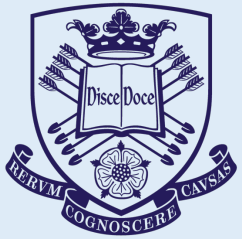
Assurance Methods for Clinical Trials: practical tools and new developments

James Salsbury[1], Jeremy Oakley[1], Steven Julious[2], Lisa Hampson[3]

[1]School of Mathematics and Statistics, The University of Sheffield
[2]School of Health and Related Research, The University of Sheffield
[3]Novartis Pharma AG, Basel, Switzerland



The University Of Sheffield.

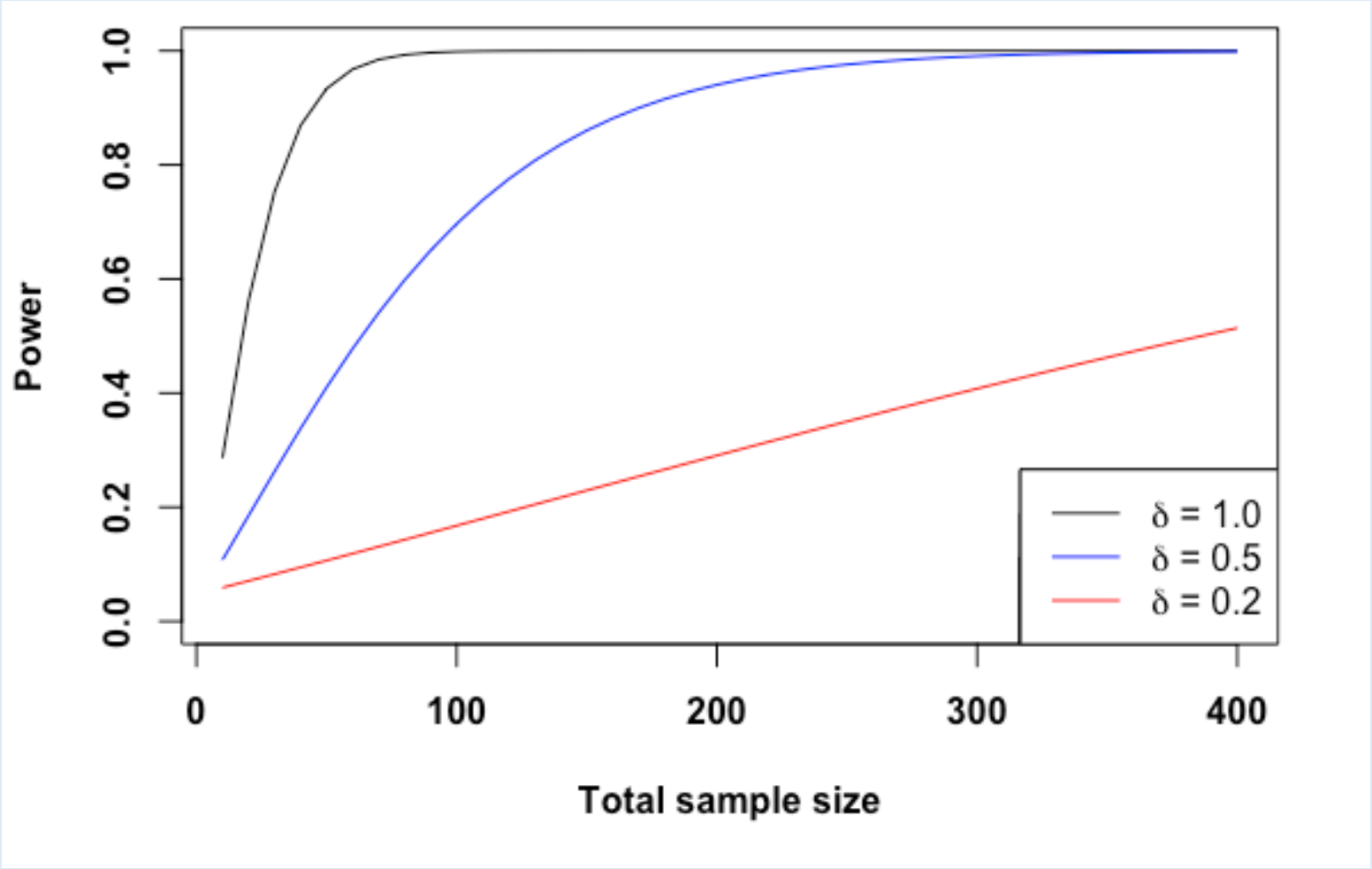
NOVARTIS

Introduction

- Overview of power/assurance
- An example of elicitation/assurance
- My current area of research

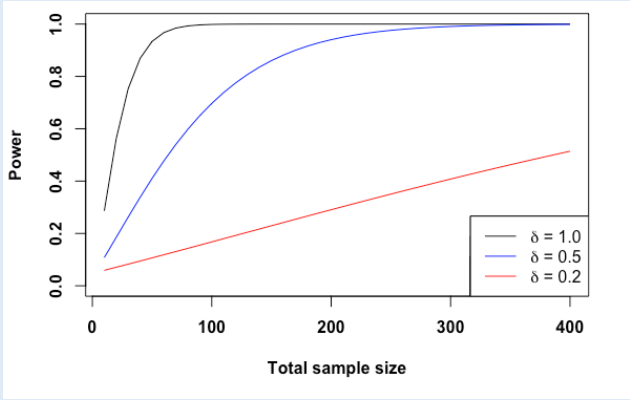
What is assurance?

- What is power?



