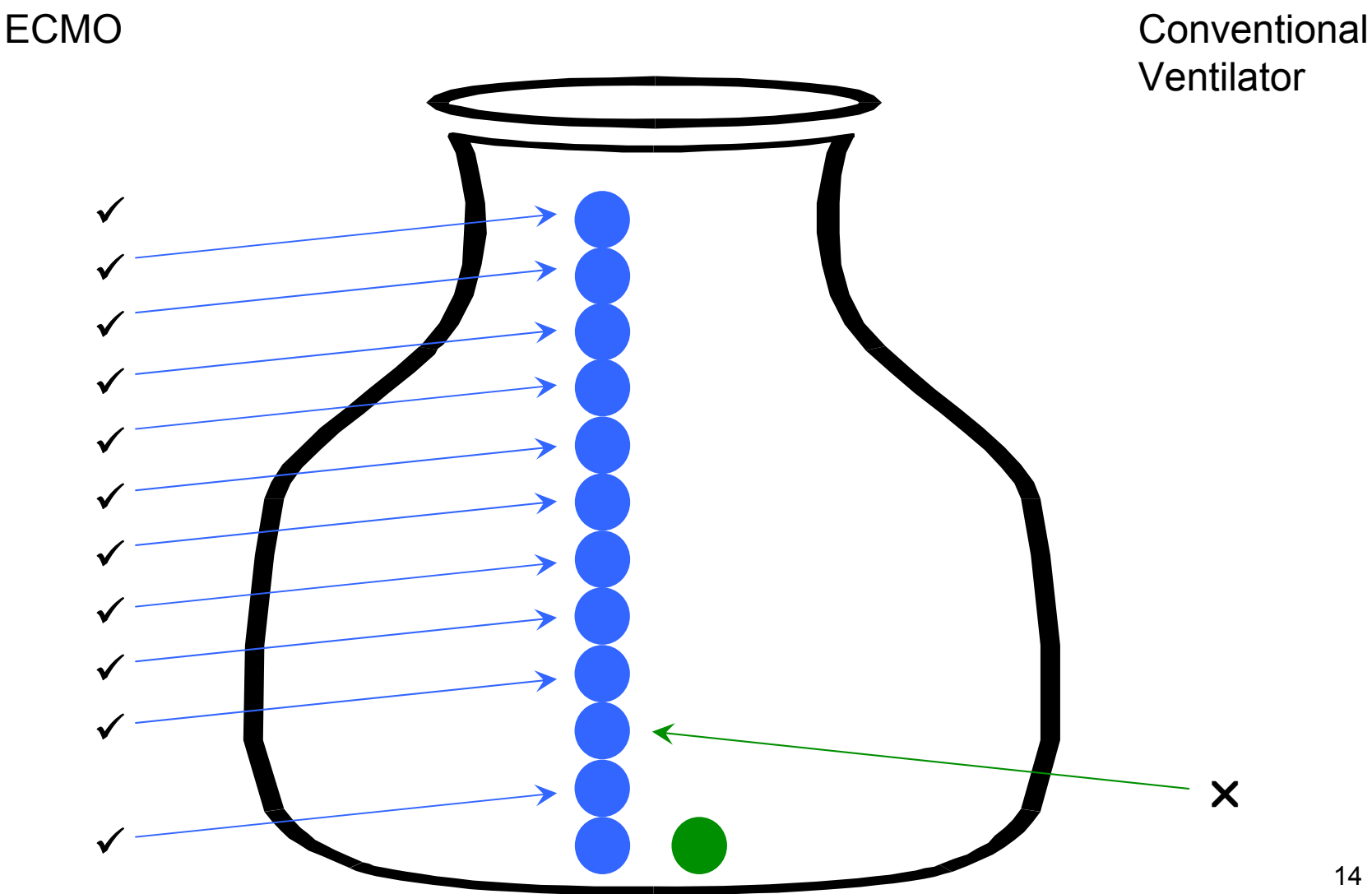
- RPW designs described by Urn models
- At beginning of trial
  - Urn contains α balls of each of two colours (W&R) representing 2 treatments
  - When a patient is to be treated a ball is chosen at random (with replacement)
  - When the response is known the urn content is updated as follows
- If the patient was allocated to treatment t (either W or R) and responds positively, β balls of colour t are added to the urn otherwise γ of colour s (the complement of t) are added.
- In time the urn will contain a higher proportion of colored balls associated with the more successful treatment
- RPW(α,β,γ) design

