Using Simulations to Optimize Drug Development Decision Making

Tom Parke

Presentation at EFSPI

12th December 2018

Drug Development

- As a human endeavour, drug development is unique:
  - Very high probability of failure
  - Value is highly time sensitive
  - It's a process of scientific enquiry
  - Value, if successful, dominates cost

Drug Development Optimization

- So optimizing it can be about
  - *Reliably* spotting failure early and transferring resources to other projects
  - Trading probability of success against time
  - Optimally reducing uncertainty
  - And not so much about reducing cost

How to select designs for early trial phase trials?

- 3+3 vs mTPI vs N-CRM
- Number of doses tested in Phase II
- What endpoint(s) to study / use for decision making in Phase II (c/w Phase III)
- Phase II size
- Post Phase II go/no-go decision threshold
- Phase II dose selection criteria
- Seamless Phase II/III
- Early stopping in Phase II & III
- Safety & tolerability vs efficacy

Estimating the value of a drug development program

- Expected revenue if successful = R(eT, Ti)
  - eT = estimate of treatment effect
  - Ti = Time to registration
- Probability of success= P(N, T, SD, Thr)
  - N = sample size
  - T = true treatment effect
  - SD = SD of endpoint (or other 'nuisance' parameter)
  - Thr = decision threshold
- Time = $T_i(N, A)$
  - A = accrual rate
- Cost = C(F, T, N)
  - F = fixed
- Value = R * P – C
- Value Phase II & Phase III = $R * P_2 * P_3 – C_2 – C_3 * P_2$

Calculation of value

- Value = R * P – C
- Value = $R(eT, T_f+T_i(N, A)) * P(N, T, SD, Thr) – C(F, T_i(N, A), N)$
- Value Phase II & Phase III = $R * P_2 * P_3 – C_2 – C_3 * P_2$
- Value Phase II & Phase III =
  $R(eT, T_f+T(N_2, A_2) + Ti(N_3, A_3)) * P(N_2, T, SD, Thr_2) * P(N_3, T, SD, Thr_3) –$
  $C(F_2, T_i(N_2, A_2), N_2) –$
  $C(F_2, T_i(N_2, A_2), N_2) * P(N_2, T, SD, Thr_2)$
- Note that increasing N increases time (decreasing revenue), increases power and increases cost.
- Note that T (treatment effect) in Phase III and eT for Revenue will often depend on decisions in Phase II such as:
  - selection of treatment/dose,
  - selection of patient population to treat / biomarker / biomarker threshold

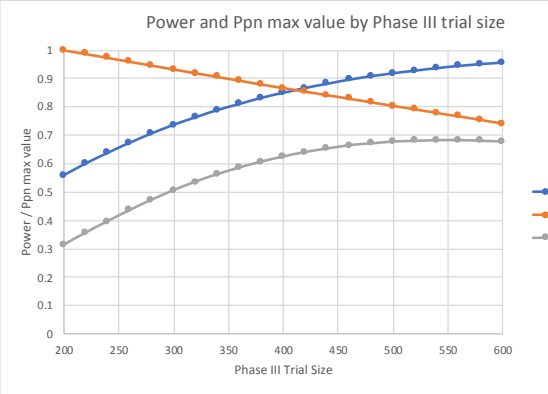
Very Simple Illustration

- We are testing a treatment with the expected mean treatment effect of 1.3-1.5 points on a cts endpoint
- SD 5.
- Our prior expectation that the treatment works is 50%.
- Two phase III trials to be significant at 0.025 (one sided)
- We discount future treatment at 8% per year.
- We have a 'revenue horizon' of 10 years … patent expiry, competition, compound discounting is 43%
- The expected market share is ~100,000 patients at a net revenue of between $5,000 and $2,500 per patient.
- The Phase III trials will be run in parallel and we can recruit into each phase III at an average rate of 150-250 subjects per year.
- Subjects in Phase III cost $20,000

Very simple illustration cont'd

- In a spreadsheet a row for each Phase III sample size to be considered.
- Compute power of the Phase III
- Time to registration/clinical use
  - Fixed time: plan Phase III, time to analyze data and submit after Phase III
  - Variable time: time to recruit Phase III subjects
- Value for each year over 10 years
  - 0 If not in use yet, 1=reduction in unmeet medical need if in clinical use in year 1.
  - 1/discount rate^{year}
  - Sum to give total discounted value
- To derive probability weighted value: multiply total discounted value by prior probability treatment is effective and by the power of the two phase IIIs
- Subtract cost of Phase III per subject * number of subjects.
- E.g. Cost:
  - probability of being untreated (Pr(control) + Pr(treatment)*Pr(treatment ineffective) ... E.g. 0.5 + 0.5 * 0.5 = 0.75
  - Weight of 1 subject = 1/potential treatment population