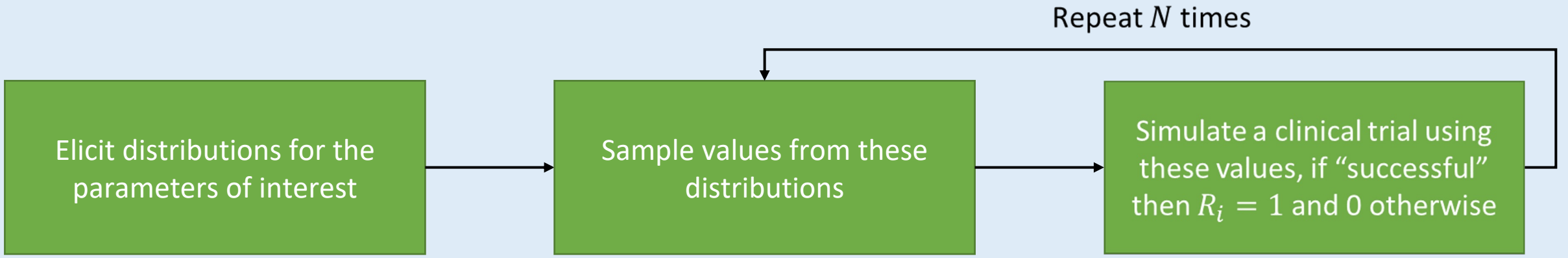
What is assurance?

- What is power?



- *"The probability of detecting a difference between groups **given** a difference exists"*
- However, this is **conditional**. Can we do better?
- By eliciting distributions for the parameters of interest then this probability is now **unconditional**
- Trial can still be analysed using frequentist methods

How do you calculate assurance?



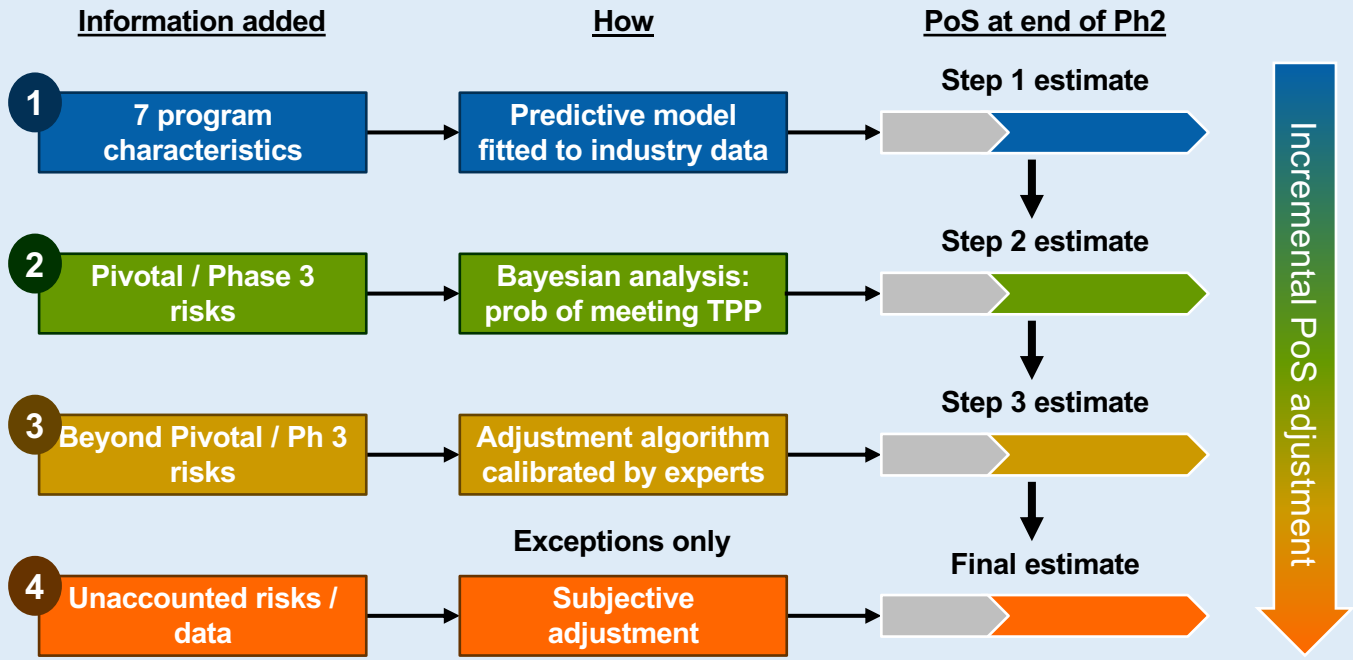
The assurance is then estimated as $\hat{P}(R) = \frac{1}{N}\sum_{i=1}^{N} R_i$

What are the barriers of using assurance in practice?

- Eliciting distributions for the parameters of interest: it can be difficult to convey beliefs/uncertainty
- Resources: time and/or money
- What work has been done in this area?
  - O'Hagan et al (2005) consider normal and binomial outcomes
  - Ren and Oakley (2014) consider time-to-event outcomes
  - Alhussain and Oakley (2020) consider uncertainty about variances
- **SH**effield **EL**icitation **F**ramework (SHELF) is a tool used for eliciting a probability distribution from an expert or a group of experts

Probability of Success (PoS) at Novartis

Within a Phase	60%	95%	70%
Probability of Approval	60% * 95% = **57%**	(potential gap)	
Probability of Success	60% * 95% * 70% = **40%**		



Hampson LV, Bornkamp B, Holzhauer B et al. *Pharmaceutical Statistics* 2022; 21:439

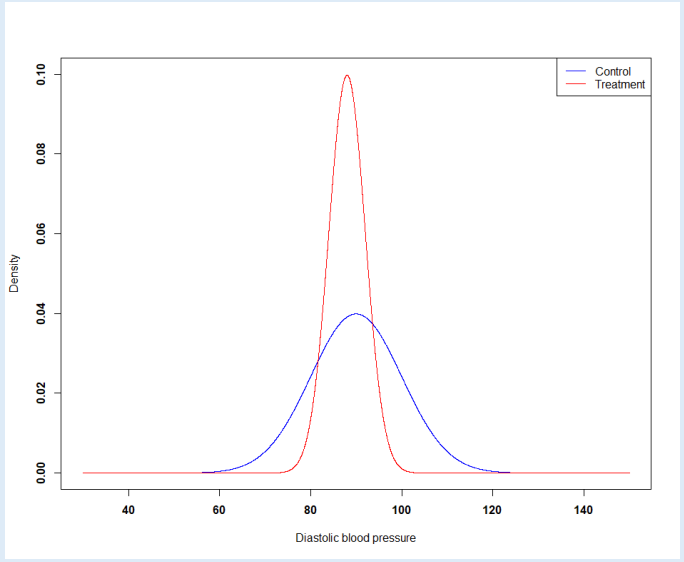
Holzhauer B, Hampson LV, Gosling JP et al. *Pharmaceutical Statistics* 2022. In press.