KING'S College LONDON
University of London

ECMO Bartlett et al (1985)

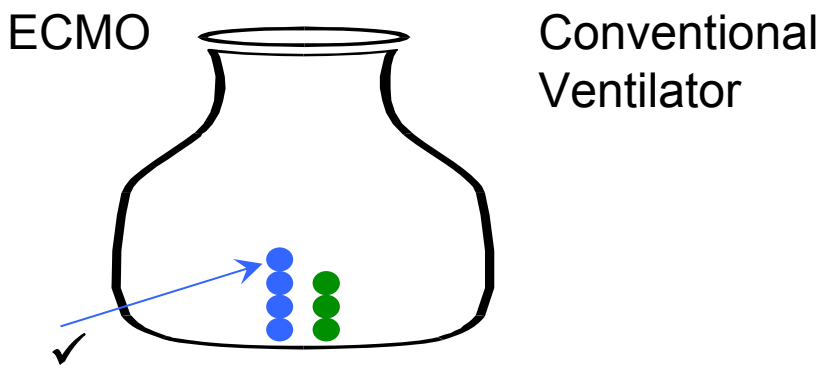
- Newborn infants with severe respiratory failure - Mortality
- Extra Corporeal Membrane Oxygenation vs Traditional Ventilator
- Phase I trials >50% survival on ECMO
- Optimal Therapy : survival < 20 %
- Chose Randomised Play-the-Winner (RPW)
  - speedy outcome - anticipated response diff -> small sample size - scientific/ethical dilemma

Randomised Play-the-Winner - Urn Model (ECMO)



Randomised Play-the-Winner Urn Model (ECMO) : Issues

- Was the urn model sensible ?
  - Other parameters
  - Begin with randomised block
- How reliable are the results - 11/11 vs 0/1 ?
  - Ranking and selection procedure
  - Minimum number of patients



- Ethics
- Tamura et al (J Am Statist Ass,1994)

KING'S College LONDON
University of London

Continuous Reassessment Method

The Continuous Reassessment Method (CRM) O'Quigley et al (Biometrics, 1990)

