```
cpp11   (0.4.7 -> 0.5.0) [CRAN]
httr2   (1.0.1 -> 1.0.3) [CRAN]
pkgdown (2.0.9 -> 2.1.0) [CRAN]

The downloaded binary packages are in
    /var/folders/7s/hgrbfd4152gfx6nyk71l8fpw0000gn/T//RtmprDNDv9/downloaded_packages
── R CMD build ─────────────────────────────────────────────────────────────────
* checking for file
‘/private/var/folders/7s/hgrbfd4152gfx6nyk71l8fpw0000gn/T/RtmprDNDv9/remotes1ada552a0eb3/Genentec
phase1b-c9c1440/DESCRIPTION’ ... OK
* preparing ‘phase1b’:
* checking DESCRIPTION meta-information ... OK
* checking for LF line-endings in source and make files and shell scripts
* checking for empty or unneeded directories
* building ‘phase1b_1.0.0.tar.gz’
```

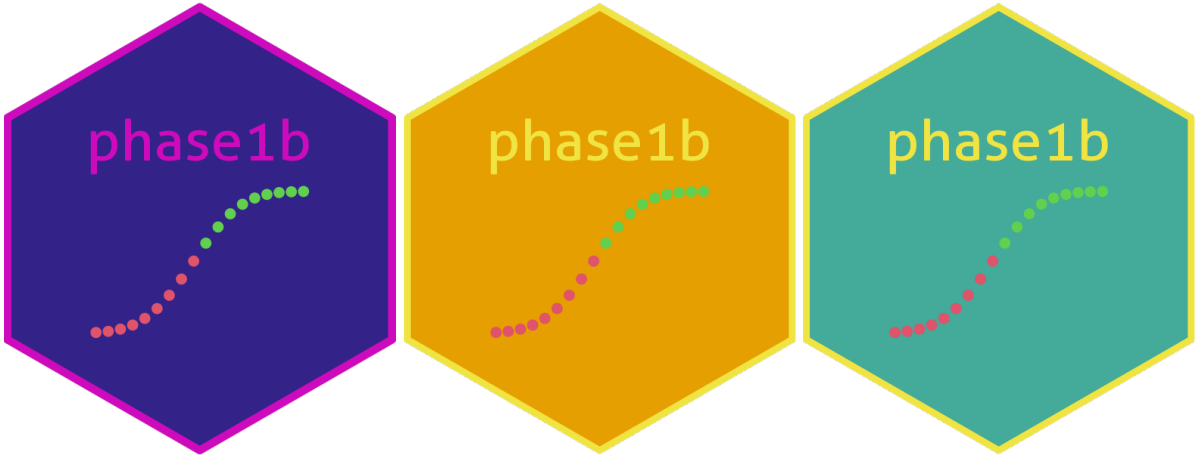
A Bayesian approach to decision making in early development clinical trials: An R solution

Audrey Yeo

Welwyn Statistics CoP, 2024

This presentation has ALT text and as much as possible, uses colour-blind friendly palettes

Early oncology trials and why `phase1b`?



Prior and Posterior of Beta Distribution for response rate

- Conjugate Prior is $f(\pi)$, where $\pi \sim Beta(\alpha, \beta)$, same family of distribution of Posterior (see below)
- We know the mean response rate (RR) is :

$$\pi = \frac{\alpha}{\alpha + \beta}$$

- Likelihood is $f(x|\pi)$, where $x \sim Binomial(x, n)$
- The updated Posterior $f(\pi|x)$ is again a $Beta$ distribution (same family as prior) :

$$\pi|\ x \sim Beta(\alpha + x,\ \beta + n - x)$$

where x is the number of responders of current trials

History and how to install :

- 2015 : Started as a need in Roche's early development group, package development led by Daniel Sabanés Bové in 2015.
- 2023 : Refactoring, Renaming, adding Unit and Integration tests as current State-of-Art Software Engineering practice.
- 100% written in R and Open Source.
- website : genentech.github.io/phase1b/

```
library(devtools)
devtools::install_github("https://github.com/Genentech/phase1b")
library(phase1b)
```

Use case:

A single arm novel therapeutic with an assumed control response rate is at most 60%