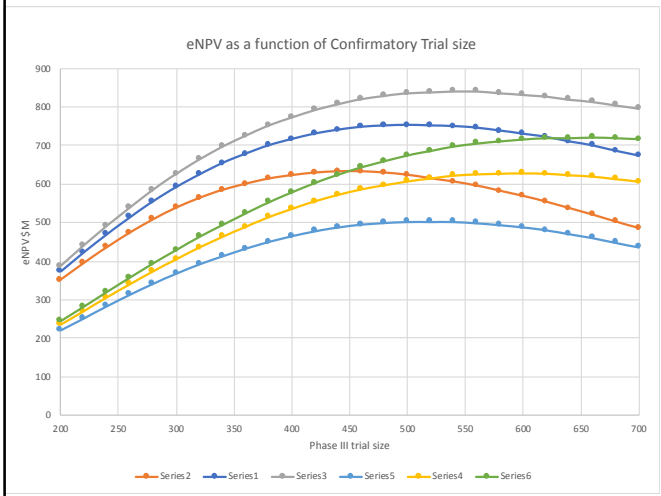
Power and value illustrated



Change in ppn of max value – declines linearly with increase in sample size.
Due to time taken.

Expected value is ppn of max value * power * power (as there are 2 phase IIIs, both of which have to be successful

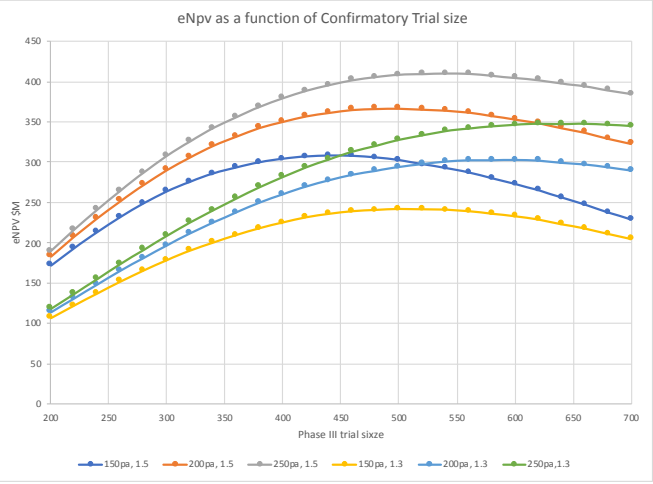
Total expected treatment value illustrated (trt revenue $5,000)



Note that curves are quite flat round the maximum
6 curves for 2 treatment effects (1.5 & 1.3) and 3 recruitment rates (250, 200, 250 pa)

Best Phase III size, are 460, 500, 540 for treatment effect of 1.5 and
500, 600, 660 for 1.3

Total expected treatment value illustrated (trt revenue $2,500)



Note that curves are quite flat round the maximum
6 curves for 2 treatment effects (1.5 & 1.3) and 3 recruitment rates (250, 200, 250 pa)

Best Phase III size, are 440, 500, 540 for treatment effect of 1.5 and
500, 600, 660 for 1.3

So despite halving the expected revenue, only one maximum changed.

	$2,500						$5,000					
	1.3			1.5			1.3			1.5		
	150	200	250	150	200	250	150	200	250	150	200	250
440	235	277	304	**308**	362	395	488	572	625	633	741	808
460	239	284	313	307	365	402	496	587	644	**633**	748	821
480	240	289	320	305	366	406	500	597	660	630	752	830
500	**241**	294	328	302	**367**	408	**503**	607	675	624	**753**	837
520	241	297	333	297	365	409	502	615	687	615	752	839
540	240	300	338	292	364	**410**	502	622	698	606	750	**841**
560	238	302	342	287	362	409	499	626	706	596	746	841
580	236	302	344	280	358	407	494	627	712	583	738	837
600	232	**302**	346	272	353	404	489	**628**	717	569	731	833
620	228	301	348	265	349	402	481	628	720	554	722	828
640	223	299	347	256	342	397	471	623	720	537	710	820
660	217	297	**348**	247	337	394	461	620	**722**	521	700	814

Selecting in uncertainty

- Say these 12 scenarios 'capture' our uncertainty, and we weight them equally
- (Clearly we could easily add more scenarios and weight them differently)
- We can simply average the eNPV for each sample size over all the scenarios
- And pick the sample size with the greatest expected eNPV

Average eNPV

Sample size	eNPV $M
440	479
460	487
480	491
500	495
520	496
540	497
560	496
580	493
600	490
620	485
640	479
660	473

How does the 'average' maximum compare to the per scenario maximum?