Trial with Normal endpoints

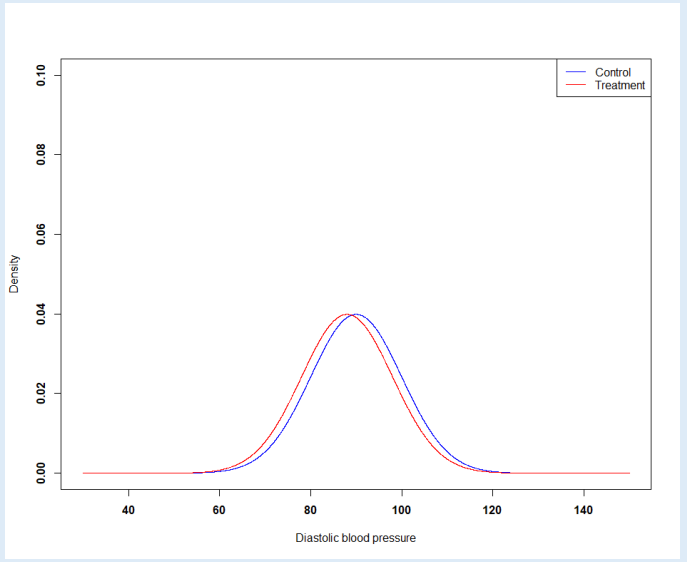
- Trial to lower diastolic blood pressure
- Control observations $X_1, \ldots, X_{n_c} \sim N(\mu_c, \sigma_c^2)$ and treatment observations $Y_1, \ldots, Y_{n_t} \sim N(\mu_t, \sigma_t^2)$
- We define the mean difference between control and treatment as $\delta = \mu_c - \mu_t$
- There is a lot of literature on eliciting judgments about δ
- However, little work has been done on eliciting beliefs about variances (σ_t^2)

Why does the variance matter? (1/2)

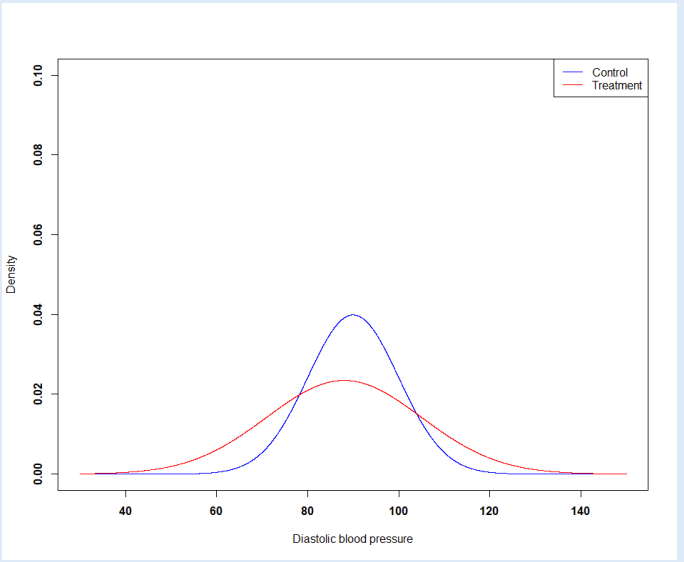
$\text{Control} \sim N(90, 10^2), \text{Treatment} \sim N(88, 4^2)$



