

# 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**



WORLDWIDE RESEARCH & DEVELOPMENT



# 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,  $\delta$  is mean difference at study end.
- For the one-sided significance test of  $H_0: \delta=0$  vs  $H_1: \delta>0$  we will, at study end, reject  $H_0$  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  $H_0$  given by:
 
$$1 - \Phi\left(\frac{\frac{Z_\alpha(n_1 + n_2)\sqrt{\frac{\sigma_D^2}{(n_1 + n_2)}} - n_1\bar{d}_1}{n_2} - \bar{d}_1}{\sqrt{\frac{\sigma_D^2}{n_1} + \frac{\sigma_D^2}{n_2}}}\right)$$



# 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

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



# Bayesian Predictive Probability

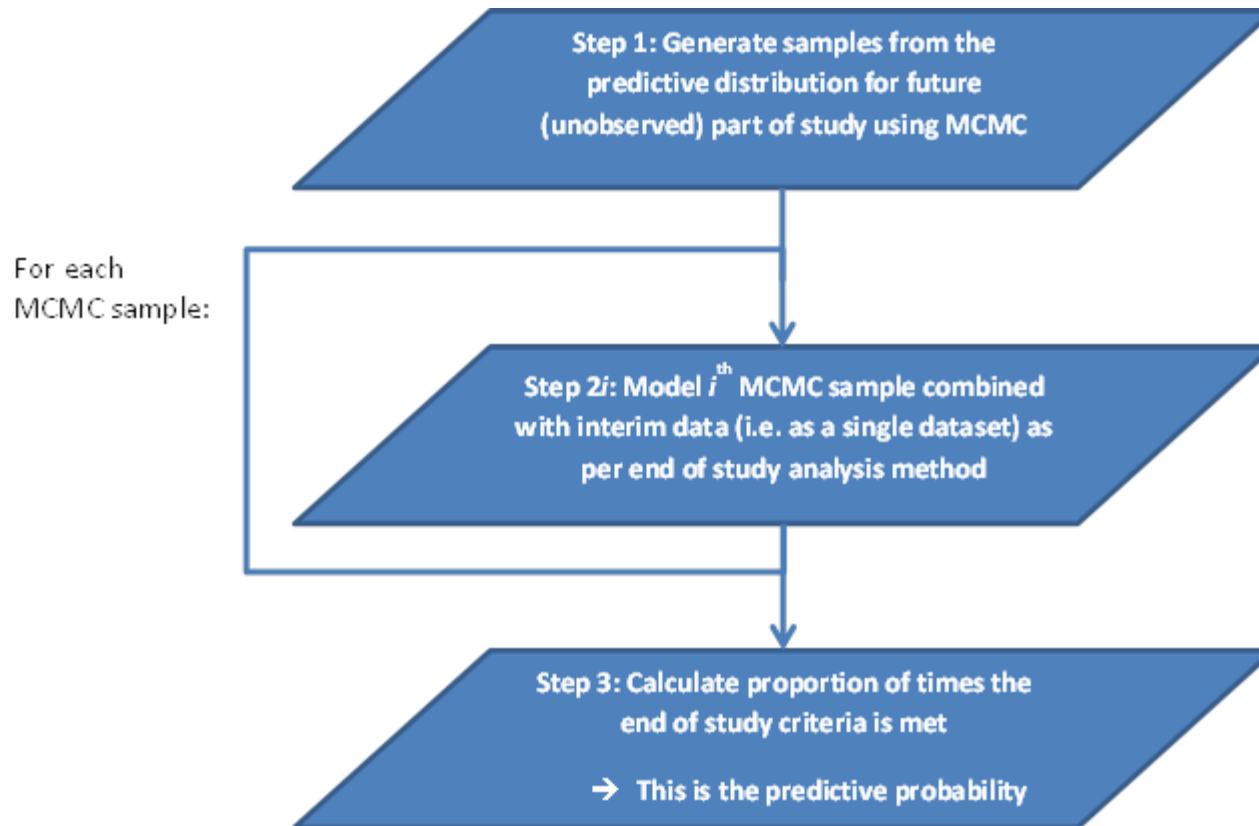
- Many practical examples are more complex
  - Non-Normal endpoints
  - Unequal N in each arm
  - Change to design or random ratio post-interim
  - Informative priors
    - Analytical formula can be extended to incorporate prior through “effective N” (see Walley et al 2015 for derivation)
    - Complex 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