	$2,500						$5,000					
	1.3			1.5			1.3			1.5		
	150	200	250	150	200	250	150	200	250	150	200	250
Per scenario	**241**	**302**	**348**	**308**	**367**	**410**	**503**	**628**	**722**	**633**	**753**	**841**
Average	240	300	338	292	364	**410**	502	622	698	606	750	**841**
Difference	-1	-2	-10	-16	-3	0	-1	-6	-24	-27	-2	0

A couple of key findings:

1. Our optimal design choice is not impacted by our uncertainty in revenue
2. An 'average' optimal design choice is not far off optimal across all our scenarios

ESTIMATING OVER PHASE II & PHASE III

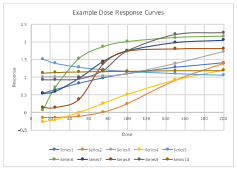
Approaches

- Spreadsheet
- Decision tools
- Bayesian Network

- These can evaluate just phase III, or very simple phase II & phase III

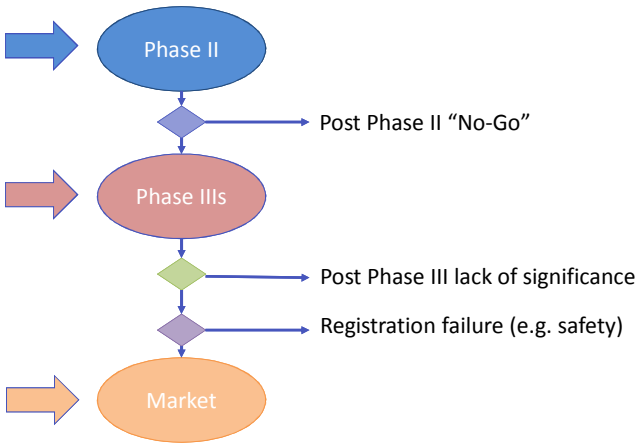
- But as soon as any complexity added we always end up with evaluation based on trial simulations, because of impact of Phase II selections on Phase III and Revenue

Considering Phase II as well

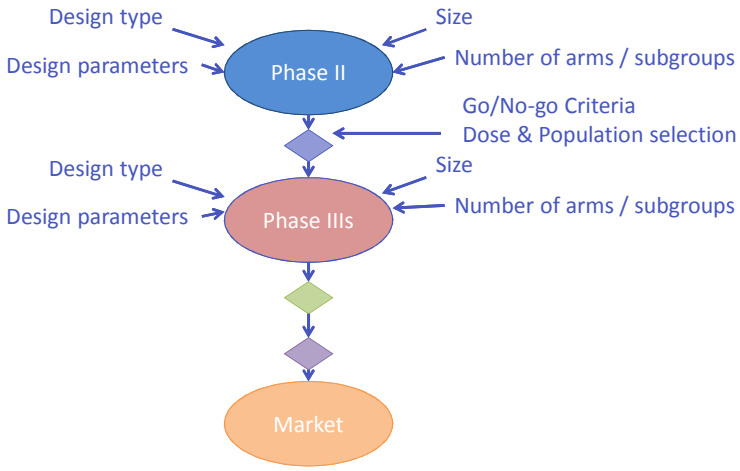


Phase III depends on actual response of selected dose, & possibly on estimate of response

Value in market depends on time to market, estimate of effect, tolerability



Allowing us to Evaluate



Impact?

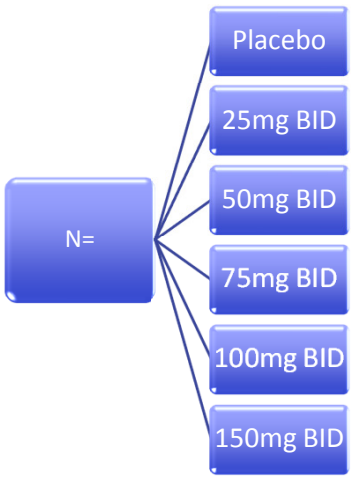
- Potentially huge
- Design trials in light of need & cost-benefit, not just plugging in std numbers
- Evaluate complex trade-offs:
  - Which sub-groups
  - Test one treatment or several
- Raise profile of statistics in Pharma to decision making level – as it is in most other industries

EXAMPLE (ISCTM PAPER – SCHIZOPHRENIA)