$\text{Control} \sim N(90, 10^2), \text{Treatment} \sim N(88, 10^2)$

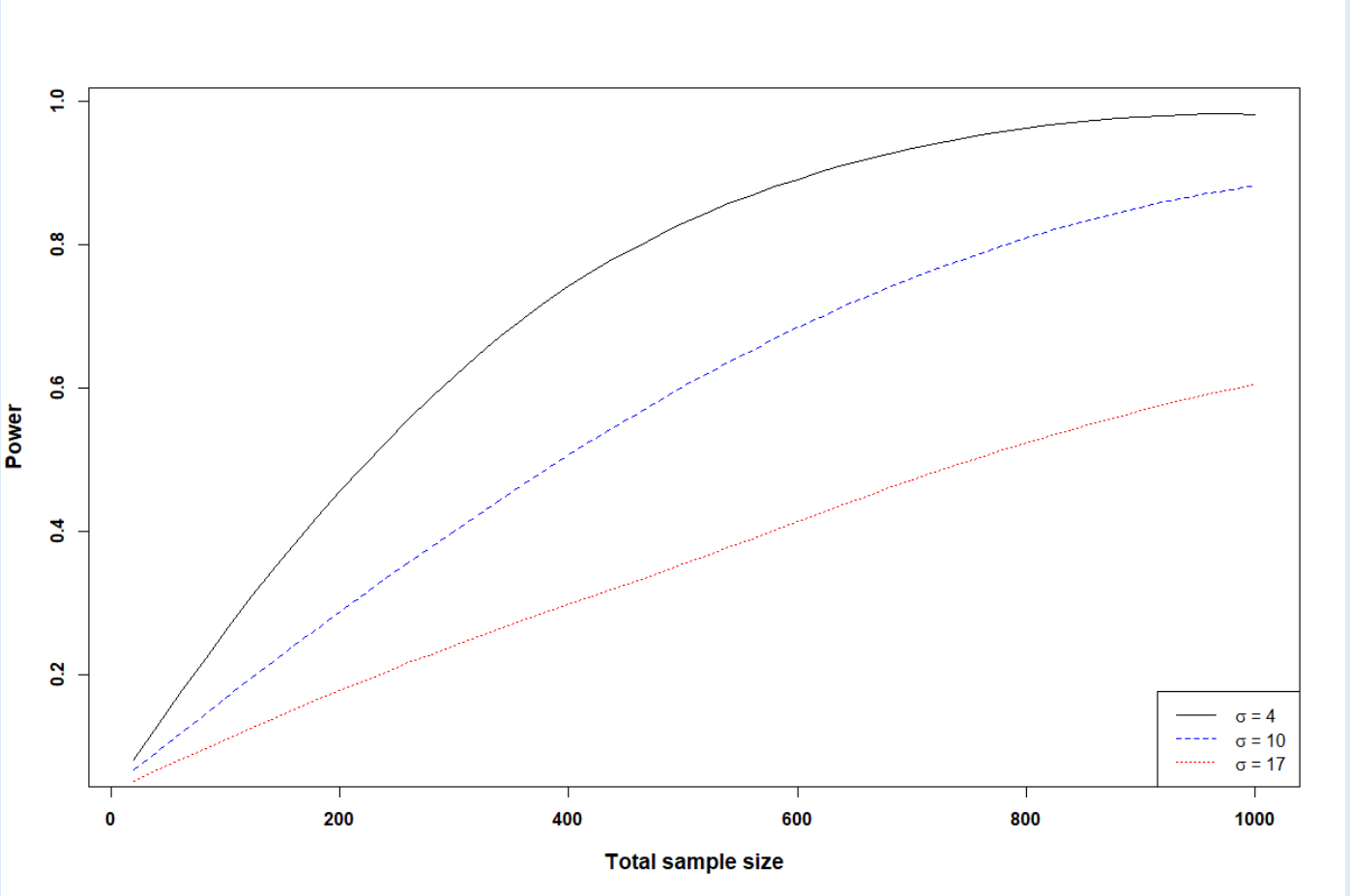


$\text{Control} \sim N(90, 10^2), \text{Treatment} \sim N(88, 17^2)$



- μ_t is the same in all three plots
- Only difference is σ_t^2
- How can we specify the spread around the mean?

Why does the variance matter? (2/2)

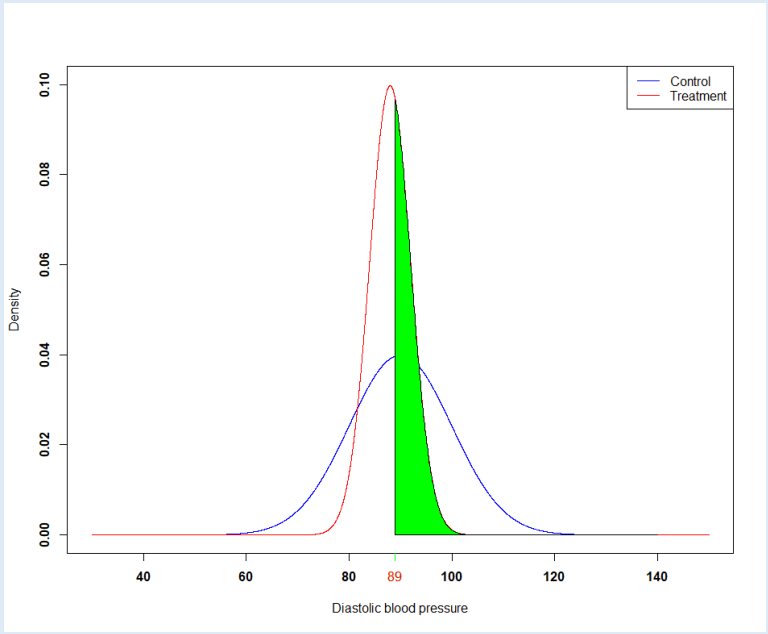


Choosing the interval

- We can specify our uncertainty about the variance by eliciting a proportion contained in fixed interval $[k_1, k_2]$
- For example, $\mu_c = 90$, $\mu_t = 88$ in the previous plots. Therefore $\delta = 2$
- We can ask the question "*Given that the drug works as expected (i.e $\delta = 2$), what proportion of patients are not expected to benefit from the drug?*"
- We assume that a response ≥ 89 will be considered to have **not** benefited from the drug
- Therefore we are specifying $[k_1, k_2] = [89, \infty]$