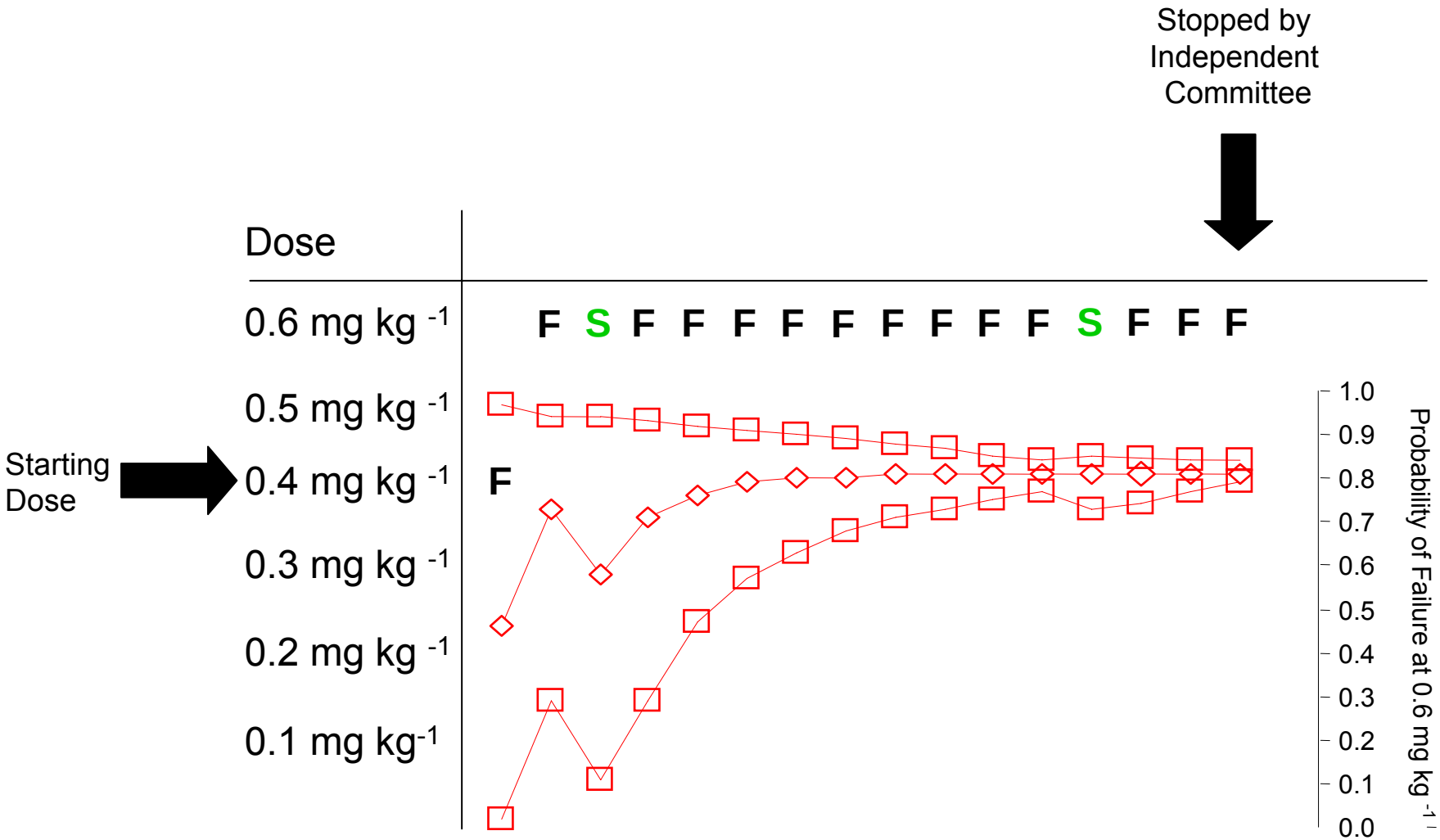
- Goal : identify a dose with the targeted toxicity as quickly as possible and focus experiment at that dose
- Doses are pre-defined : $d_1, d_2, \ldots, d_k$
- Outcome is binary : DLT / No DLT
- Assumption : There exists a monotone dose-response function $\psi(d;\theta) = \text{Prob}(DLT|d,\theta)$ depending on a single parameter θ
- The number of patients N is fixed in advance

- Sedation of Infants / Cardiac Catheterisation
  - Fabre et al (Br J Clin Pharm ,1998)
  - Aim : Find ED90 (90% sedated)
  - Bayesian approach
  - One parameter (α) logistic dose response
  - Choose dose to "optimise" gain (utility) function
    - predictive probabilities

$$\pi_i = \int_a p(Y = 1 \mid d_i, \alpha) p(\alpha \mid \underline{x}, \underline{y}) d\alpha$$

  - Choose as next dose the one which gives π_i closest to the target π (ED90)

CRM Design Infant Sedation



Fabre et al (Br J Clin Pharm ,1998)

Examples of Clinical CRMS

- Farge et al (Br. J Haematology, 2002)
  - Autologous BMT in refractory systemic sclerosis
- Desfrere et al (J. Clin Pharm. & Ther. 2005)
  - Ibuprofen in patent ductus arteriosus - 2 cohorts

Cohort no.	Patients (*n*)	Allocated dose	Success/ failures	Ibuprofen loading dose (mg/kg) Prior estimated probabilities of success 0·60	**0·80**	0·90	0·95
				5	10	15	20
1	3	10	2/1	0·481	0·683	**0·812**	0·891
2[a]	1	5	0/1	0·370	0·544	0·682	**0·787**
3[b]	3	15	3/0	0·539	**0·744**	0·861	0·925
4	3	10	2/1	0·512	0·717	**0·840**	0·915
5	3	15	2/1	0·467	0·667	**0·799**	0·882
6	2	15	2/0	0·500	0·703	**0·829**	0·903
7[c]	1	10	1/0	0·519	0·723	**0·845**	0·914
8	3	15	3/0	0·553	**0·757**	0·870	0·931
9	1	10	1/0	0·567	**0·771**	0·880	0·938