Study Proposal

- Multicenter, double-blind, randomized, placebo-controlled, parallel-group, dose-response study in male and female subjects with schizophrenia
- The primary objective is to evaluate the efficacy via change from baseline in the total Positive and Negative Syndrome Scale (PANSS) total score of multiple fixed doses of Compound X
- Primary outcome= change from baseline to endpoint in PANSS total score
- Placebo controlled, 6 wks in duration
- Minimal effectiveness = 10 point difference from placebo (SD = 20)

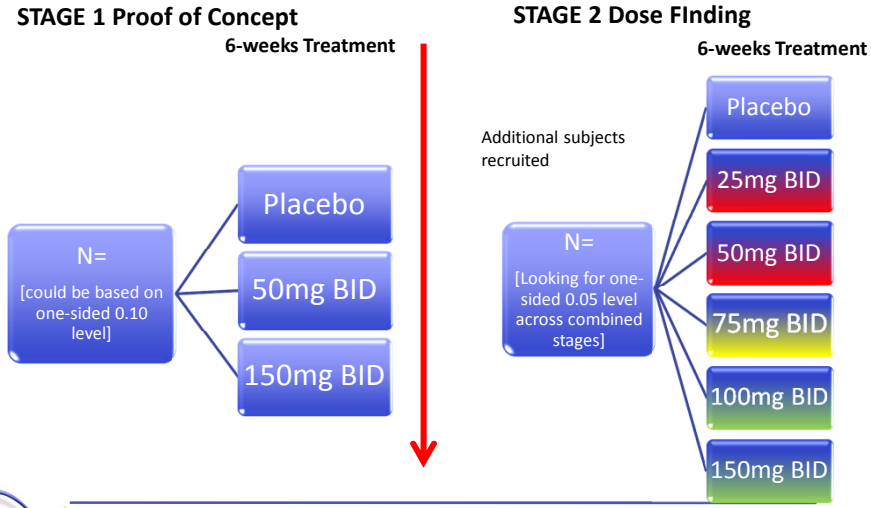
Available Study Doses for the New Trial



75mg BID is not currently available

Strong believe that both 100mg and 150mg are effective

Possible Phase II Study Design in 2 stages:



Fixed Design – Stage 1

- Conventional, fixed 2a
  - 3 arms (placebo, 50mg & 150mg dose), 25 per arm, if successful run 2b
  - Use posterior probability of being better than control to judge success (non-informative prior – approximately equivalent to a one-sided p-value test with Bonferonni adjustment)
  - If $Pr(\theta_{50mg} < Control) > 0.9$, run 2b with: 4 arms (placebo, 25, 50mg & 150mg doses),
  - Otherwise if $Pr(\theta_{150mg} < Control) > 0.9$, run 2b with: 3 arms (placebo, 100, 150mg doses),
  - Otherwise neither dose successful in phase 2a, don't run phase 2b.

Fixed Design – Stage 2

- Conventional fixed 2b,
  - Sample size 250 (currently same sample size used regardless of # of arms)
  - equal allocation between control and 2/3 chosen doses
  - if dose with Max effect, $Pr(\theta_{max} < Control -7) > 0.5$ run Phase III
  - If no dose > CSD, phase 2 futile

Example adaptive design

- Adaptive combined 2a and 2b
  - Allocate first 75 subjects equally to control, 50mg and top dose, then add 2 more doses 25, & 100mg
  - After recruiting the 75th subject perform the first interim
  - Update randomisation every 2 weeks, favoring the dose most likely to be the minimum effective dose

Adaptive Design

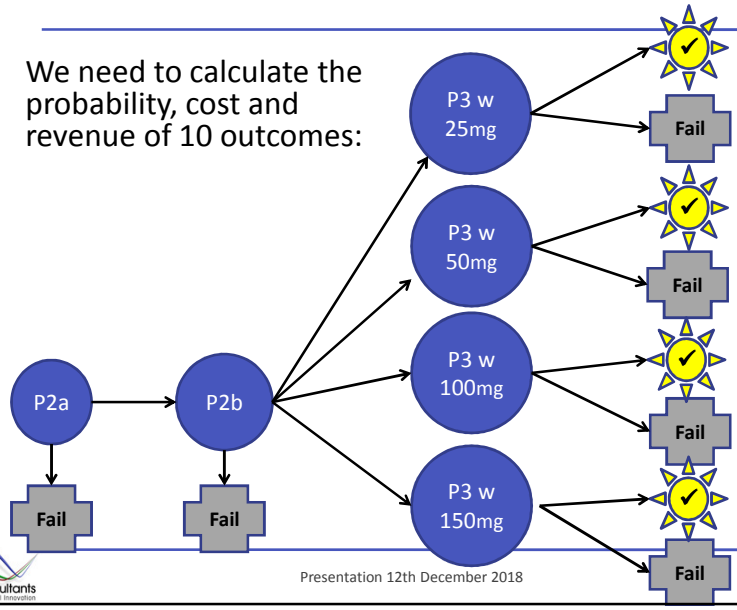
- Model dose response using 2nd order NDLM (allow response to be non-monotonic)
- Force tau to be 'on the high side' to ensure not too much smoothing
- Allocate 1/3rd subjects to control
- Stop for futility if
  - P(treatment difference < CSD) < 0.2
- Stop for success if
  - P(treatment difference < CSD) > 0.875
  - P(MED) > 0.6

