Enabling the use of Bayesian Methods for Interim Analyses in Early Clinical Development at Pfizer

Mark Whitlock

Statistics Group Head – ECD Internal Medicine, Pfizer

PSI Conference 2019

Early Clinical Development

Pfizer WORLDWIDE RESEARCH & DEVELOPMENT

SCIENCE FOR LIFE-CHANGING IMPACT

Contents

- Motivation
- Bayesian Predictive Probability
- Example
- Challenges
- Guidance Document
- Software
  - Worked Example
- Summary

Motivation

- Situation:
  - At Pfizer, Bayesian methods are widely used for decision making in early (Phase 2) clinical trials.
  - Studies have pre-defined criteria for success. Pfizer approach;
    - C1: 95% (75%) confident effect > placebo (LRV)
    - C2: Observed effect > Target Value
  - Interim Analyses often included to enable study adaptations (e.g. stop for futility, accelerate development planning, add or drop an arm)
  - Bayesian framework allows intuitive approach. Decisions are based on the probability of meeting end-of-study decision criteria based on the existing interim data.
    - E.g. Stop for futility if probability of meeting C1 is < x%, accelerate planning if probability of meeting C2 is > y%.
- Problem:
  - Teams encouraged to use Bayesian predictive probability approach but implementation inhibited by;
    - Technical challenges
    - Time to create and validate approach / code
    - Potential for errors
- Solution:
  - Team set up (Donal Gorman / Yao Zhang + input from other experts) to develop internal guidance document and software (R-Shiny Apps) to standardise approach and enable implementation.

Bayesian Predictive Probability

- The basic framework for predictive probabilities comes from Grieve 1991
- Analytical derivation for simple case;
  - Normally distributed response, 2-arm PG, n_1 subjects per arm at interim, n_2 remaining subjects per arm, $\bar{d}_1$ mean difference at interim, $\bar{d}_2$ mean difference of remaining subjects, δ is mean difference at study end.
- For the one-sided significance test of H0: δ=0 vs H1: δ>0 we will, at study end, reject H0 if:

$$\frac{n_1\bar{d}_1 + n_2\bar{d}_2}{(n_1+n_2)\sqrt{\frac{\sigma_D^2}{(n_1+n_2)}}} > Z_\alpha$$

- In terms of the random variable $\bar{d}_2$ we require:

$$P(\bar{d}_2 > \frac{Z_\alpha(n_1+n_2)\sqrt{\frac{\sigma_D^2}{(n_1+n_2)}} - n_1\bar{d}_1}{n_2})$$

- It can be shown (given predictive distribution of $\bar{d}_2$) that predictive probability of rejecting Ho given by:

$$1-\Phi\left(\frac{\frac{Z_\alpha(n_1+n_2)\sqrt{\frac{\sigma_D^2}{(n_1+n_2)}} - n_1\bar{d}_1}{n_2} - \bar{d}_1}{\sqrt{\frac{\sigma_D^2}{n_1}+\frac{\sigma_D^2}{n_2}}}\right)$$

Grieve AP (1991) Predictive probability in clinical trials: Biometrics, 47(1) 323-9..

Example 1: Dental Pain Study

- Parallel group study with 4 treatment arms (PF 1000mg, PF 2000mg, Placebo, Ibuprofen 400mg)
  - N=35 per group planned
- Primary endpoint TOTPAR[6]
  - Area under the pain relief curve (0-4 scale) through 6 hours after dosing following surgical extraction of wisdom tooth
- Primary analysis: ANCOVA model with treatment and baseline pain
- Decision Criteria:
  - C1: At least 95% confident that PF will be superior to placebo.
  - C2: At least 25% confident that PF has a TOTPAR[6] greater than 6, compared with placebo.
- Interim analysis conducted after N=90 of 140 subjects completed
- Bayesian predictive probability of achieving the C1 criteria at the end of the study, given the data observed at the interim, calculated.
- The following decision rule was pre-defined:
  - *STOP for futility if the Bayesian predictive probability of passing C1 is less than 10%.*
- The criteria must be met for both PF doses.
  - If the criteria is not met then the study will continue without any modification to the design.

Example 1: Dental Pain Study

- ANCOVA carried out on interim data (90 subjects):

	1000 mg (N=22)	2000 mg (N=23)	Ibuprofen (N=22)	Placebo (N=23)
TOTPAR [6]				
LS mean	4.3	4.4	8.8	5.3
SE	1.5	1.5	1.5	1.5
90% CI	1.8, 6.8	2.0, 6.9	6.3, 11.3	2.9, 7.8
Treatment differences (vs placebo)				
LS mean	-1.1	-0.9	3.5	
SE	1.9	1.9	1.9	
90% CI	-4.3, 2.2	-4.1, 2.3	0.3, 6.7	

- Predictive probability of passing C1 (At least 95% confident that PF will be superior to placebo) at end of study, given interim data, calculated for each dose.
- The predictive probability was 0.14% for 1000mg and 0.13% for 2000mg
  - Both probs < 10%. Study stopped for futility
- Note that ibuprofen did show a statistically significant effect demonstrating trial success

$$1-\Phi\left(\frac{\frac{Z_\alpha(n_1+n_2)\sqrt{\frac{\sigma_D^2}{(n_1+n_2)}}-n_1\bar{d}_1}{n_2}-\bar{d}_1}{\sqrt{\frac{\sigma_D^2}{n_1}+\frac{\sigma_D^2}{n_2}}}\right)$$