Example	Interim	Final
Responders	16	23
n	23	40
Response rate	69.57 %	57.5 %
Posterior probability*	ask phase1b	ask phase1b
Predictive posterior probability*	ask phase1b	-
Decision to develop molecule further : Go/Stop/Grey Zone	ask phase1b	ask phase1b

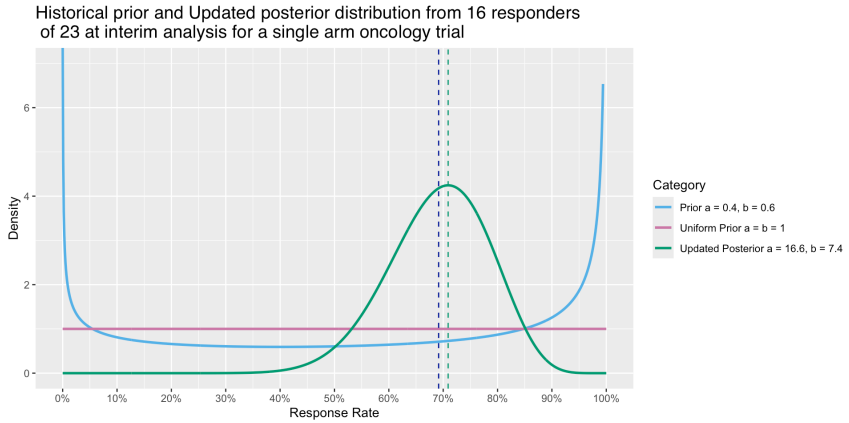
* Posterior Probability : $P(\pi > 60\% | \alpha + x, \beta + n - x)$
* Predictive Posterior Probability : $P(success\ or\ failure\ at\ final)$

Updating the Posterior

- Using the formula for the mean, where $\alpha = 0.6, \beta = 0.4$ and at interim x = 16, n = 23 :

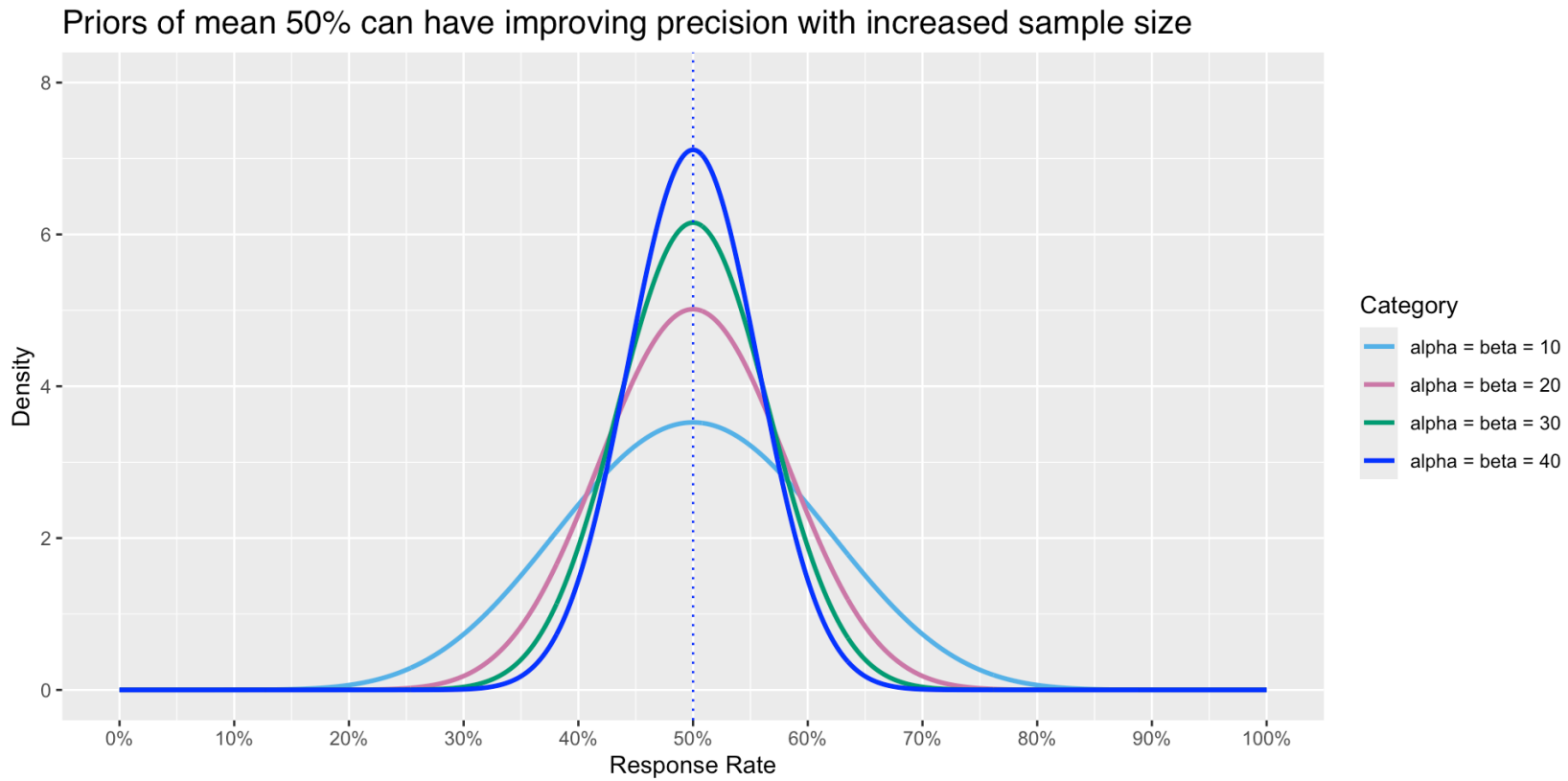
$$\pi = \frac{\alpha}{\alpha + \beta} = \frac{\alpha_{updated}}{\alpha_{updated} + \beta_{updated}} = \frac{16.6}{16.6 + 7.4} \approx 69.17\%$$