Response scenarios

- Null and Weak – there is no revenue for this scenario, P3 trials so likely to fail and revenue if P3 successful so small that it can be ignored.
  - Thus best design for these scenarios will be the one that minimizes costs
- 3 with monotonic response (but different MEDs: 50mg, 100mg, 150mg)
- 2 non-monotonic (and different peaks 50mg, 100mg)
- Assume weights of 30, 15, 1, 1, 1, 1, 1
  - I.e. prior expectation of a successful drug: 10%
- In scenarios where a lower dose is 'good enough', higher doses are simulated having intolerability rates reducing revenue by 15% and 30%.

Thus in each success scenario

- We need to calculate the probability, cost and revenue of 10 outcomes:

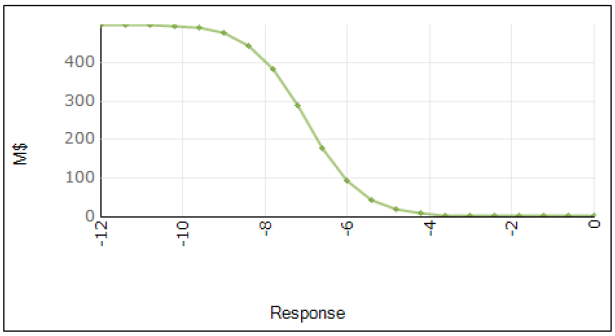


Success depends on dose

- Probability of success depends on the dose, and its effect size in that scenario
- To simply we assume P3 is the same for both fixed and adaptive programs, and fixed size (independent of result of phase 2)
- 2 phase 3 trials, each one:
  - 2 arms
  - 2-sided alpha 0.05
  - Power 0.9 for assumed mean difference of 8 points
  - 132 per arm
  - Actual power depends on true effect size

Expected Revenue depends on true treatment effect & time taken

Peak revenue modelled as sigmoid with a maximum of $500M, 50% of maximum at a -7pt treatment difference, slope of -1.5 to give ~0 revenue at -5pts and ~100% at -10pts



NOTE: revenue based on post-Phase III estimate of treatment effect, not "true"

NPV

- Total Revenue:

  Peak revenue * remaining patent life * discount
- Remaining patent life:

  10yrs – development time
- Discount = $(1 – 0.09)^{\text{time to revenue}}$

Time assumptions (fixed)

- 2 years elapsed already
- 6 months to prepare P2a
- 30 subjects per month (3 months to reach mean rec rate)
- 6 weeks follow up (P2a – 5.5 months elapsed)
- 6 months to prepare P2b
- … (P2b - 10.5 months elapsed)
- 6 months to prepare P3
- … (2xP3 – 17.5 months elapsed)
- 6 months to prepare submission
- 12 months to register
- ~5 yrs patent life remain

Time assumptions (adaptive)

- 2 years elapsed already
- 9 months to prepare P2a/b
- 30 subjects per month (3 months to reach mean rec rate)
- Take mean sample size given scenario and outcome, rounded up
- 6 weeks follow up
- 6 months to prepare P3
- … (2xP3 – 17.5 months elapsed)
- 6 months to prepare submission
- 12 months to register
- ~5.5 yrs patent life remain

Cost assumptions

- Development overheads: $1M pa
- Fixed P2a and P2b $1M overhead cost each
- Adaptive P2a/b $2.5M overhead cost
- P3 $4M overhead cost
- Trial cost per subject
  - P2 $60K
  - P3 $49K
- Cost of launch: $10M