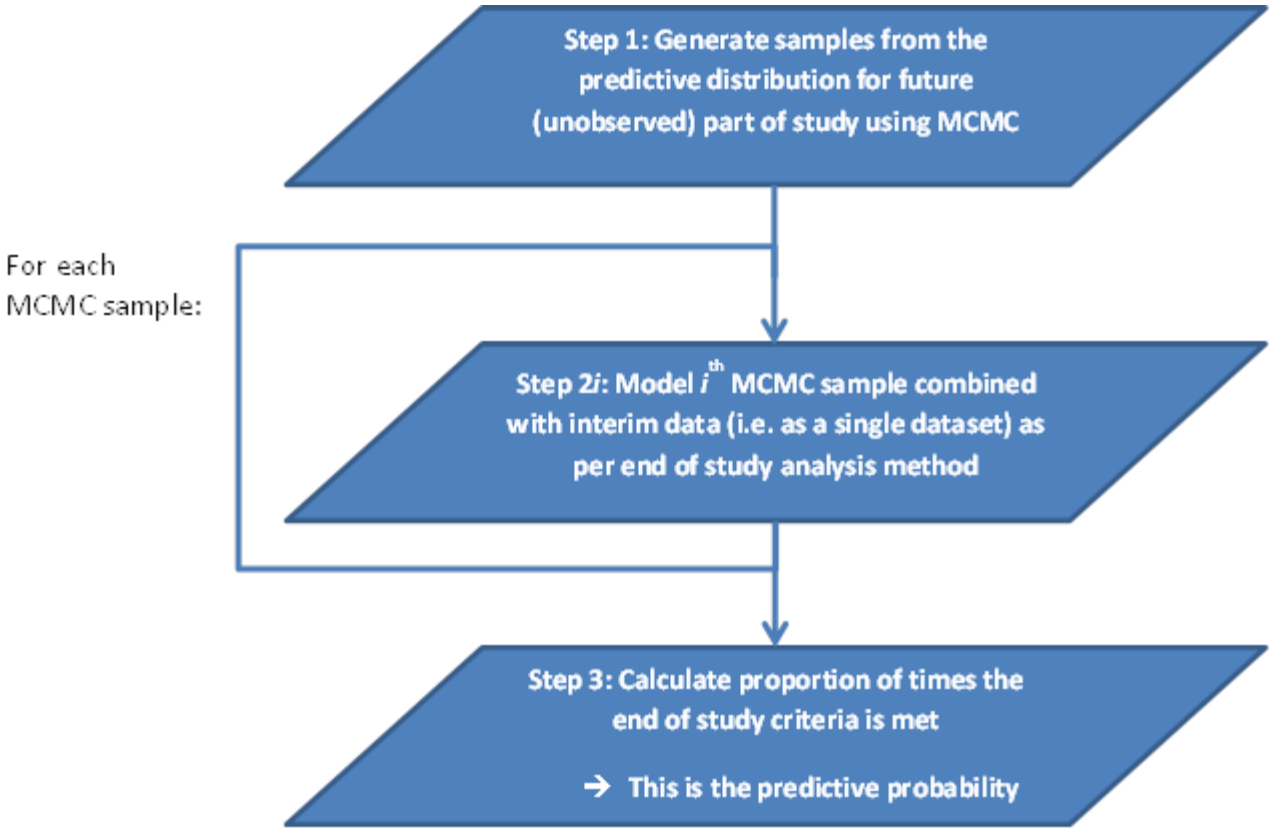
Bayesian Predictive Probability

- Many practical examples are more complex
  - Non-Normal endpoints
  - Unequal N in each arm
  - Change to design or rando ratio post-interim
  - Informative priors
    - Analytical formula can be extended to incorporate prior through “effective N” (see Walley et al 2015 for derivation)
    - Complex if using robust prior method (e.g. weighted mixture of vague/informative)
  - Multiple decision criteria
  - Analysis method
    - E.g. how to deal with MMRM analysis where at interim stage some subjects complete but others may only have partial data.
- Solutions:
  - Analytical formula but with assumptions/approximations, e.g. ratio of information at interim stage relative to end of study
  - Simulation based approach

Walley RJ, Smith CL, Gale JD & Woodward P (2015) Advantages of a wholly Bayesian approach to assessing efficacy in early drug development: a case study: Pharm Stat, 14(3) 205-15.