$$mode(\pi) = \frac{\alpha_{updated} - 1}{\alpha_{updated} + \beta_{updated} - 2} = \frac{16.6 - 1}{16.6 + 7.4 - 2} \approx 70.90\%$$



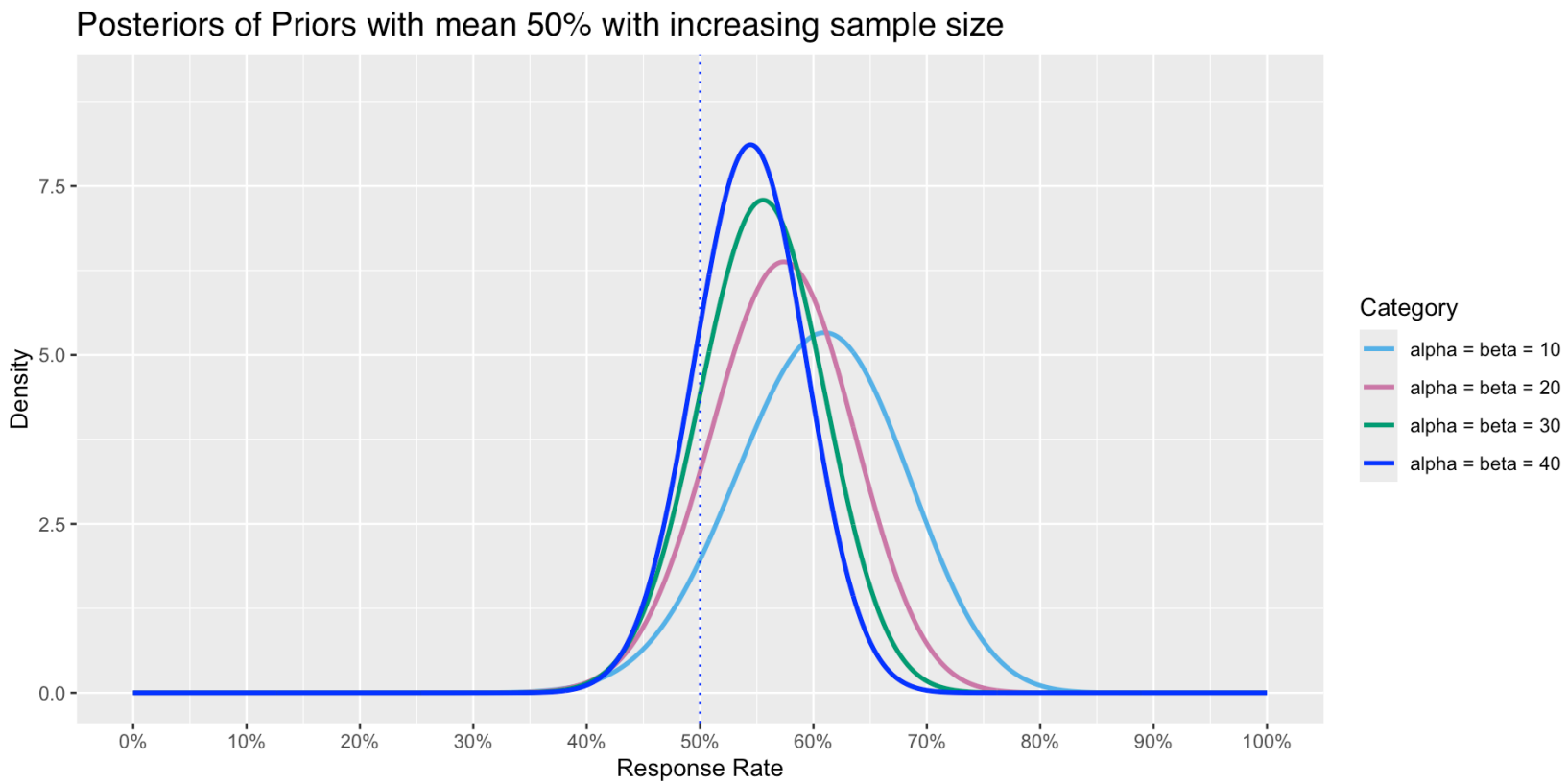
A variety of Priors

- To illustrate how density of Prior changes with increased sample size even though mean is the same



A variety of Posteriors

- To illustrate how density of Posterior changes with increased sample size even though mean is the same



Terminology

- A “look” is a stop
- A stop is when a rule is applied
- The rule is specified for Go, Stop or Evaluate (Gray zone)
- If the rule is met, the result is Go, Stop or Evaluate (Gray zone)
- Rules are applied at interim or final
- Go = Success = Efficacious
- Stop = Failure = Futile
- Rule = Criteria, e.g. Go rule is a Go Criteria

`postprob()` example (Lee & Liu, 2008)

Example	Interim
Responders	16
n	23
Response rate	69.57 %
Standard of Care Response rate	60 %
Posterior probability	postprob( ) call from phase1b

```
1 postprob(x = 16, n = 23, p = 0.60, par = c(0.6, 0.4))
```

```
[1] 0.8359808
```

Posterior Probability

- Interim trial is efficacious if posterior probability exceeds 70% or P(RR ≥ 60 % | data) > 70%