# 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)

**Bayesian Predictive Probabilities for Interim Analyses**  
Pfizer - Early Clinical Development (ECD) Statistics

Home | Resources

**Guidance Doc**

The latest version of the Guidance Document can be found [here](#)

**Links to Apps**

- \* [Predictive Probability Calculator \(Beta Version\)](#)
- \* [Interim OC Curve Generator \(Alpha Version\)](#)

**Contacts**

**Home**

Welcome to the landing page for guidance on Bayesian Predictive Probabilities for Interim Analyses. This page primarily contains links to the Apps created to help facilitate the application of predictive probabilities for Interim Analyses in either clinical or non-clinical studies. Also included is a link to the latest version of the Guidance Document itself and associated resources (see Resources Tab).

**Guidance Document**

A Guidance Document on the use of Bayesian Predictive Probabilities for Interim Analyses was first written in May 2018. The document, entitled "Bayesian Interim Predictive Analysis Guidance", includes technical details on the predictive probabilities along with practical considerations on how to apply such methods. The latest version can be found [here](#)

**Recent Updates**

- 2018/July/05 - Alpha version of Interim OC Curve Generator now available for testing
- 2018/May/21 - Beta version of Predictive Probability Calculator now available for testing
- 2018/Apr/16 - Pre-alpha version of Predictive Probability Calculator now available for demonstration
- 2018/Apr/16 - Pre-alpha version of Interim OC Curve Generator now available for demonstration
- 2018/Apr/16 - Launch of this Landing Page!

Copyright © 2014. Design by [HTML Templates](#)



# 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)

Study   **Criteria**   Design   Output   Advanced

[Help & Resources](#)

---

### Decision criteria:

End of Study   **Interim**

**Number of criteria**

1 decision criterion  
 2 decision criteria

**Direction of treatment effect**

Greater than TV    Less than TV

**Criterion 1 (C1) options:**

*Typically 'C1 value' is 0 and 'C1 confidence' is 95%*

**C1 value:**

  
**C1 Confidence (%):**

# App #1: Interim OC Curve Generator

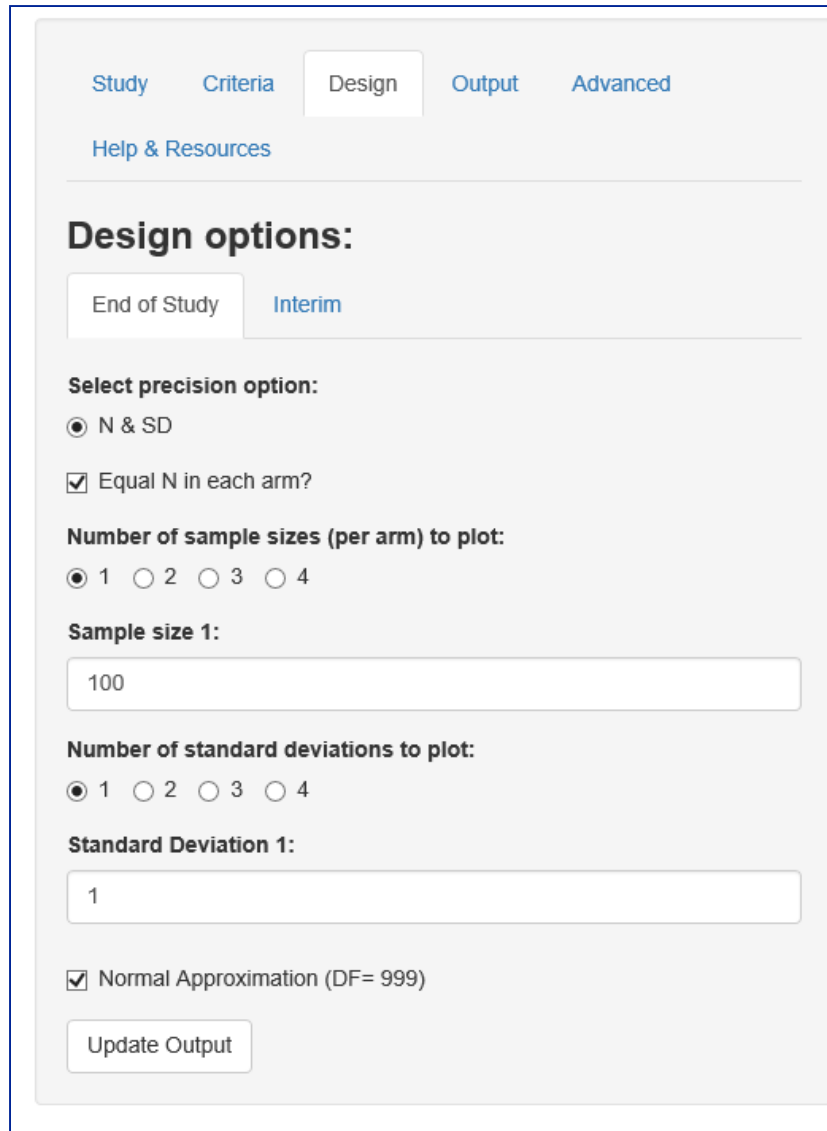
- This App generates OC curves for interim analyses so that value and timing can be assessed.
- Key inputs are:
  - End of study decision criteria (e.g. C1: 95% confident effect > 0)
  - **Interim decision criteria (e.g. STOP for futility if predictive probability < 10%)**