Bayesian Predictive Probability

- Simulation Approach

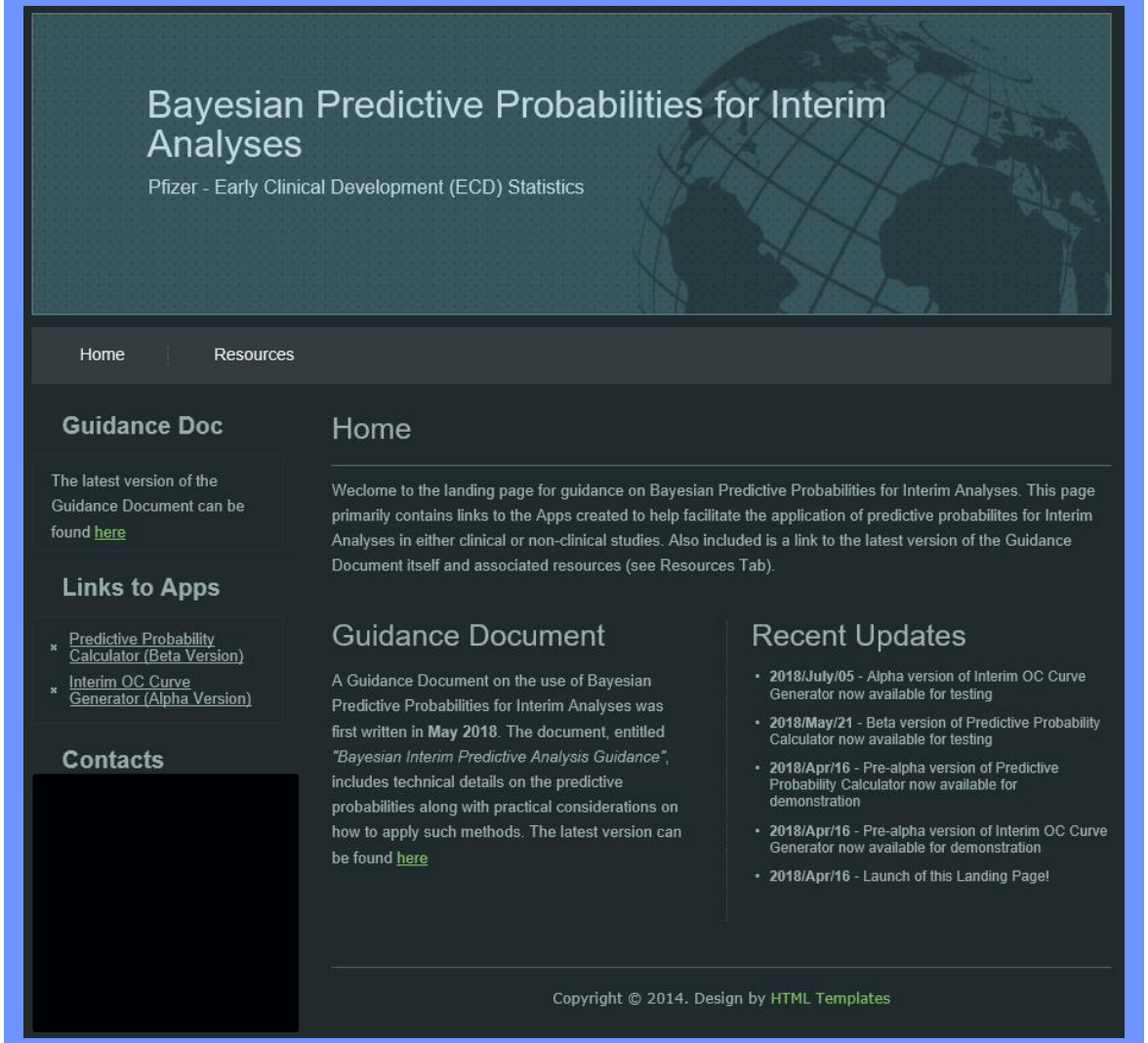


Guidance Document

- Complexity of approach means guidance and software required to aid implementation, reduce time burden and errors
- Team led by Donal Gorman and Yao Zhang developed “how-to” guidance document covering:
  - Introduction (Scope / Benefits of approach)
  - Methodology
    - Approaches and formulas covering simple no prior approach, incorporation of prior, simulation based approach
  - Design considerations
    - Timing of interim (OCs and recommendations)
    - Recommended decision criteria (e.g. for futility)
  - Operational considerations
    - Planning / Programming documentation
    - Best practices – e.g. interim charter, independent interim analysis team
  - Examples (with R-code)
  - Link to internal website and R-SHINY Apps