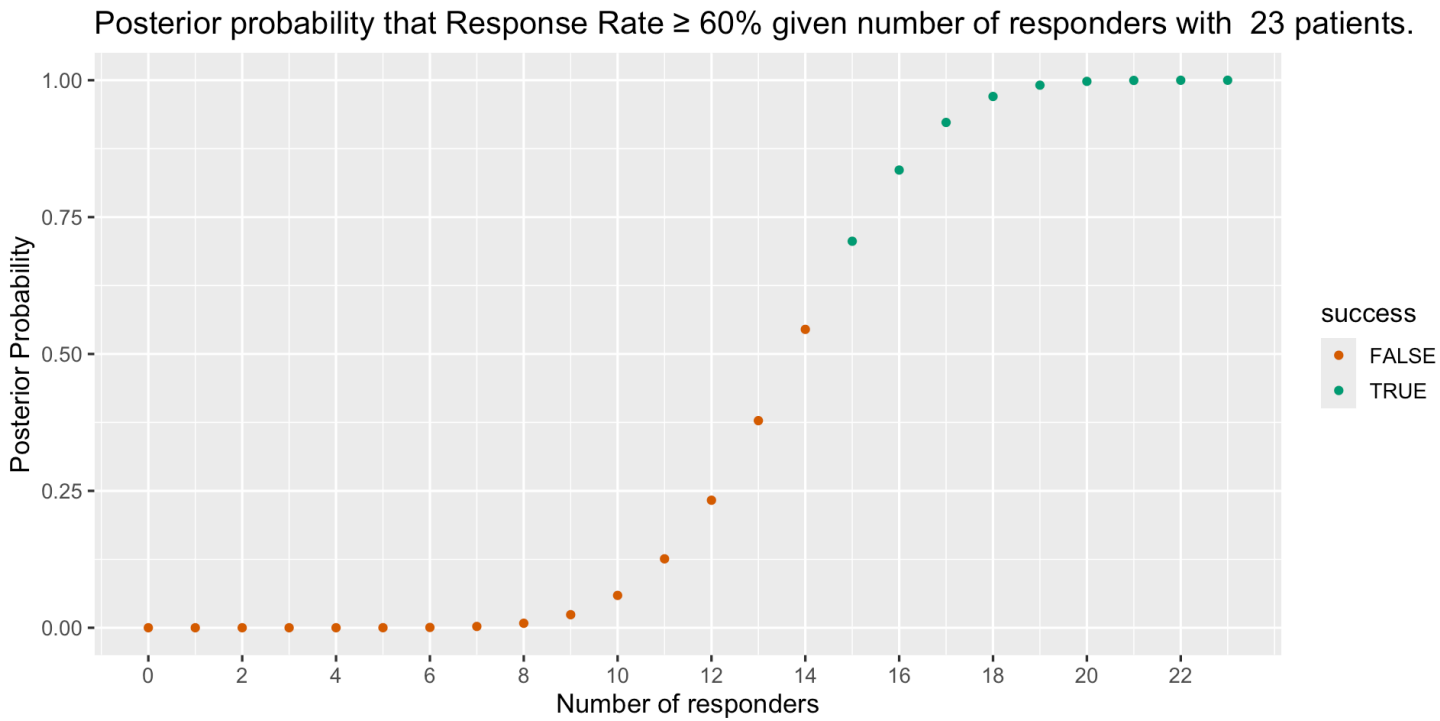


Figure 1

Beta prior mixture

- `phase1b` facilitates the flexibility of usingvarious priors and its respective weightings:

```
Prior is P_E ~ sum(weights * beta(parE[, 1], parE[, 2]))
```

```r
a = postprob(x = 16,
             n = 23,
             p = 0.60,
             parE = c(0.6, 0.4), weights = 1)

b = postprob(x = 16,
             n = 23,
             p = 0.60,
             parE = c(2, 4), weights = 1)

0.5*(a + b)

postprob(x = 16,
             n = 23,
             p = 0.60,
             parE = rbind(c(0.6, 0.4), c(2, 4)), weights = c(0.5, 0.5))