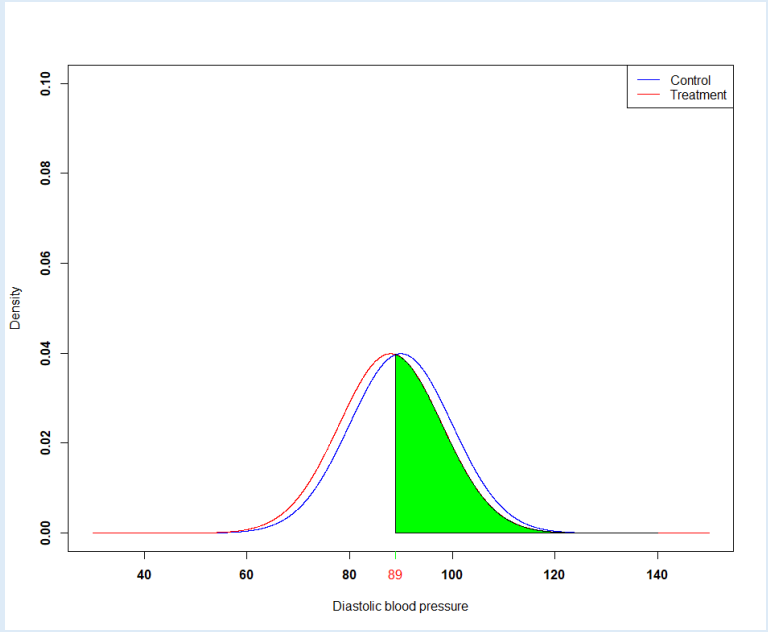
What this interval corresponds to

$\text{Control} \sim N(90, 10^2), \text{Treatment} \sim N(88, 4^2)$



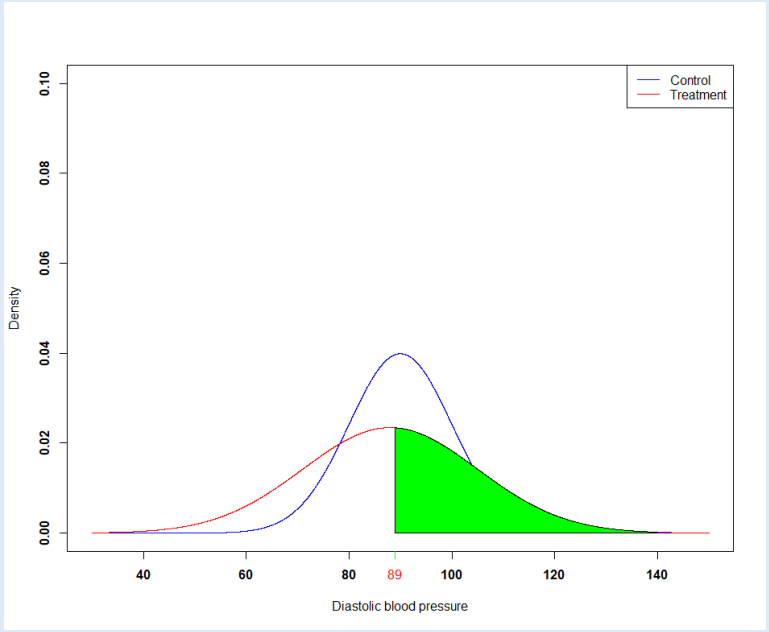
40% of patients in Green Region

$\text{Control} \sim N(90, 10^2), \text{Treatment} \sim N(88, 10^2)$



46% of patients in Green Region

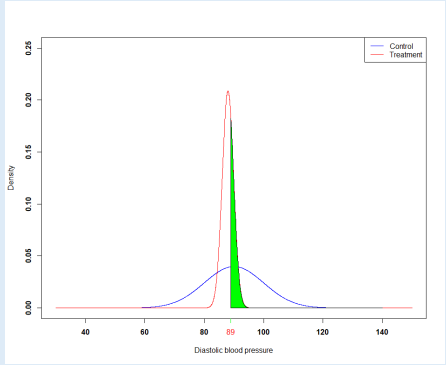
$\text{Control} \sim N(90, 10^2), \text{Treatment} \sim N(88, 17^2)$

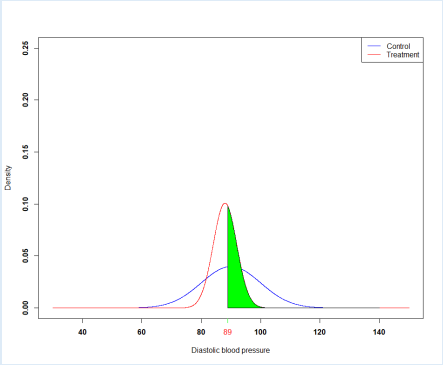


48% of patients in Green Region

Eliciting σ_t^2

- We have asked the question "*Given that the drug works as expected (i.e* $\delta = 2$*), what proportion of patients are not expected to benefit from the drug?*"
- Therefore we are specifying $[k_1, k_2] = [89, \infty]$
- A: "*Between 30% and 40% of patients are not expected to benefit from the drug*"
- We can then (numerically) solve for σ_t^2 (by finding σ_t^2 that means that the area under the curve equals 0.3 and 0.4)
- σ_t^2= [3.65, 15.60]

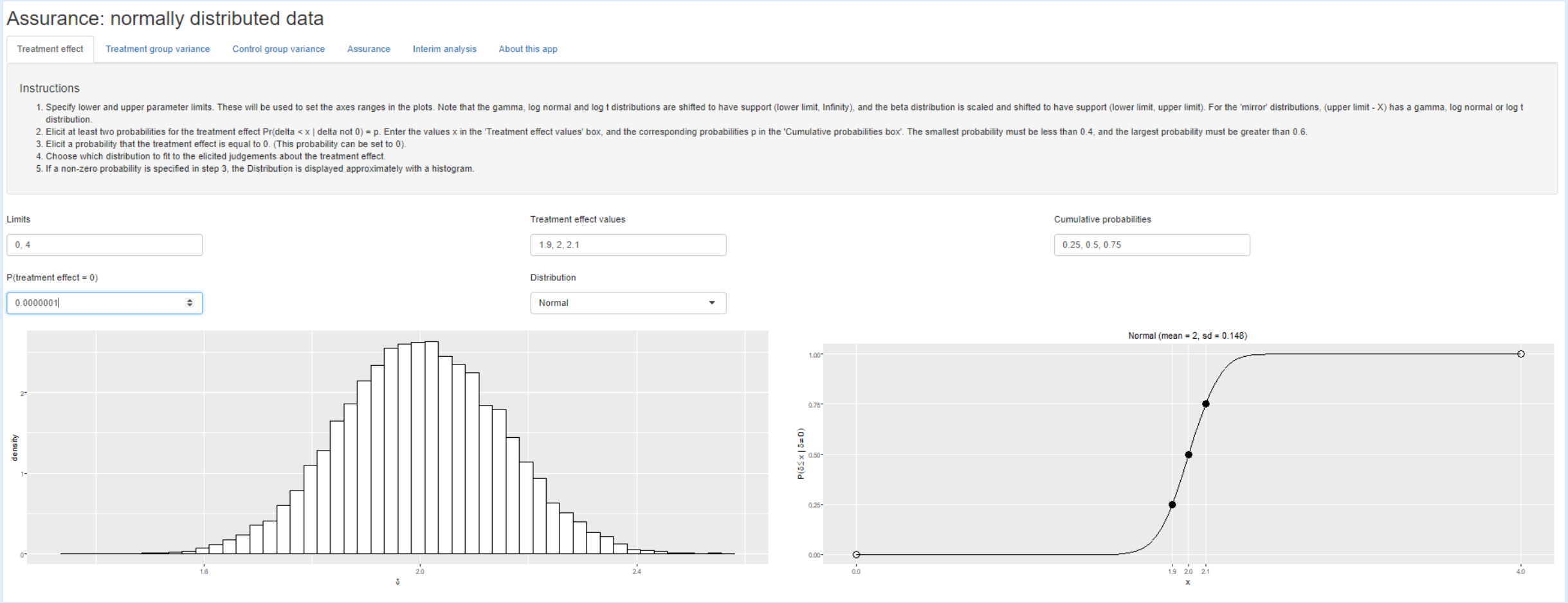




Calculating assurance

- We can then specify a probability distribution for σ_t^2 (using these elicited judgments)
- Can feed this distribution into the flow-chart and calculate a value for the assurance
- However, there is an R package {*assurance*} which uses a Shiny app to elicit these beliefs