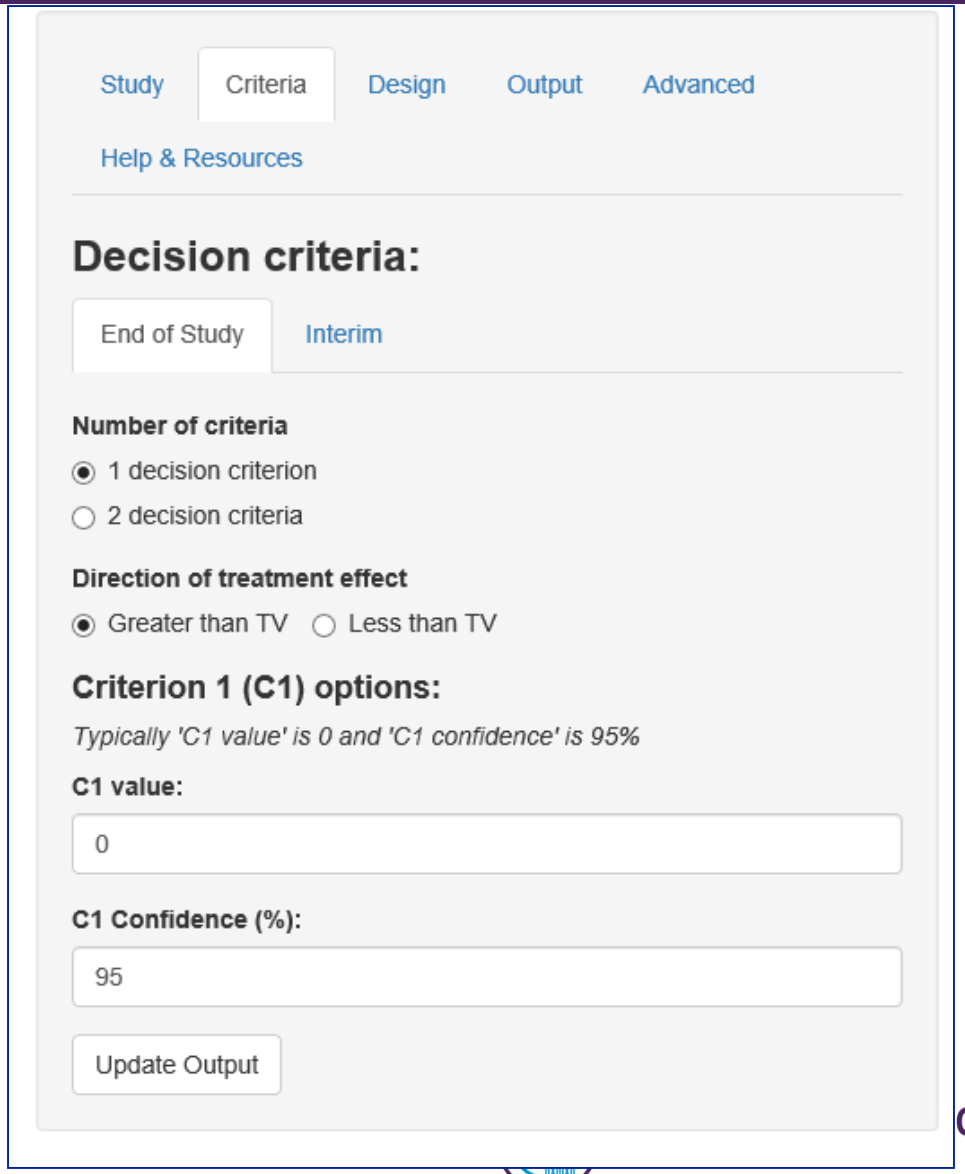
Website for the Apps

- Includes:
  - Latest version of the guidance document
  - Links to two R-SHINY Apps
    - Interim OC generator
    - Predictive probability calculator
  - PDF links of the key references in the document (e.g. other interim guidance)



App #1: Interim OC Curve Generator

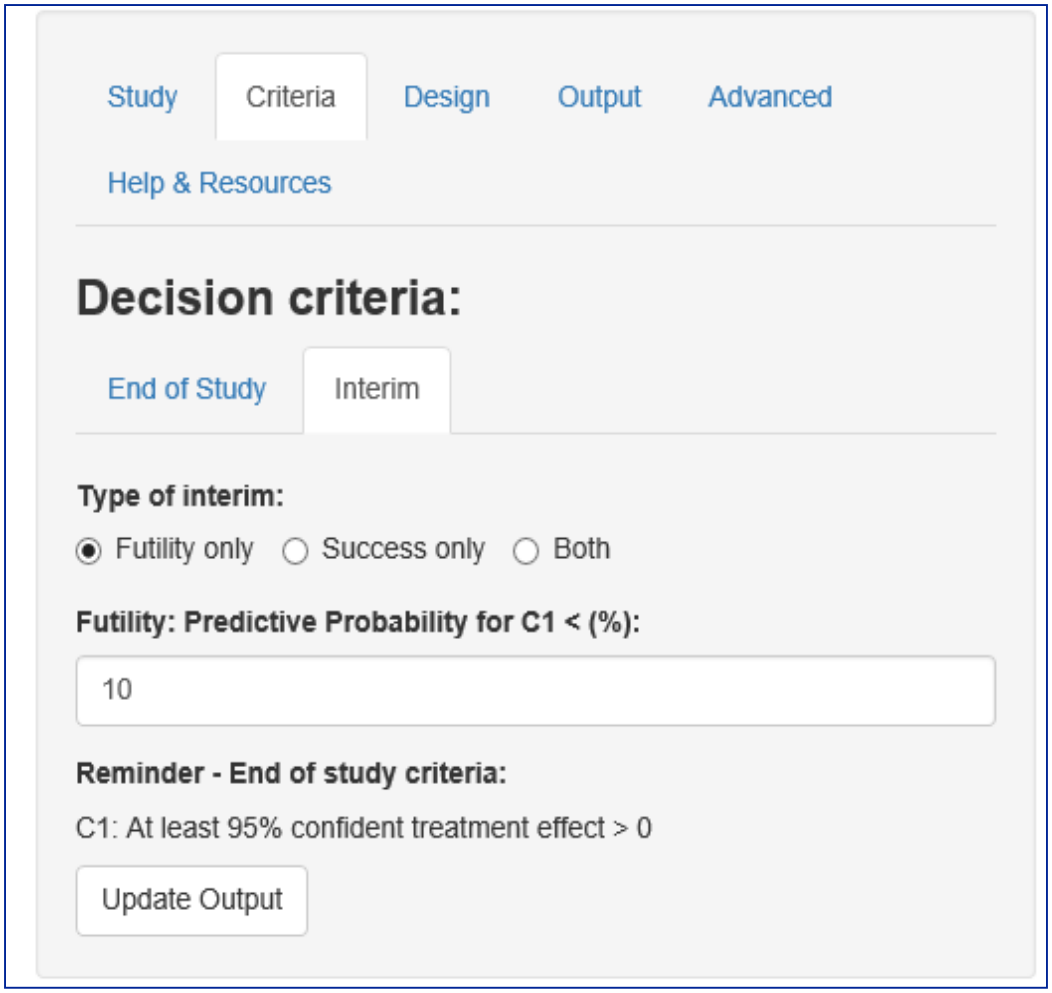
- This App generates OC curves for interim analyses so that value and timing can be assessed.
- Key inputs are:
  - End of study decision criteria (e.g. C1: 95% confident effect > 0)



App #1: Interim OC Curve Generator

- This App generates OC curves for interim analyses so that value and timing can be assessed.