```

- Posterior formulation :

$$f(\pi|x) \propto \pi^{x}(1-\pi)^{n-x} \sum_{j=1}^{k} w_j \frac{1}{B(\alpha_j, \beta_j)} \pi^{\alpha_j - 1}(1-\pi)^{\beta_j - 1}$$

`predprob()` example (Lee & Liu, 2008)

Example	Interim
Responders	16
n	23
Response rate	69.57 %
Standard of Care Response rate	60 %
Predictive Posterior probability	predprob( ) call from phase1b

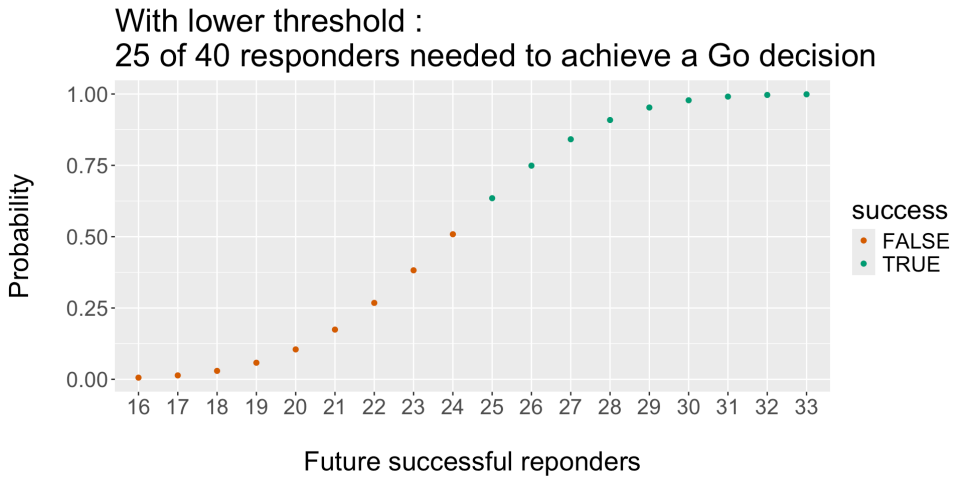
```r
control = 0.6
confidence_sixty = 0.7
result <- predprob(
  x = 16, n = 23, Nmax = 40, p = control, thetaT = confidence_sixty,
  parE = c(0.6, 0.4)
)
result$result
```

```
[1] 0.8211011
```