Shiny app (1/3)


Assurance: normally distributed data
Treatment effect
Treatment group variance
Control group variance
Assurance
Interim analysis
About this app
Instructions
1. Specify lower and upper parameter limits. These will be used to set the axes ranges in the plots. Note that the gamma, log normal and log t distributions are shifted to have support (lower limit, Infinity), and the beta distribution is scaled and shifted to have support (lower limit, upper limit). For the 'mirror' distributions, (upper limit - X) has a gamma, log normal or log t distribution.
2. Elicit at least two probabilities for the treatment effect Pr(delta < x | delta not 0) = p. Enter the values x in the 'Treatment effect values' box, and the corresponding probabilities p in the 'Cumulative probabilities box'. The smallest probability must be less than 0.4, and the largest probability must be greater than 0.6.
3. Elicit a probability that the treatment effect is equal to 0. (This probability can be set to 0).
4. Choose which distribution to fit to the elicited judgements about the treatment effect.
5. If a non-zero probability is specified in step 3, the Distribution is displayed approximately with a histogram.
Limits
0, 4
Treatment effect values
1.9, 2, 2.1
Cumulative probabilities
0.25, 0.5, 0.75
P(treatment effect = 0)
0.0000001
Distribution
Normal
density
δ
Normal (mean = 2, sd = 0.148)
P(δ ≤ x | δ≠0)
x

Shiny app (2/3)

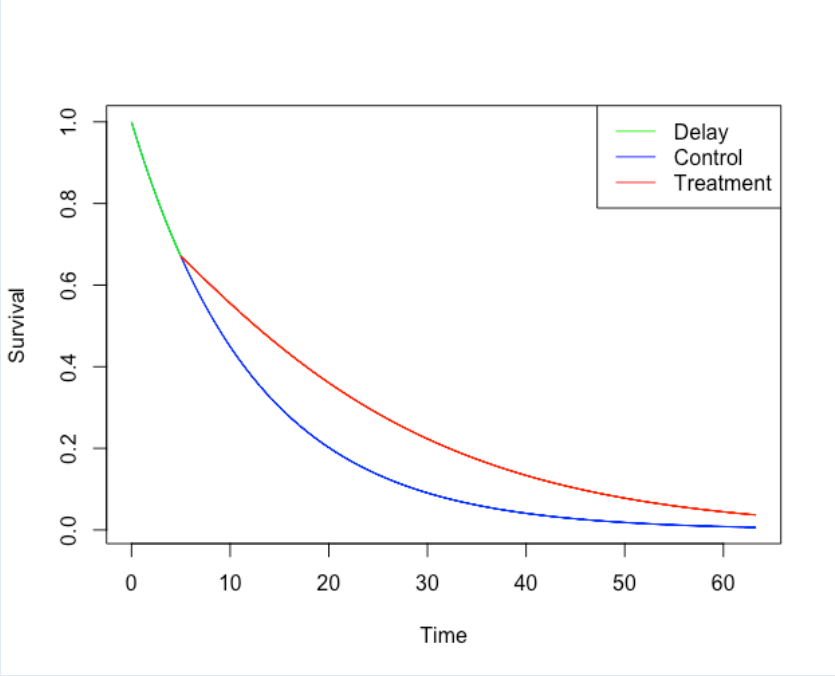

Assurance: normally distributed data
Treatment effect
Treatment group variance
Control group variance
Assurance
Interim analysis
About this app
Instructions
1. Specify a hypothetical value for the treatment group median
2. Specify an interval of treatment group responses. One of (-Inf, a), (a, median), (median, b), (b, Inf), for constants a and b.
3. Elicit lower and upper values for the proportion of patients with responses in the specified interval, corresponding to the choices of 'Proportion quantiles'.
4. Select either a gamma or log normal distribution, to be fitted to the treatment group precision, using the two elicited values for the proportion.
Treatment group median
88
Treatment group interval
89, Inf
Lower/upper proportions
0.3, 0.4
Proportion quantiles
0.05, 0.95
lower (0.05 quantile) proportion = 0.3
upper (0.95 quantile) proportion = 0.4
x
Elicited distributions for the standard deviation
density
σ
Distribution for precision
gamma
lognormal
Selected distribution
Gamma

Shiny app (3/3)


Assurance: normally distributed data
Treatment effect
Treatment group variance
Control group variance
Assurance
Interim analysis
About this app
1.00
0.75
0.50
0.25
0.00
assurance
25
50
100
250
500
1000
sample size per arm
Custom sample sizes
Treatment group sample size
50
Control group sample size
50
Assurance
0.88

My current research

- I have been working in collaboration with Novartis to develop assurance methods for when delayed treatment effects (DTE) are present
- DTE exhibits behaviour shown in the following plot:



What makes DTE hard?