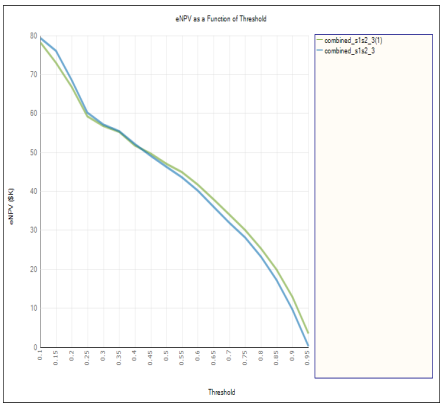
eNPV by scenario ($M)

Scenario	Fixed Design	Adaptive seamless P2a/b
Null	-11	-14
Weak	-10	9
High Dose	740	1,038
Middle Dose	752	1,287
Low Dose	558	995
Peak at 50mg	444	849
Peak at 100mg	338	1,105
Aggregate	**46**	**104**

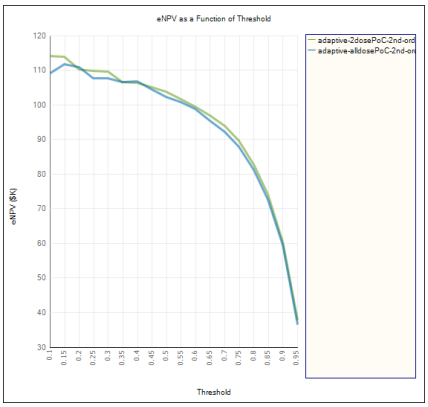
Adaptive designs advantages

- ~30% Greater revenue: nearly a year quicker
- Greater power in Phase II (while still keeping probability of success in Phase II ~0.026)
- Greater probability of winning in Phase III (given success in Phase II) due to better dose selection

Optimal end of phase II go/no-go decision threshold

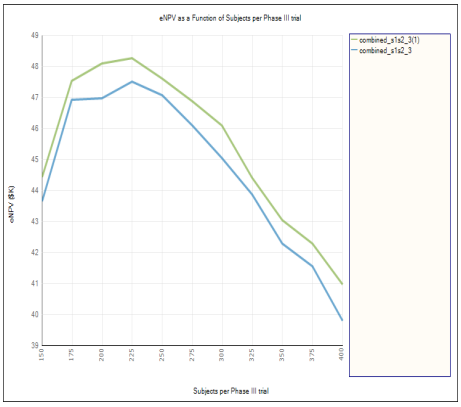


Fixed two stage design

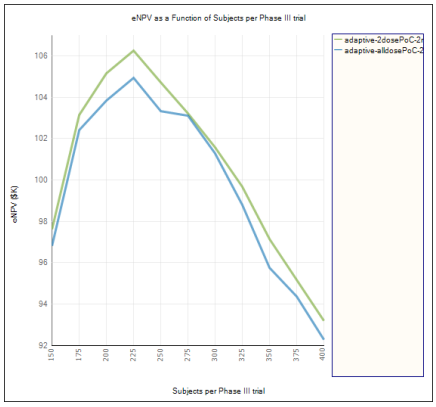


Adaptive design

Optimal pre-specified phase III sample size

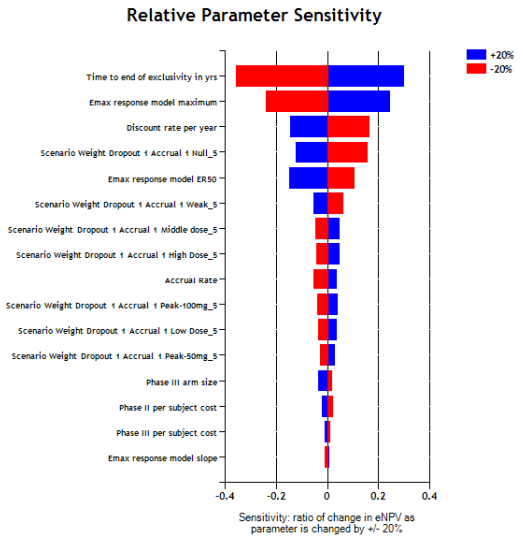


Fixed two stage design



Adaptive design

Parameter sensitivity analysis



Berry Consultants
Statistical Innovation

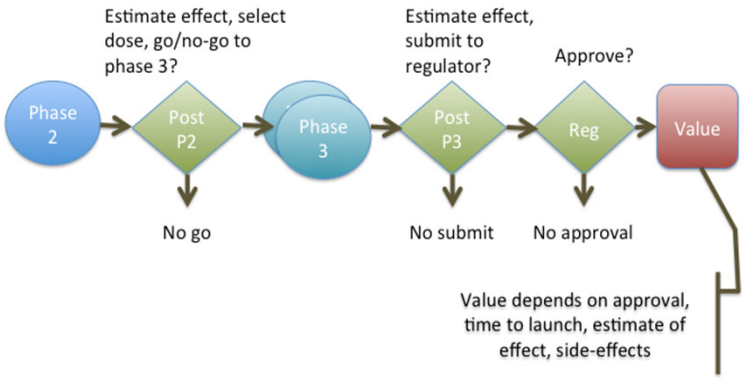
"QUOTES"