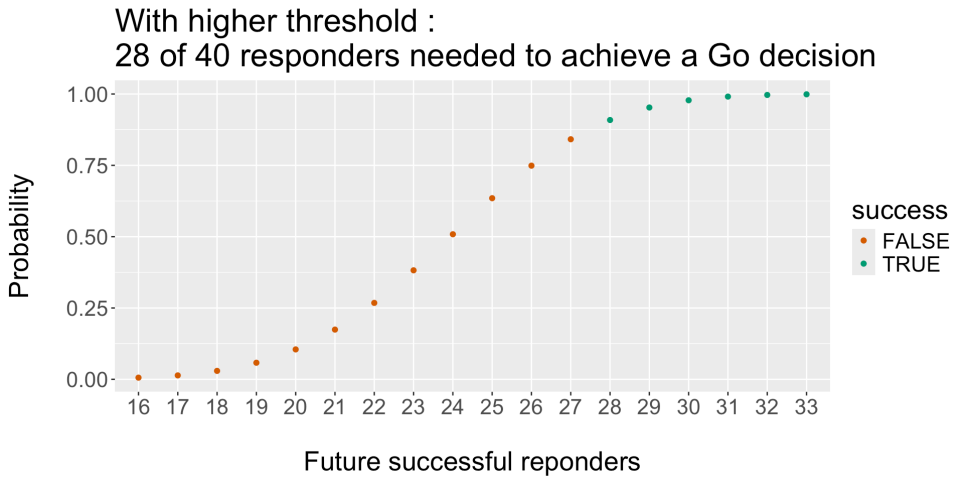
```r
confidence_ninety = 0.9
result_high_thetaT <- predprob(
  x = 16, n = 23, Nmax = 40, p = control, thetaT = confidence_ninety,
  parE = c(0.6, 0.4)
)
result_high_thetaT$result
```

```
[1] 0.5655589
```

Predictive Posterior Probability



(a) $P(\pi > 0.6 | \ data) > 70\%$



(b) $P(\pi > 0.6 | \ data) > 90\%$

Figure 2: Predictive Posterior CDF for different Efficacy Rules

Operating Characteristics

Operating Characteristics : threshold for Success (and failure):

- Efficacy criteria, e.g. we would stop for Efficacy if :

$$Pr(RR > p1) > tU$$

- Futility criteria, eg. we would stop for Futility if :

$$Pr(RR < p0) > tL$$

Rules and Operating characteristics. A use case for `ocPostprob()`:

- Look for Efficacy: Go if $P(\pi > 60\% | \ data) > 90\%$
- Look for Futility: Stop if $P(\pi < 60\% | \ data) > 70\%$
- Prior of treatment arm $Beta(0.6, 0.4)$.

```
set.seed(2025)
res <- ocPostprob(
  nnE = c(23, 40), truep = 0.60, p0 = 0.60, p1 = 0.69, tL = 0.70, tU = 0.90, parE = c(0.6, 0.4),
  sim = 500, wiggle = TRUE, nnF = c(23, 40)
)
res$oc
```