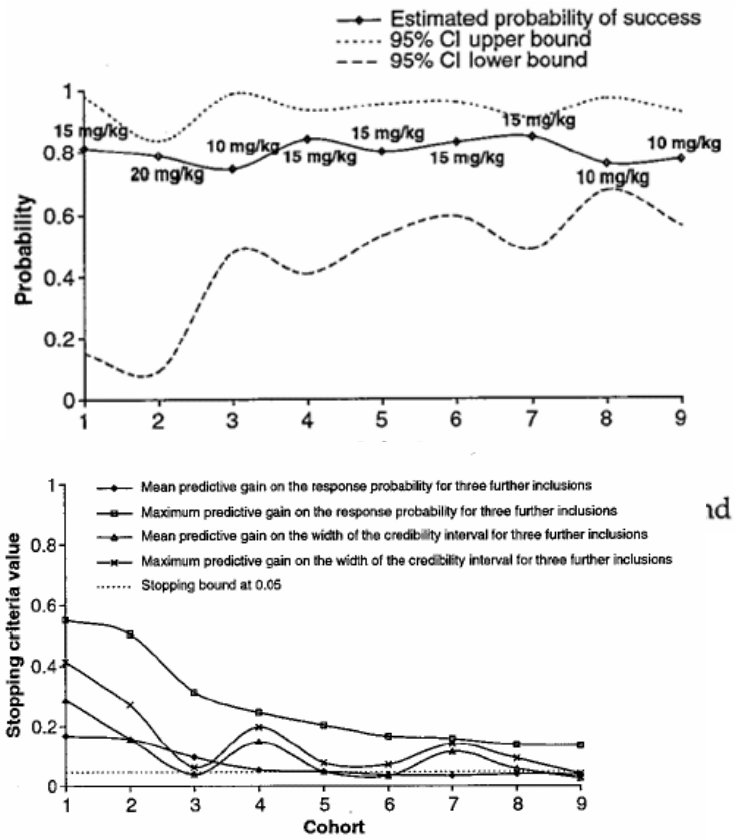


Fig. 2. Stopping criteria based on the predictive gain (mean and maximum) of further patient inclusions on the estimated response probability and the width of its 95% credibility intervals for the ≥27-week PMA group.

KING'S College LONDON
University of London

Adaptive Interim Designs

Adaptive Interim Designs
Bauer & Koehne (Biometrics,1994)

- General Strategy - One or two adaptive interims
  - Adaptive
    - re-assessing sample size
    - reduction of the set of multiple end-points
    - selecting a sub-set of doses in a dose-response design
    - in a placebo controlled trial changing the dose based on safety with ok efficacy

Adaptive Interims

- Basic Idea - Single Adaptive Interim - Maintain global α
  - Conduct a first stage - observed p-value - p_1
    - stop for futility - ($p_1 > \alpha_0$)
    - stop for efficacy - ($p_1 < \alpha_1$)
    - re-estimate sample size for second stage
  - Conduct second stage - observed p-value - p_2
  - Combine p_1 and p_2 by Fisher's method

    $p_1 p_2 < c_\alpha = \exp(- ½ \bullet \chi^2_{4,\alpha})$

  - $\alpha_1 + c_\alpha (\ln \alpha_0 - \ln \alpha_1) = \alpha$ (to be solved for α_1 by iteration)
- Bauer and Kieser (SIM, 1999)
  - Combining different phases in a single trial
  - Should each stage be given equal weight
  - Optimal choices for α_0 and α_1
  - k stage

Adaptive interim
An example

Cite this article as: BMJ, doi:10.1136/bmj.38356.655266.82 (published 11 February 2005)

Papers

Acute treatment of moderate to severe depression with hypericum extract WS 5570 (St John's wort): randomised controlled double blind non-inferiority trial versus paroxetine