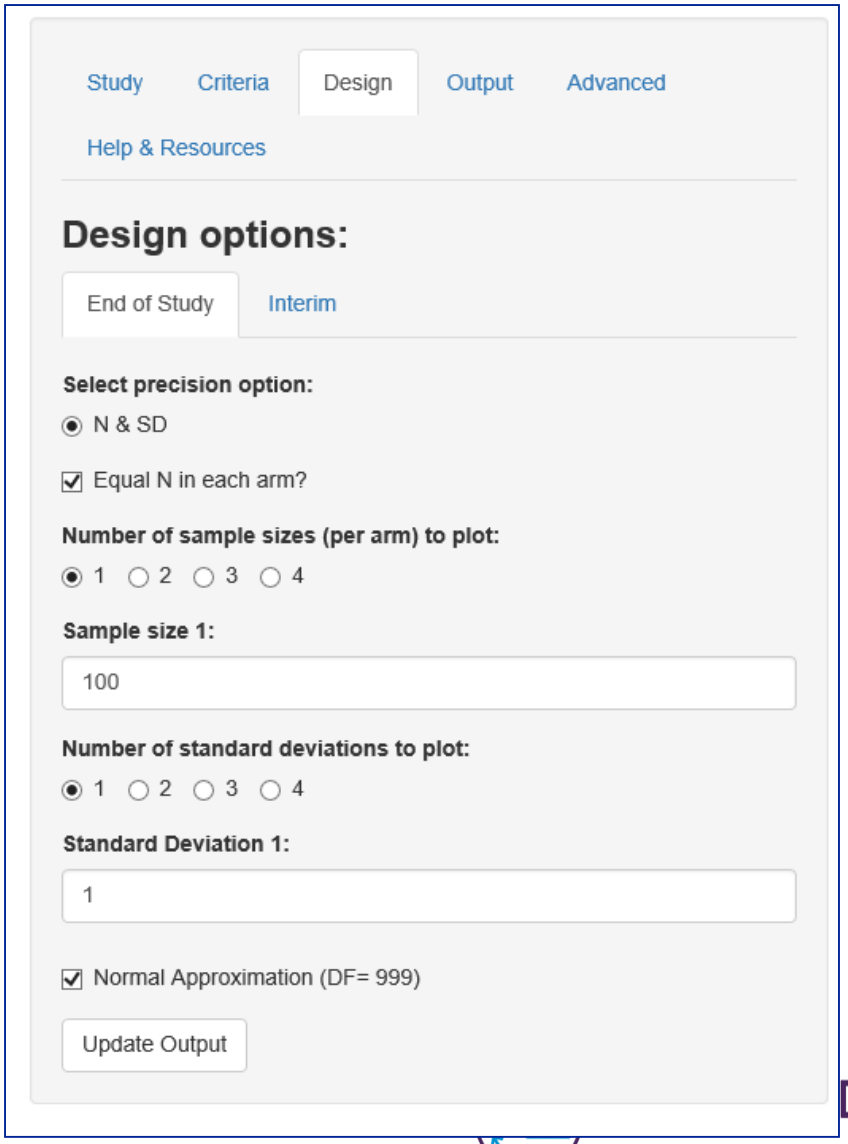
- Key inputs are:
  - End of study decision criteria (e.g. C1: 95% confident effect > 0)
  - Interim decision criteria (e.g. STOP for futility if predictive probability < 10%)



App #1: Interim OC Curve Generator

- This App generates OC curves for interim analyses so that value and timing can be assessed.

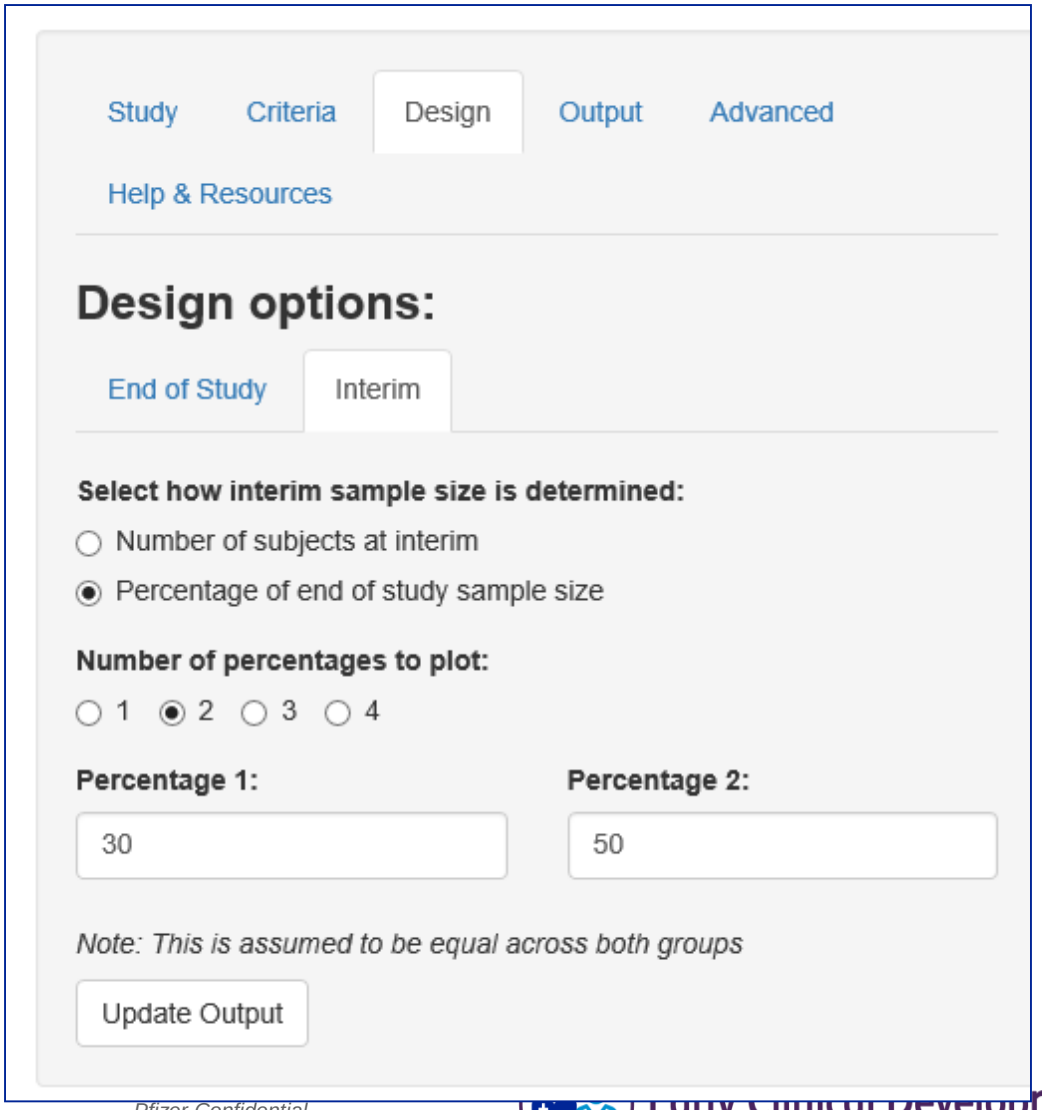
- Key inputs are:
  - End of study decision criteria (e.g. C1: 95% confident effect > 0)
  - Interim decision criteria (e.g. STOP for futility if predictive probability < 10%)
  - End of study sample size and SD estimate



App #1: Interim OC Curve Generator

- This App generates OC curves for interim analyses so that value and timing can be assessed.