The screenshot shows the 'Criteria' tab of the 'Interim OC Curve Generator' app. The interface includes a navigation bar with 'Study', 'Criteria', 'Design', 'Output', and 'Advanced' tabs. Below the navigation bar is a 'Help & Resources' link. The main section is titled 'Decision criteria:' and has two sub-tabs: 'End of Study' and 'Interim'. Under the 'Interim' tab, there is a 'Type of interim:' section with three radio buttons: 'Futility only' (selected), 'Success only', and 'Both'. Below this is a 'Futility: Predictive Probability for C1 < (%):' section with a text input field containing '10'. At the bottom, there is a 'Reminder - End of study criteria:' section with the text 'C1: At least 95% confident treatment effect > 0' and an 'Update Output' button.



# 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



The screenshot displays the 'Design' tab of the Interim OC Curve Generator app. The interface includes a navigation bar with tabs for 'Study', 'Criteria', 'Design', 'Output', and 'Advanced'. Below the navigation bar is a 'Help & Resources' link. The main section is titled 'Design options:' and contains several settings:

- End of Study / Interim:** Two radio buttons, with 'Interim' selected.
- Select precision option:** Radio buttons for 'N & SD' (selected) and 'Equal N in each arm?' (checked).
- Number of sample sizes (per arm) to plot:** Radio buttons for 1 (selected), 2, 3, and 4.
- Sample size 1:** A text input field containing the value '100'.
- Number of standard deviations to plot:** Radio buttons for 1 (selected), 2, 3, and 4.
- Standard Deviation 1:** A text input field containing the value '1'.
- Normal Approximation (DF= 999):** A checked checkbox.
- Update Output:** A button at the bottom of the form.



# 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**

The screenshot shows the 'Design' tab of the Interim OC Curve Generator app. The interface includes a navigation bar with 'Study', 'Criteria', 'Design', 'Output', and 'Advanced' tabs. Below the navigation bar is a 'Help & Resources' link. The main section is titled 'Design options:' and contains two sub-tabs: 'End of Study' and 'Interim'. Under the 'Interim' tab, there are three sections: 'Select how interim sample size is determined:' with radio buttons for 'Number of subjects at interim' and 'Percentage of end of study sample size' (selected); 'Number of percentages to plot:' with radio buttons for 1, 2 (selected), 3, and 4; and two input fields for 'Percentage 1:' (30) and 'Percentage 2:' (50). A note states 'Note: This is assumed to be equal across both groups'. At the bottom is an 'Update Output' button.



# 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:**    **Percentage 2:**

*Note: This is assumed to be equal across both groups*

Interim OC

Summary Table

Data Downloads

End of study decision criteria:

C1: At least 95% confident treatment effect  $> 0$

Interim criteria:

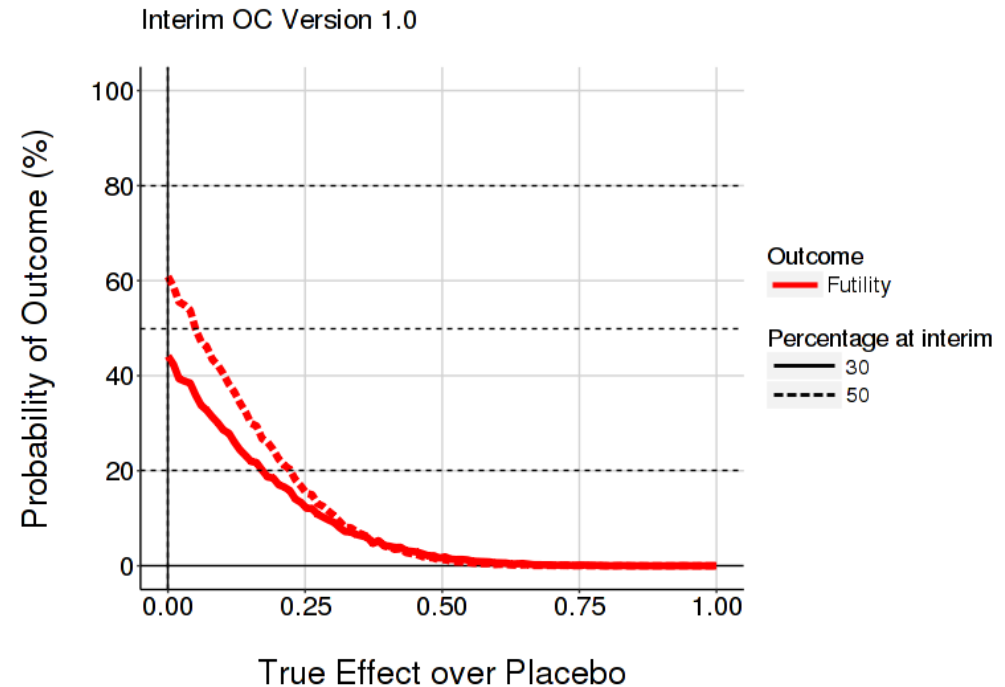
Futility: Predictive probability for C1  $< 10\%$

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 reproduce 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)

End of study
Interim Results
Help & Resources

---

**End of study design:**

Direction of treatment effect  
 Greater than control    Less than control

Target value:       Confidence (%):

Number of active subjects:       Number of control subjects:

Control prior incorporated at interim?  
 Yes    No

End of study
Interim Results
Help & Resources

Result
Codes

---

End of study decision criteria: At least 95% confident that treatment effect > 0

Interim conducted with: 50% of control subjects and 50% of active subjects

Predictive probability = 97.1%

```
data;
/* SAS Code to calculate predictive probability from 'Interim Predictive Probability Calculator' */
*Version: 1.01;
*Created: 14-May-2019;

/*Inputs*/;
L_e_dir = 1;
L_e_lv = 0;
L_e_sig = 0.95;
L_e_nAtot = 100;
L_e_nInt = 100;
L_r_nAint = 50;
L_r_nPint = 50;
L_r_meanD = 0.5;
L_r_seD = 0.2;

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

/* Derived*/;
i = L_r_nAint/L_e_nAtot;
if L_e_dir=1 /*Greater than*/ THEN
z = QUANTILE(NORMAL,L_e_sig);
if L_e_dir=2 /*Less than*/ THEN
z = QUANTILE(NORMAL,1-L_e_sig);

/*Calculate numerator of equation*/;
postSD = sqrt(i)*L_r_seD;
num_eq = L_e_lv + z*postSD - L_r_meanD;

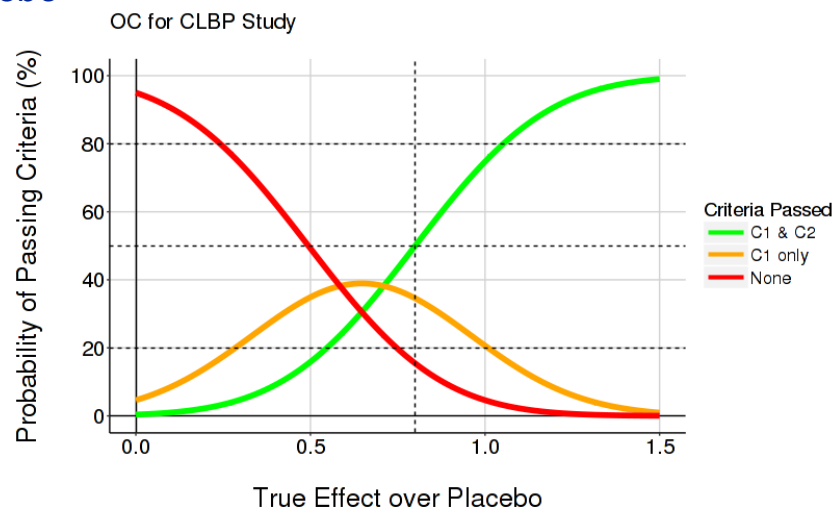
/*Calculate denominator of equation*/;
den_eq = sqrt(1-i)*L_r_seD;

/*Determine predictive probability*/;
eq_result = num_eq/den_eq;
if L_e_dir=1 /*Greater than*/ THEN
predProb = 1-CDF(NORMAL,eq_result);
if L_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

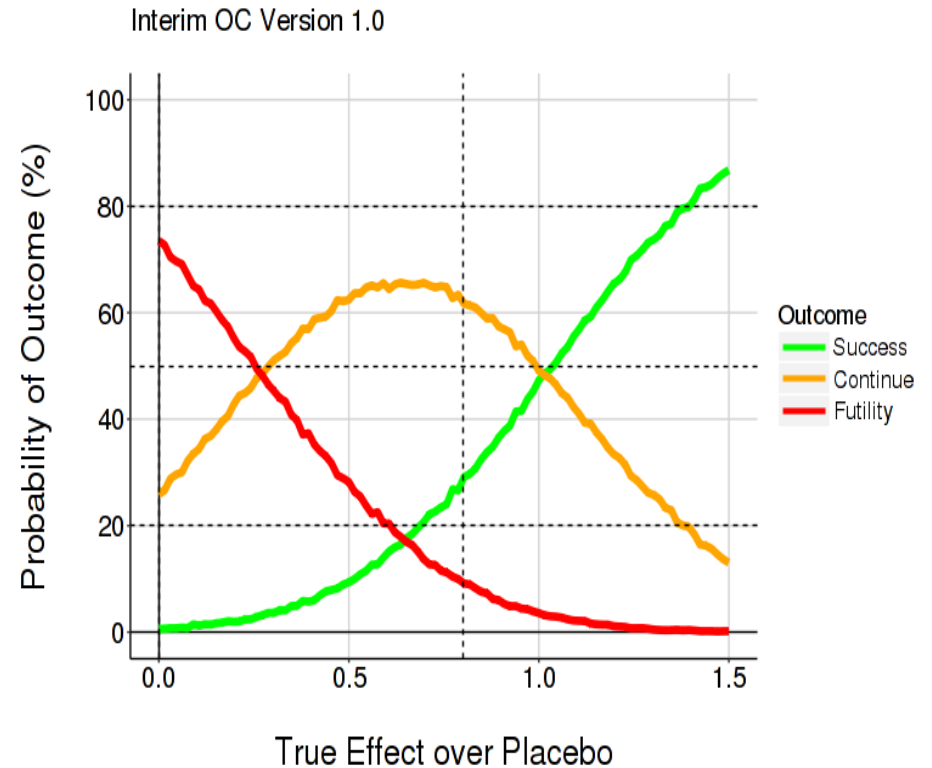
Select how interim sample size is determined:  
 Number of subjects at interim  
 Percentage of end of study sample size

Number of subjects at the interim (treatment arm) to plot:  
 1  2  3  4

Interim Sample Size 1:

Number of subjects at the interim (control arm) to plot:  
 1  2  3  4

Interim Sample Size 1:



OC Type: Interim (Absolute difference of 2 normal means)  
 End of study: N treatment = 116, N control = 100, SD = 2.2, SE = 0.3  
 At the interim: N treatment = 66, N control = 50  
 Success: Predictive probability for C2 > 80%  
 Futility: Predictive probability for C1 < 20%  
 C1: At least 95% confident treatment effect > 0; C2: At least 50% confident treatment effect > 0.8

Plot created: 17-May-2019 05:38

Created by: MW  
 Version 1.0



# 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%

End of study | Interim Results | Help & Resources

**Interim Results:**

Number of active subjects:  Number of control subjects:

Based on:  
 Group Means  Treatment Difference

Active Mean observed at interim:  Control Mean observed at interim:

Standard deviation at interim:

Result Codes

End of study decision criteria: At least 95% confident that treatment effect < 0

Interim conducted with: 56.9% of control subjects and 50% of active subjects

Predictive probability = 0.0346%

# 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