A Szegedi, R Kohnen, A Dienel, M Kieser

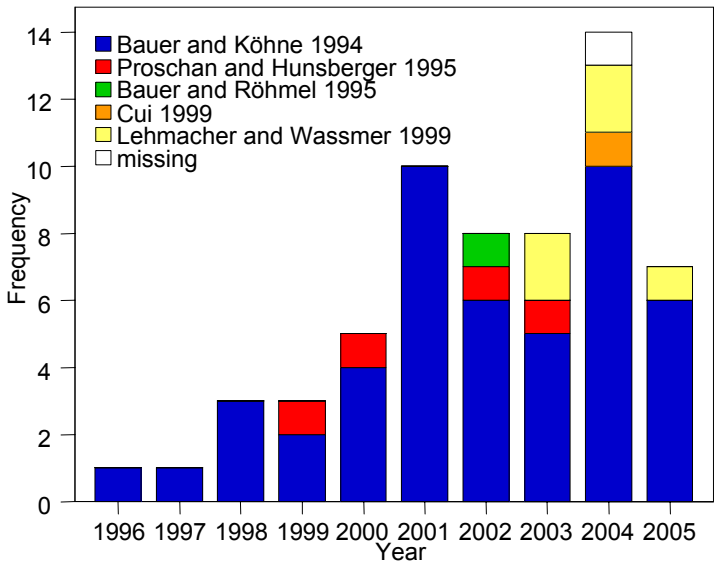
- Non-inferiority trial, Δ=2.5 (change in HAMD)
- Recruited 100 patients into part 1
- At end of part 1, p_1=0.084 for non-inferiority
- To gain overall significance, need p_2<0.0038/0.084 = 0.045
- Need 150 patients in part 2 for 80% power for significance at one-sided α=0.045 for non-inferiority
- Achieved p_1p_2<0.0038 – details not available in publication

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

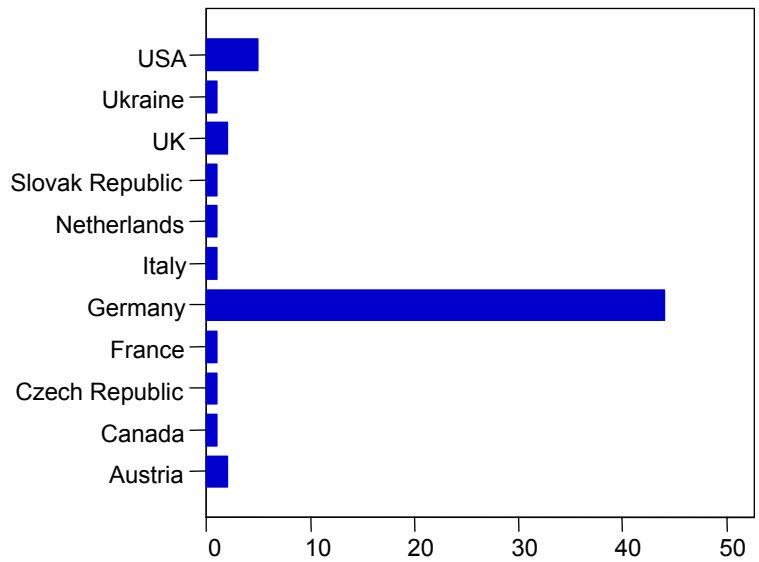
- Identified 75 papers dealing with adaptive designs : 1989-2004
  - combination tests
  - conditional error function
  - did not consider Bayesian approaches

- 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

KING'S College LONDON
University of London

Bauer and Einfalt Conclusions

- 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

Bayesian Adaptive Randomisation

KING'S College LONDON
University of London

Bayesian Adaptive Randomisation Thall and Whalen (Eur J Cancer, 2007)

- Back to the idea of Thompson (1933)
- Similar to RPW – binary outcome
- Randomisation to treatment A on the basis of