Berry Consultants
Statistical Innovation

3 Modules

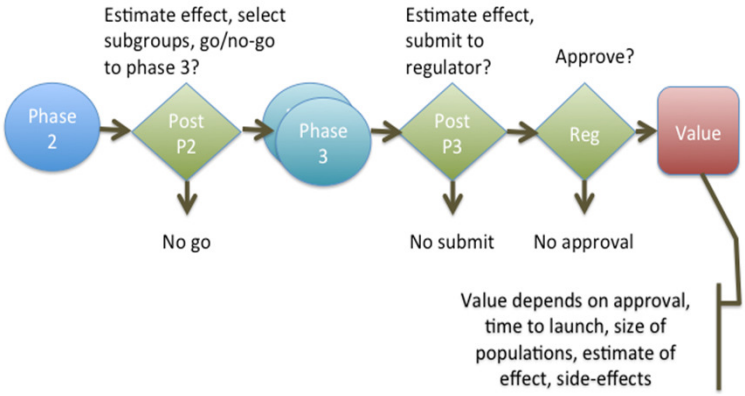
1. Core Module: Evaluating a conventional sequence of a phase 2 followed by phase 3 trial(s) or a single complex phase 2 or 3 trial.
2. Umbrella Module: This module allows the evaluation of less conventional trials: Basket and Platform trials
3. Staged Module: This module allows for a sequence of complicated, innovative, trials namely an adaptive phase 2 followed by a possibly adaptive phase 3, with possibly multiple doses/treatments being tested in phase 3.

Program Decision type: 1



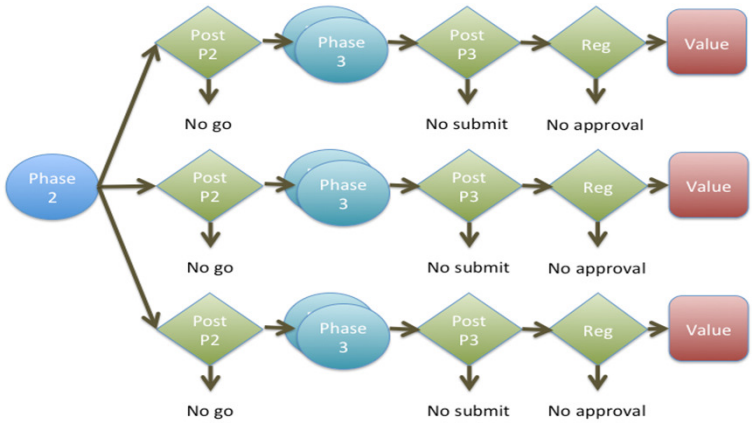
Dose/Treatment selection, followed by Phase 3 in the selected treatment

Program Decision type: 2



Sub-group finding, basket trials, followed by Phase 3 in the selected population (made up of the selected sub-groups)

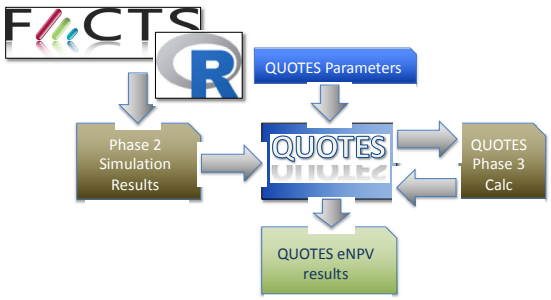
Program Decision type: 3



Simple platform trials followed by separate Phase 3 in each of the successful treatments.

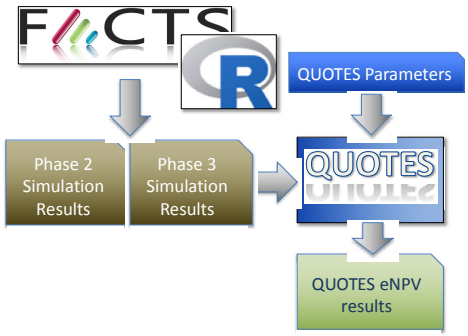
Architecture: Phase 3 in QUOTES

Core Module, FACTS or R simulated Phase 2, simple Phase 3 simulated trials within QUOTES



Staged Design Module

FACTS or R simulated combined Phase 2 and Phase 3 trials



SUMMARY

Supporting drug development decisions using simulation and estimation of NPV

- There are many uncertainties!
- BUT the most important seems to be the “true” drug effect, then accrual rates
- Differences in expected revenue make a big difference to overall expected NPV … but has little impact on “what is optimal” for a development program
- The eNPV is rarely super sensitive to the typical development program design parameter values
- Incremental improvements in the trial design yield incremental improvements in eNPV, but cumulatively they can yield significant improvements.