- Key inputs are:
  - End of study decision criteria (e.g. C1: 95% confident effect > 0)
  - Interim decision criteria (e.g. STOP for futility if predictive probability < 10%)
  - End of study sample size and SD estimate
  - Interim timing



App #1: Interim OC Curve Generator

Study | Criteria | Design | Output | Advanced

Help & Resources

Design options:

End of Study | Interim

Select how interim sample size is determined:

- ○ Number of subjects at interim
- ◉ Percentage of end of study sample size

Number of percentages to plot:

○ 1 ◉ 2 ○ 3 ○ 4

Percentage 1: 30

Percentage 2: 50

Note: This is assumed to be equal across both groups

Update Output

Interim OC | Summary Table | Data Downloads

End of study decision criteria:

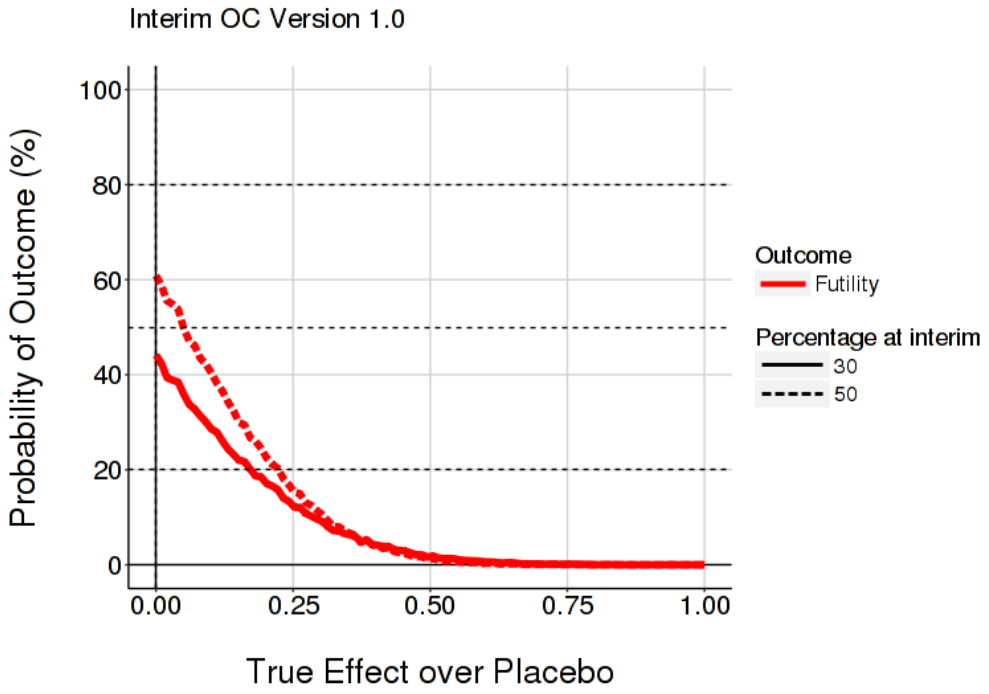
C1: At least 95% confident treatment effect > 0

Interim criteria:

Futility: Predictive probability for C1 < 10%

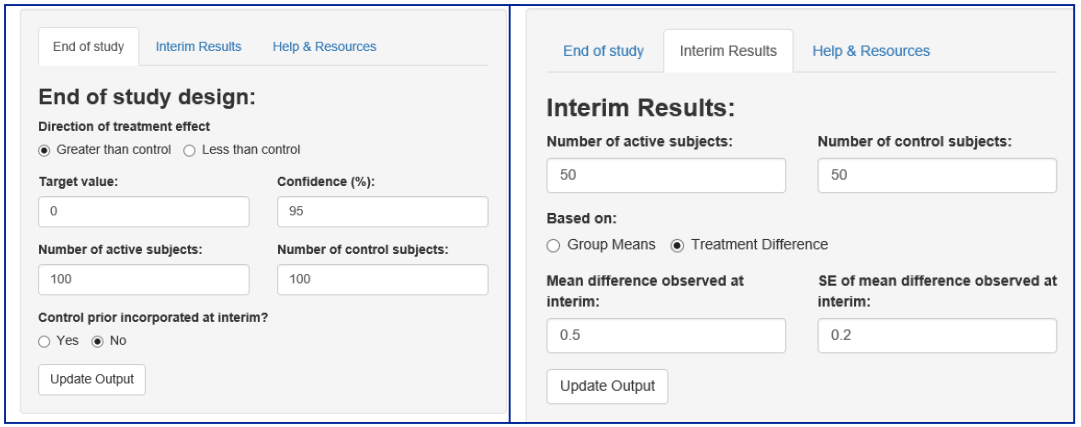
Sample sizes:

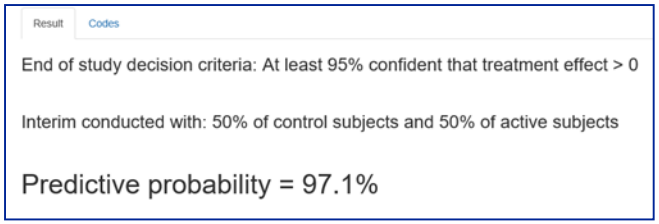
End of study: N (treatment) = 100; N (control) = 100
For 30% subjects at the interim: N (treat) = 30; N (control) = 30
For 50% subjects at the interim: N (treat) = 50; N (control) = 50



App #2: Predictive Probability Calculator

- This app calculates the predictive probability based on interim data:
- Inputs:
  - End of study criteria (e.g. C1: 95% confident effect > 0) and sample size.
  - Number of subjects at interim analysis and analysis results.
- Output:
  - Predictive probability of meeting end of study criteria
- R/SAS code is generated to re-produce the result that could also be used as a template for interim SAP/programming plan
  - Formula used is based on user inputs (e.g. whether a placebo prior is included or not)

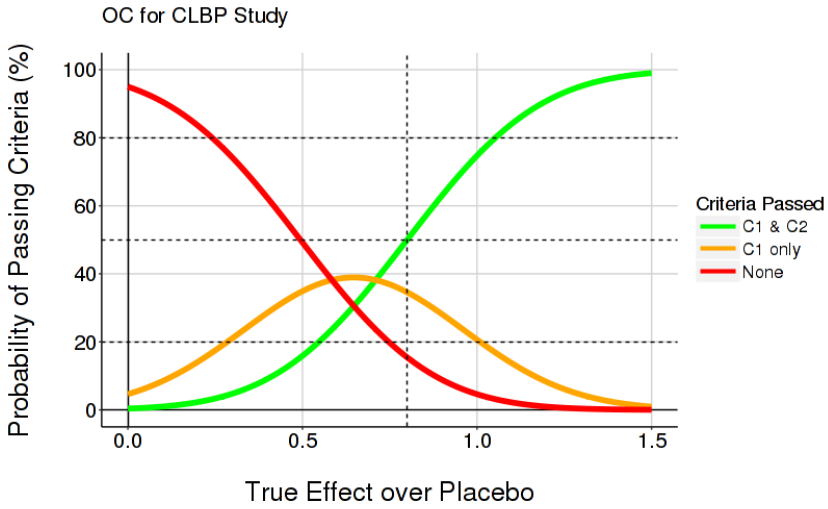





```
data;
/* SAS Code to calculate predictive probability from 'Interim Predictive Probability Calculator' */

*Version: 1.01;
*Created: 14-May-2019;

/*Inputs:*/;
i_e_dir = 1;
i_e_tv = 0;
i_e_sig = 0.95;
i_e_nAtot = 100;
i_e_nPtot = 100;
i_r_nAint = 50;
i_r_nPint = 50;
i_r_meanD = 0.5;
i_r_seD = 0.2;

/* Taken from equations 1 & 9-11 from 'Bayesian Interim Predictive Analysis Guidance' document (v1.0 19 September 2018)*/;

/* Derived:*/;
i = i_r_nAint/i_e_nAtot;
if i_e_dir=1 /*Greater than*/ THEN
z = QUANTILE('NORMAL',i_e_sig);
if i_e_dir=2 /*Less than*/ THEN
z = QUANTILE('NORMAL',1-i_e_sig);

/*Calculate numerator of equation:*/;
postSD = sqrt(i)*i_r_seD;
num_eq = i_e_tv + z*postSD - i_r_meanD;

/*Calculate denominator of equation:*/;
den_eq = sqrt(1-i)*i_r_seD;

/*Determine predictive probability:*/;
eq_result = num_eq/den_eq;
if i_e_dir=1 /*Greater than*/ THEN
predProb = 1-CDF('NORMAL',eq_result);
if i_e_dir=2 /*Less than*/ THEN
predProb = CDF('NORMAL',eq_result);

run;
```

Example 2: CLBP Example with informative placebo prior

- Chronic Low Back Pain (CLBP) example
- 3 arm parallel group study (PF, placebo, naproxen)
- Primary endpoint, change from baseline to week 4 in the daily Low Back Pain Intensity (LBPI) as measured by an 11-point NRS.
- Primary analysis: ANCOVA with treatment as factor and baseline pain as covariate
- Decision Criteria
  - C1: At least 95% confident that PF effect is greater than placebo
  - C2: Observed effect is 0.8 units better than placebo
- An informative $N(-2.36, 0.54^2)$ prior used for the placebo effect based on the results of 9 internal studies
  - Given assumed SD of 2.2 the prior equates to an effective N of 16 placebo subjects
- Sample size of 100 subjects per arm (effectively 116 on placebo) gave acceptable operating characteristics.



Example 2: CLBP Example with informative placebo prior

- Interim Analysis planned
- Stop the study for futility if:
  - the predictive probability of meeting C2 is less than 20%.
- Accelerate development activities (blinded to study team) if:
  - the predictive probabilities of meeting both criteria C1 and C2 are greater than 80%

- Question: Timing of interim?
  - Want high chance of stopping if compound is ineffective
  - OC generator can be used to compare different timings (e.g. 33% v 50% v 66%)
    - Note speed of recruitment also considered

Example 2: CLBP Example with informative placebo prior

Decision criteria:

Decision criteria:

Design options:

Design options:

End of Study | Interim

Select how interim sample size is determined:

- (•) Number of subjects at interim
- ( ) Percentage of end of study sample size

Number of subjects at the interim (treatment arm) to plot:

(•) 1 ( ) 2 ( ) 3 ( ) 4

Interim Sample Size 1:

66

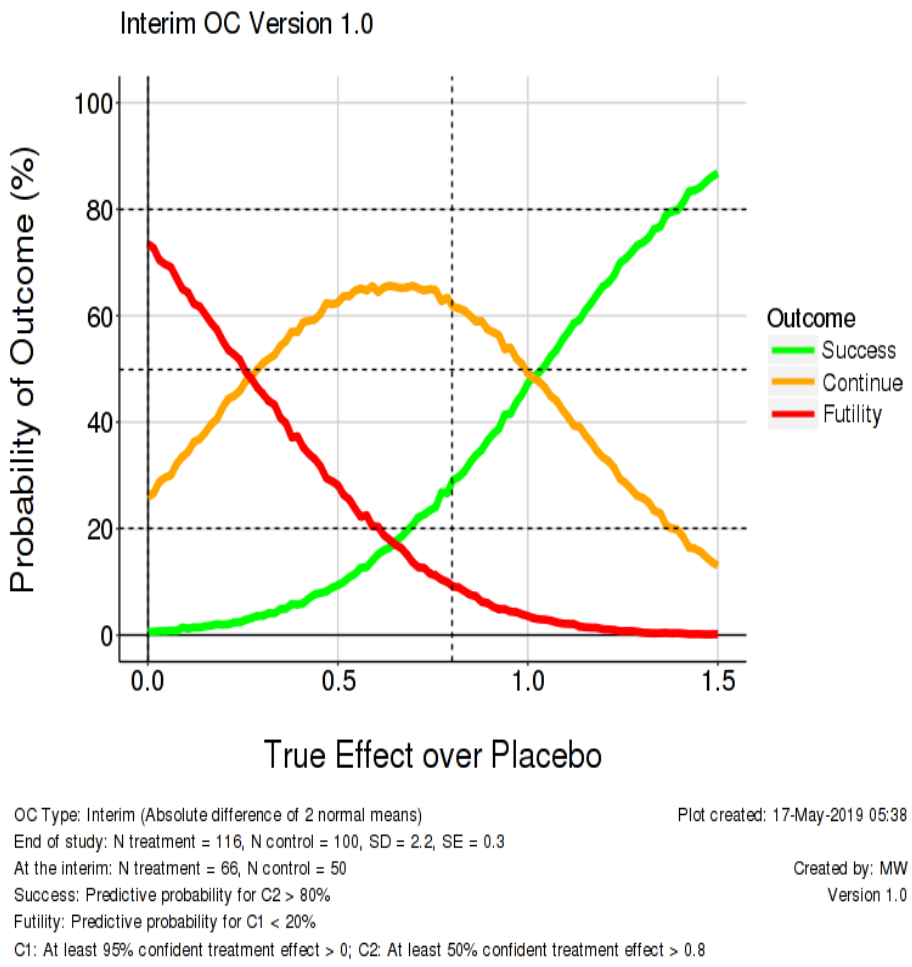
Number of subjects at the interim (control arm) to plot:

(•) 1 ( ) 2 ( ) 3 ( ) 4

Interim Sample Size 1:

50

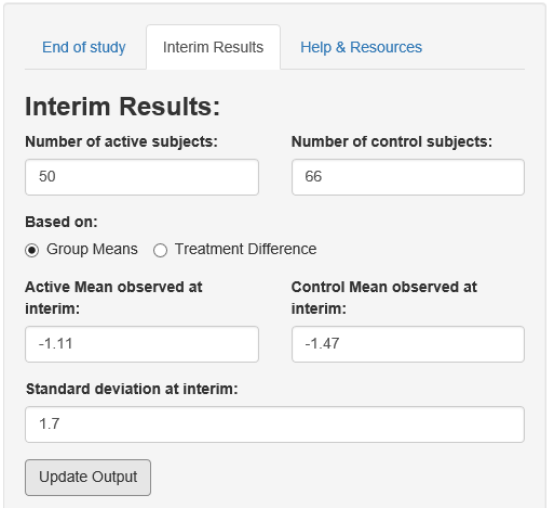
Update Output

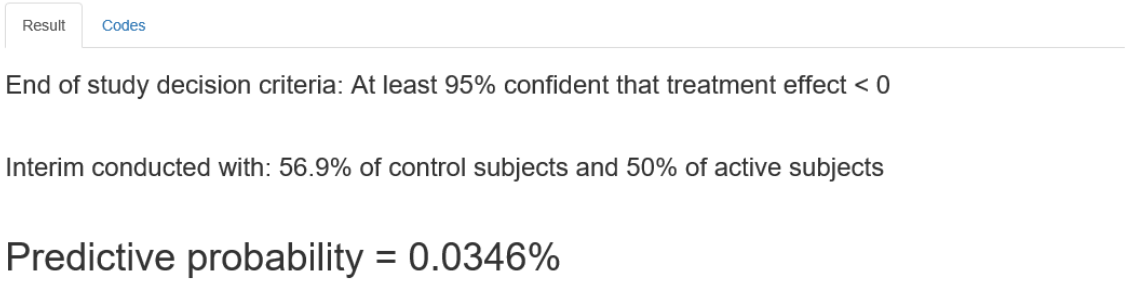


Example 2: CLBP Example with informative placebo prior

- Interim Analysis results → Study stopped for futility

	Primary Analysis		Predictive Probability	
	Mean	SE	C1	C2
Placebo	-1.47	0.22		
PF	-1.11	0.25		
PF-Placebo	0.33	0.33	0.035%	<0.001%





Summary

- Bayesian predictive probabilities are an intuitive approach to decision making for interim analyses in early stage studies.
- Implementation can be complex
- Internal guidance document developed to aid implementation.
- R-SHINY apps enable large time-savings in creation of operating characteristics and calculation of predictive probability.

Acknowledgements

- Donal Gorman
- Yao Zhang
- Satrajit Roychoudhury
- Phil Woodward
- Simon Kirby

Questions