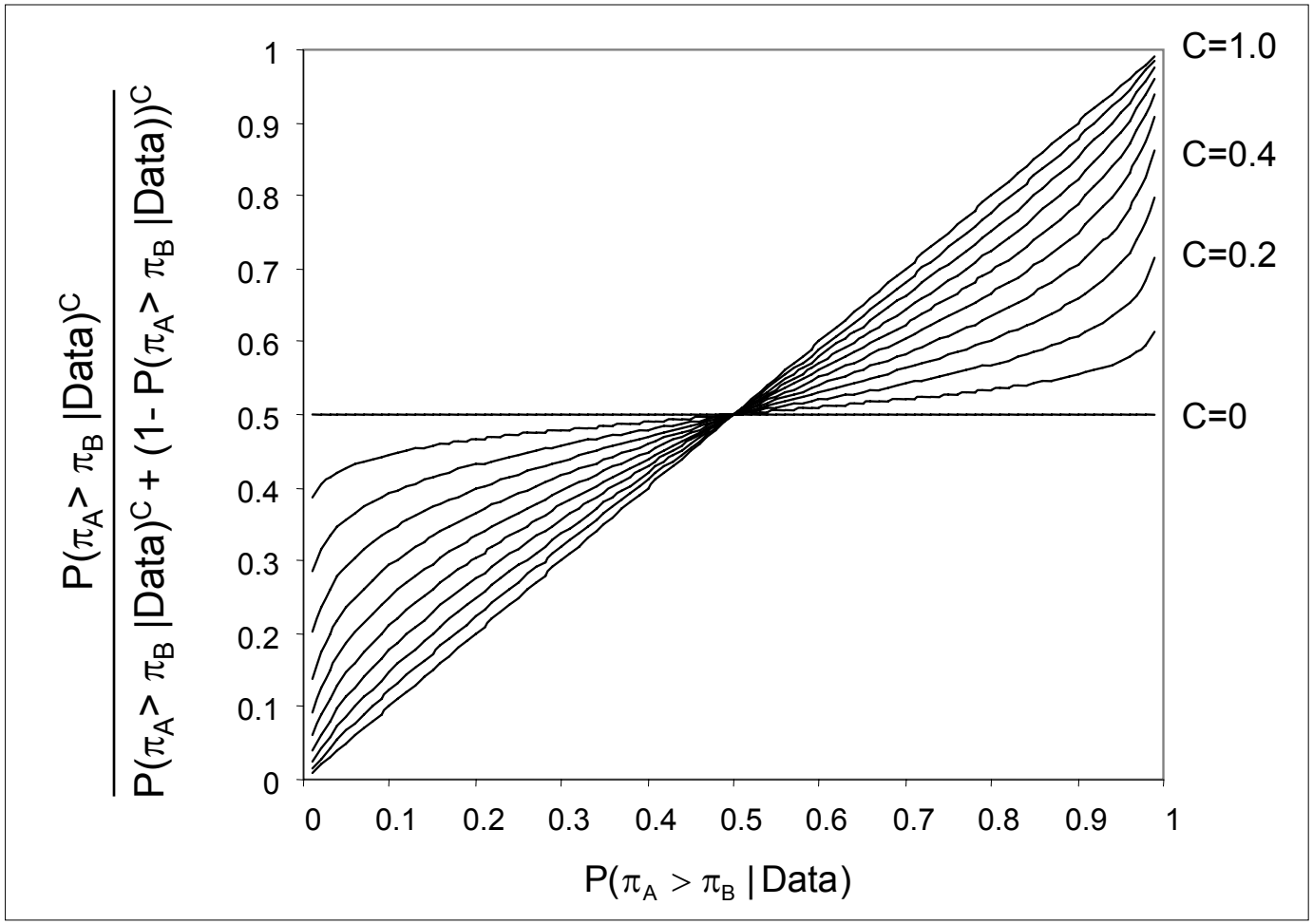
$$P(\pi_A > \pi_B \mid Data)$$

- Unstable

$$\frac{P(\pi_A > \pi_B \mid Data)^C}{P(\pi_A > \pi_B \mid Data)^C + \left(1 - P(\pi_A > \pi_B \mid Data)\right)^C}$$