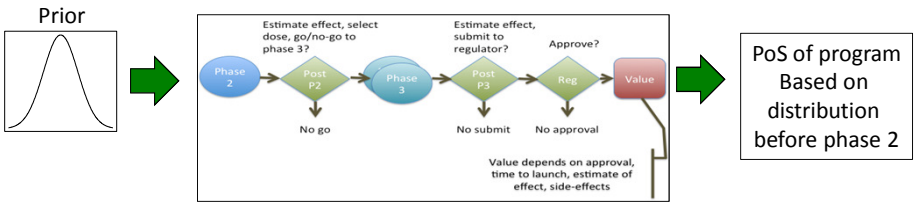
QUOTES BACKUP SLIDES

External Simulations

- Trial Simulations are external to PDMS (but it can do simple Phase III simulations internally).
- Simulations are supplied as a file per 'design' with a row per simulation.
- You can map columns in the input to the columns PDMS requires:
  - The true response (of each arm)
  - The true SD of the response (if continuous) or hazard rate on control (if time-to-event).
  - The observed response
  - The observed SD of the response (if continuous) or hazard rate on control (if time-to-event).
  - The duration of the simulated trial
  - The number of subjects
  - The statistical estimate for use in go/no-go decision (e.g. p-value, posterior probability)
  - The arm selected (if a multi-arm trial)

Bayesian PoS

- In order to model the probability of success (PoS) of a trial or program you must have the ability to simulate truths from a prior distribution:



- This Bayesian PoS will be implemented in each module

PDMS model parameters

- The phase III
- Toxicity and (in) tolerability of the arms / groups
- Scenario weights
- Development time & costs
- Registration time & costs
- Peak net revenue (e.g. as a function of observed treatment effect compared to control)
- Net revenue profile over time

Phase III Parameters

- Length of follow up, accrual rate, simple Poisson based simulation of accrual
- Fixed or Group Sequential
  - Alpha, Significance Margin, Control or Objective Control, Number of Trials, Superiority/Non-inferiority
  - GS: number of interims, alpha spending function, fraction of information before first interim, futility only, success only or both.
- Fixed Size or Size based on observed effect in Phase II
- Possible dilution of treatment effect in Phase III

Toxicity and (In)tolerability

- For specific scenarios and doses, subgroups or indications a
  - Probability of excessive toxicity being observed in Phase III (and no submission to regulator)
  - Ppn of market share lost through lack of tolerability
- If that dose, group or indication is taken to Phase III

Scenario Weighting

- Either the supplied sims can be 1 scenario with underlying mean response rate drawn for each simulation from a distribution.
- Or the supplied sims can be from a set of scenarios where each scenario uses a fixed underlying mean response rate.
- In the latter case the different scenarios can be given different weights.

Development Time

- Stages between trials are assumed fixed in length (user specified)
- Trial durations are taken from the simulations
- Stages are:
  - Pre Phase II
  - Phase II
  - Pre Phase III
  - Phase III (multiple Phase IIIs assumed to be in parallel)
  - Pre-registration
  - Registration

Development Costs

- Can be specified as a fixed cost per stage
- Can be variable cost dependent on
  - Time
  - Subjects
  - Number of Sites

  Plus a fixed cost

  Specified per stage
- Separate discount rates for costs and revenues

Registration Time and Costs

- Possible open label extension costs
- Probability of registration
- Time for Normal registration or Priority Registration plus Pr(Normal Registration)
- Cost of Market launch

Revenue

- We model Revenue net of manufacturing, distribution, sales and marketing costs
- A peak revenue model can be defined dependent on the observed response in Phase III
- A Revenue Profile over time can be specified (% of peak revenue per year)
- Expect time to end of market exclusivity (patent expiry / significant competitor)
- This time specified from start point of simulation of development
- Ramp down at end of exclusivity

Calculate eNPV

- And ROI & PI
- For all the loaded designs

Optimization

- Currently
  - End of Phase II go/no-go decision threshold
  - Phase III parameters
    - Size
    - Min/max size and required power
    - GS parameters