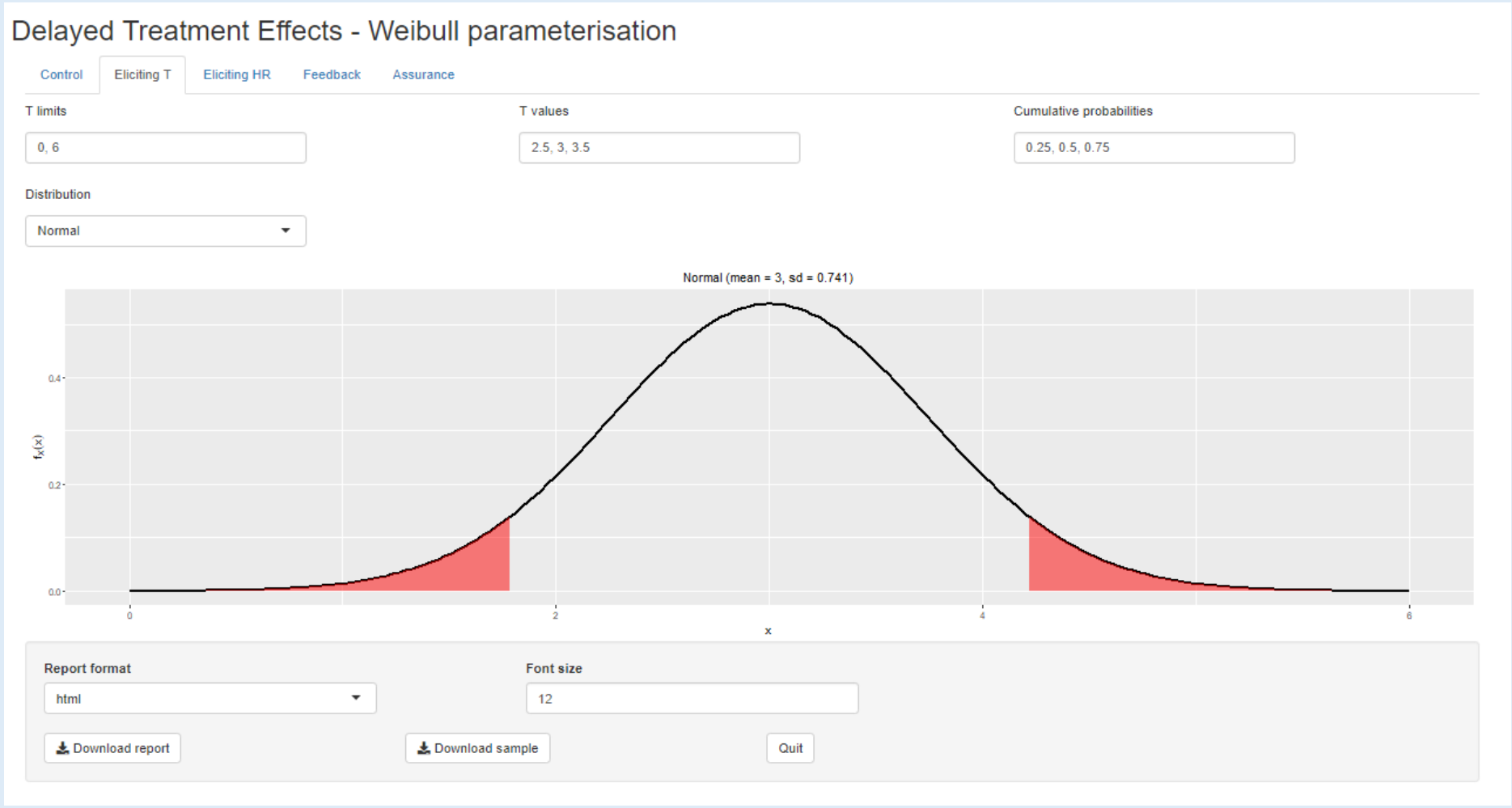
- In design:
  - When are the curves going to separate?
  - When to plan for interim analyses?

- In analysis:
  - Proportional hazards are violated, how do you account for this?
  - Weighted log-rank test etc, RMST..