KING'S College LONDON
University of London

Bayesian Adaptive Randomisation
Impact of Choice of C



KING'S College LONDON
University of London

Bayesian Adaptive Randomisation Impact of Choice of C

- Thall and Whalen recommend C= n/(2N)
  - n=current sample size
  - N=study's maximum sample size
- Begins with C=0, ends with C=1/2
- C=1/2 "works well in many applications"
- Giles et al (J Clin Oncology,2003)
  - Similar idea
  - $P(m_1<m_0|data)$ - m_1, m_0 median survaival times

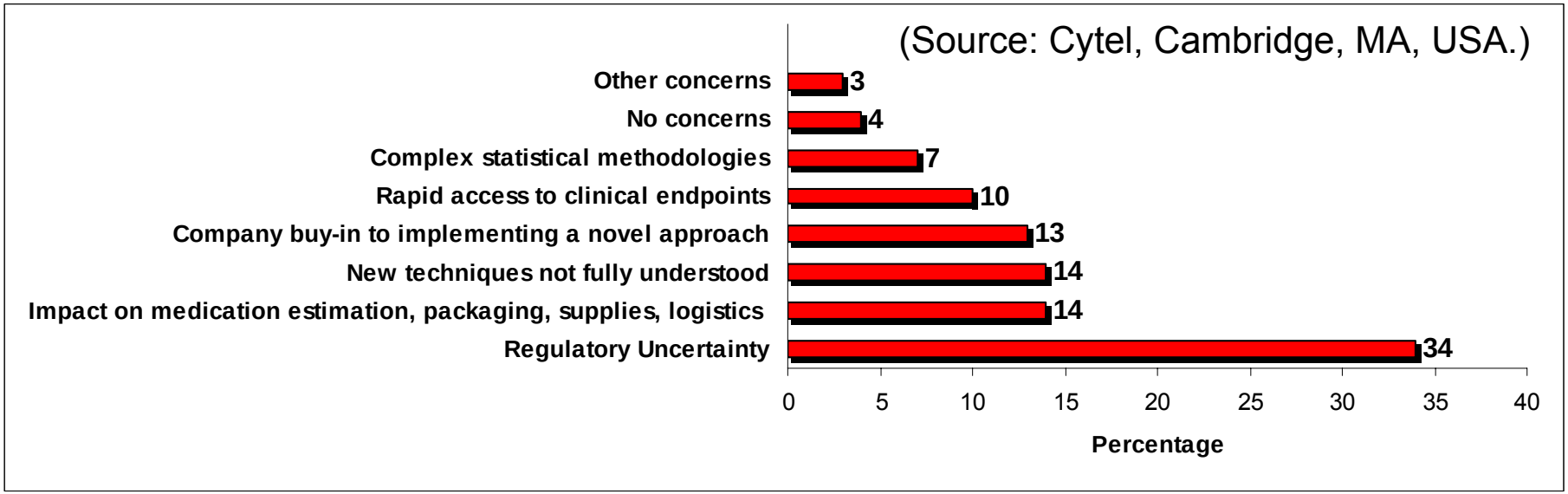
Fixed Sample vs Planned Adaptive Designs
Dragalin (Drug Information J, 2006)

4 Rules define an adaptive design –

Planned Adaptive Designs	Fixed Sample
Allocation Rule ◆ Defines how pts are allocated to arms. Can be fixed but can change based on accruing data	◆ Randomisation remains fixed throughout study
Sampling Rule ◆ How many subjects sampled at next stage (cohort size)	◆ Only one stage
Stopping rule ◆ When to stop a trial: efficacy, futility	◆ No stopping
Decision Rule ◆ Final analysis or interim changes not covered by the above 3 (eg dropping arms)	◆ No Changes

Why Does the debate Continue as to their Worth?

Perceived Barriers to Implementing Adaptive Designs



Christopher Thomas Scott & Monya Baker
Overhauling Clinical Trials
Nature Biotechnology, Volume 25(3),2007, 287-292.