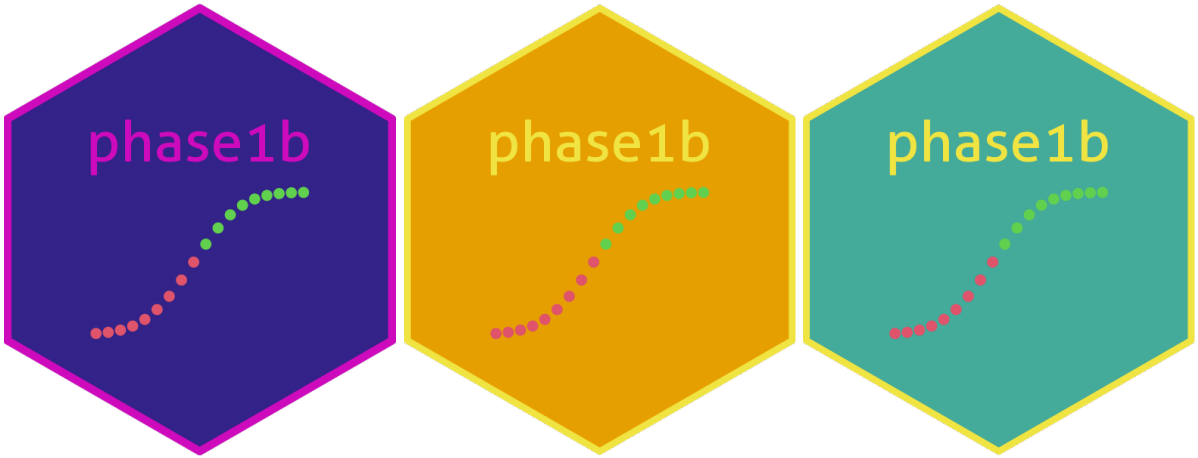
```
  ExpectedN PrStopEarly PrEarlyEff PrEarlyFut PrEfficacy PrFutility PrGrayZone
1    35.182       0.294      0.014       0.28      0.018        0.4      0.582
```

Expanded features

…. and wiggle room!

	SOC uncertainty	single-arm	two-arm	simulation	plotting	boundaries
postprob		✓				
postprobDist	✓	✓				
predprob		✓				
predprobDist	✓	✓				
ocPostprob		✓		✓		
ocPostprobDist	✓	✓		✓		
ocPredprob		✓		✓		
ocPredprobDist	✓	✓		✓		
ocRctPostprobDist	✓	✓	✓	✓		
ocRctPredprobDist	✓	✓	✓	✓		
plotBeta				✓	✓	
plotDecision					✓	
plotOc					✓	
plotBounds					✓	
boundsPostprob						✓
boundsPredprob						✓

Concluding remarks



- Big thank you to Daniel Sabanés Bové for mentorship. Data Science Acceleration colleagues Isaac Gravestock, John Kirkpatrick, Craig Gower-Paige et al who collaborated and supported.
- Extension to other therapeutic areas that use response rate as endpoint if beta priors are appropriate
- Open issues here. Contact me to collaborate.
- Statistical Engineering Journey, PDD Tech Talk 16th (EU), 19th June (SH/SSF) for more on the `phase1b` journey

Some references

- Thall P F, Simon R (1994), Practical Guidelines for Phase IIB Clinical Trials, Biometrics, 50, 337-349
- Lee J J, Liu D D (2008), A Predictive probability design for phase II cancer clinical trials, 5(2), 93-106, Clinical Trials
- Yeo, A T, Sabanés Bové D, Elze M, Pourmohamad T, Zhu J, Lymp J, Teterina A (2024). Phase1b : Calculations for decisions on Phase 1b clinical trials. R package version 1.0.0, https://genentech.github.io/phase1b
- Inclusive Speaker Orientation Linux Foundation
- Zeileis, Fisher, Hornik, Ihaka, McWhite, Murrell, Stauffer, Wilke (2020). colorspace: A Toolbox for Manipulating and Assessing Colors and Palettes. Journal of Statistical Software.
- Code for this presentation

Backupslides

Standard of Care Distribution when unknown

Using the approach by Thall and Simon (Biometrics, 1994), this evaluates the posterior probability of achieving superior response rate in the treatment group `E` compared to standard of care `S`.

- The desired improvement is denoted as `delta`. There are two options in using `delta`. The absolute case when `relativeDelta = FALSE` and relative as when `relativeDelta = TRUE`.

Desired improvement `delta`, two approaches:

1. The absolute case is when we define an absolute delta, greater than `P_S`, the response rate of the standard of care or control or `S` group such that the posterior is `Pr(P_E > P_S + delta | data)`.
2. In the relative case, we suppose that the treatment group's response rate is assumed to be greater than `P_S + (1-P_S) * delta` such that the posterior is `Pr(P_E > P_S + (1 - P_S) * delta | data)`.