KING'S College LONDON
University of London

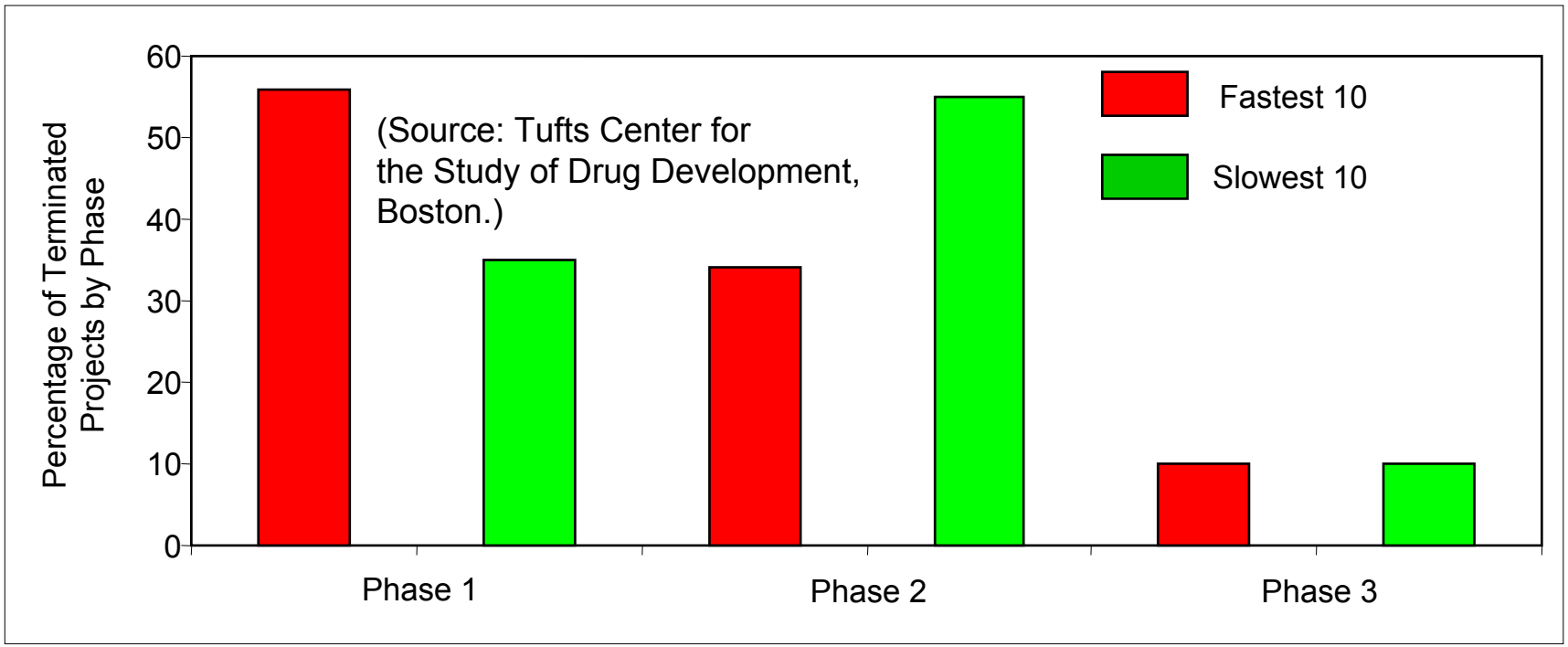
Objectives

- Faster
- Cheaper
- Better

Learning

- At an individual study level
  - May not be faster – because speed is anathema too adaptive learning
  - May not be cheaper – we may need more doses to properly characterise dose response
    - EMAX models
- BUT
  - Reduce the risk of recycling
  - Shorten the whole development program
  - Deliver better information
  - Increase Phase III success

Percentage Projects Terminated by Phase 1994-2005



Christopher Thomas Scott & Monya Baker
Overhauling Clinical Trials
Nature Biotechnology, Volume 25(3),2007, 287-292.

KING'S College LONDON
University of London

The Benefits of the Adaptive Debate

- Better planning
  - Comparison of adaptive designs with "traditional designs" will improve the latter if they are the most appropriate"
  - Sample Size Re-estimation
    - We don't know enough about the nuisance parameters
    - Acknowledge that in planning
    - Uncertainty implies large sample sizes
    - Combine with early stopping for futility or efficacy