User facing functions with this feature include

- `postprobDist()`
- `predprobDist`
- `ocPostprobDist()`
- `ocPredprobDist()`

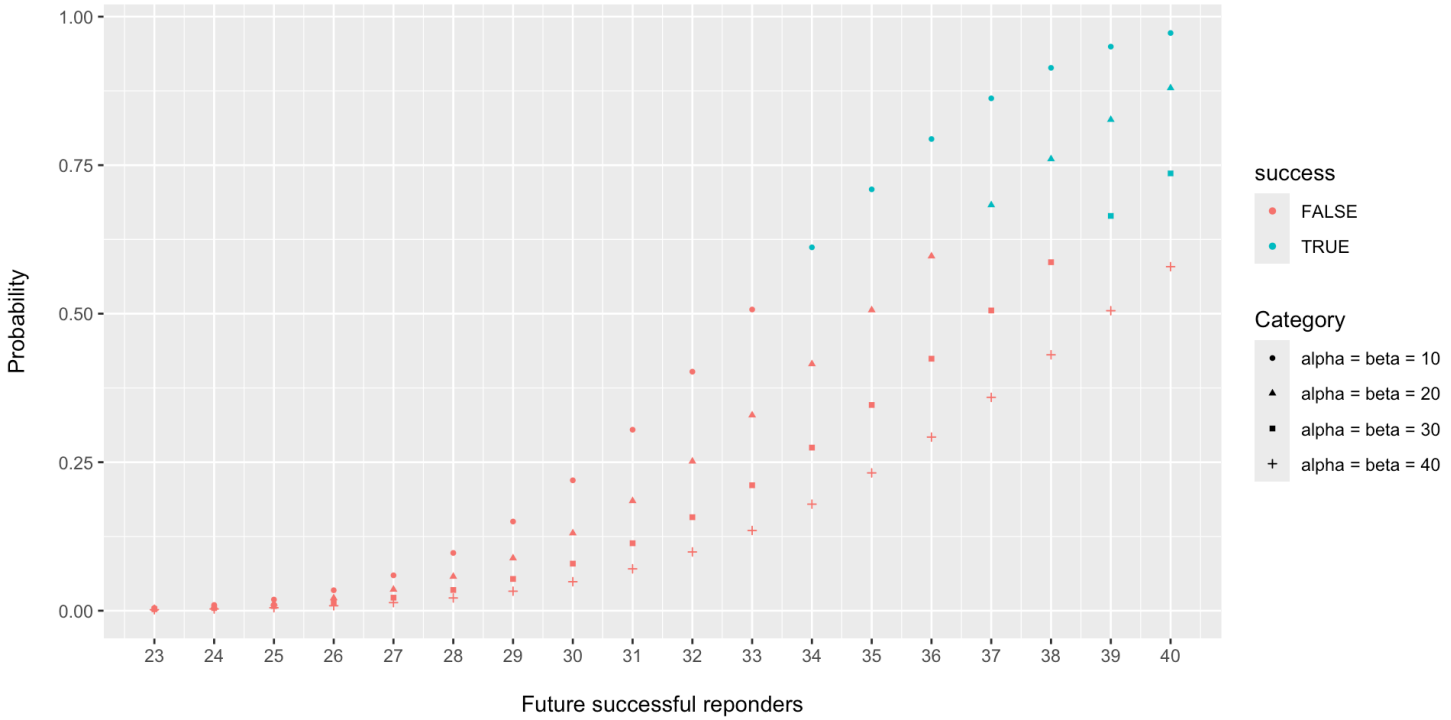
Does CDF improve with increased “data availability” of priors ?

Here we have an interim result of 16 and 23 patients. (input for Likelihood)

We have a threshold of 60% success. Our standard of care is 60%. Our priors have a mean of 50%.

Does CDF improve with increased “data availability” of priors ?

No.



Predictive posterior probability for Decision 1 :

The criteria for Decision 1 for Interim looks are :

- interim GO : `P( success at final) > phiU`
- interim STOP : `P( success at final) < phiL`

The criteria for Decision 1 for Final looks are:

- Final GO : `P( response rate > p0 | data) > tT`
- Final STOP : `P( response rate > p0 | data ) < tT`

Predictive posterior probability for Decision 2:

The criteria for Decision 2 for Interim looks are :

- Interim GO : `P ( success at final) > phiU`
- Interim STOP : `P (failure at final ) > phiFu`

The criteria for Decision 2 for Futility looks are :

- Final GO : `P( response rate > p0) > tT`
- Final STOP : `P( response rate  < p1) > tF`