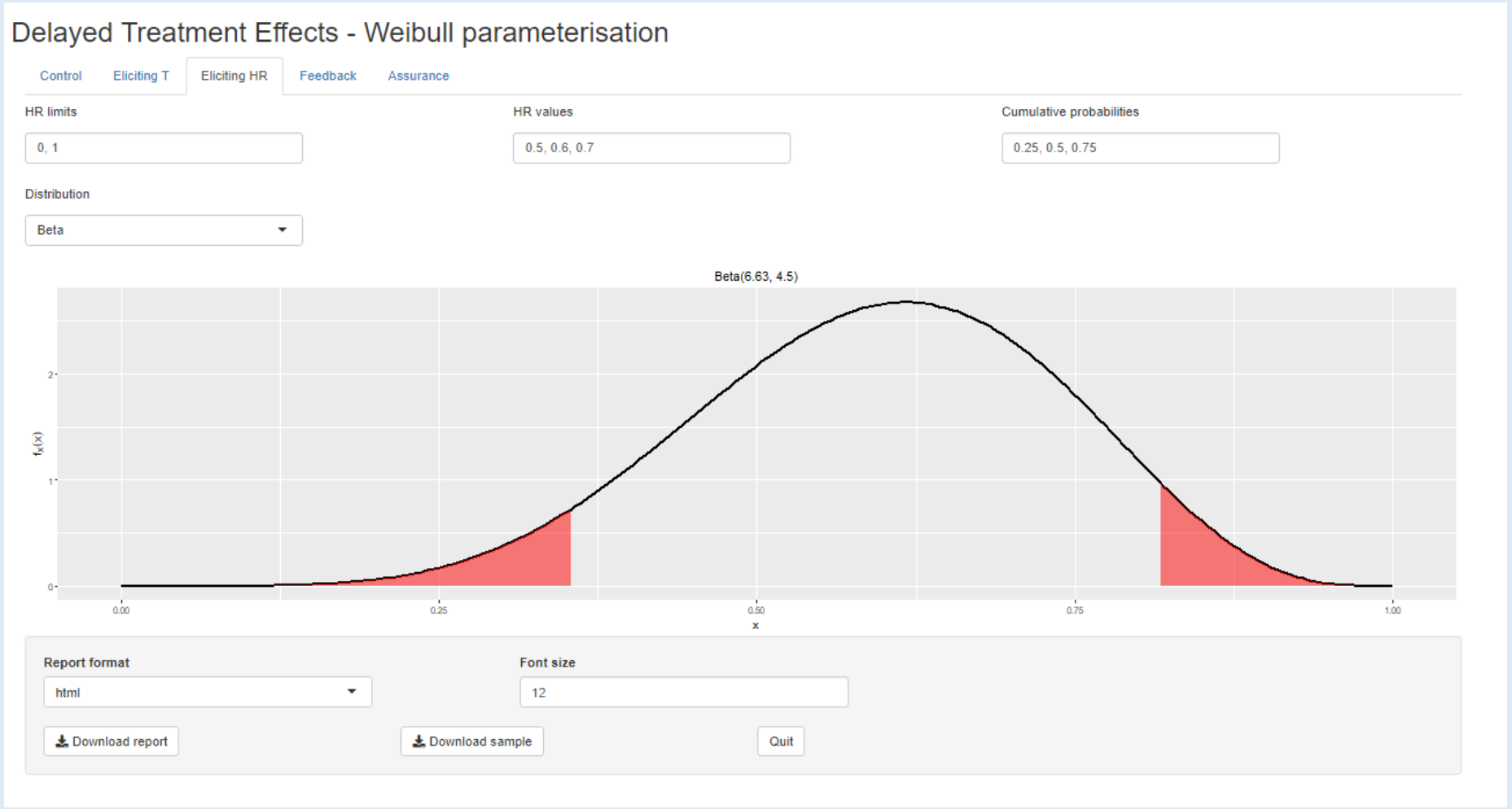
DTE assurance

- Can parameterise this problem using piecewise Weibull distributions
- We elicit distributions for T (time that the treatment starts to take effect) and the post-treatment HR
- I have created a Shiny app that facilitators can use for this problem
- Calculates an assurance based on the elicited distributions and questions about the trial

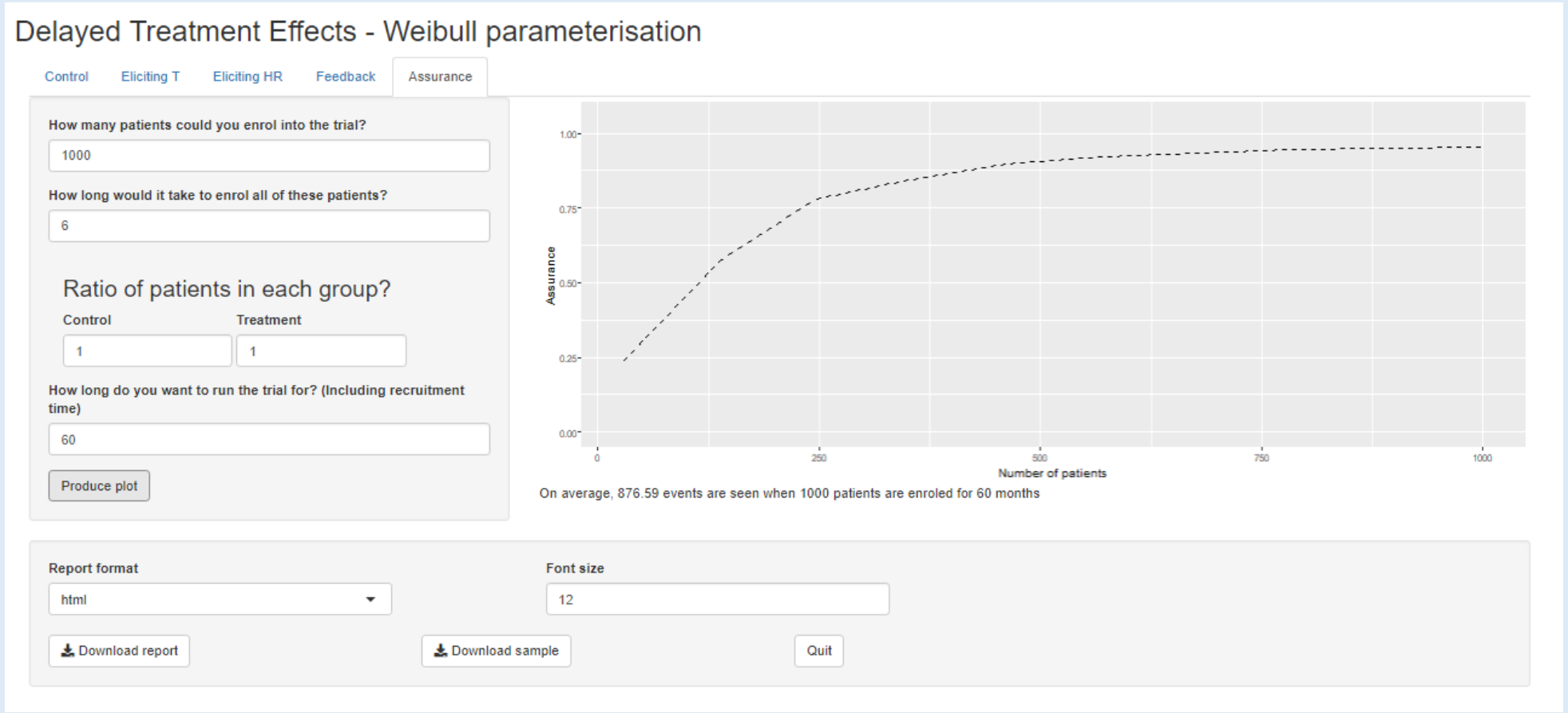
DTE assurance shiny app (1/3)



DTE assurance shiny app (2/3)



DTE assurance shiny app (3/3)



Thank you! Any questions?

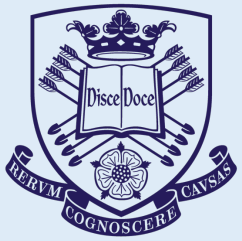
SHELF website

jsalsbury1@sheffield.ac.uk

jamesalsbury.github.io

james-salsbury

My shiny app!



The University Of Sheffield.

NOVARTIS

References

- O'Hagan, Anthony & Stevens, John & Campbell, Michael. (2005). Assurance in clinical trial design. Pharmaceutical Statistics. 4. 187 - 201. 10.1002/pst.175.

- Ren, Shijie & Oakley, Jeremy. (2014). Assurance calculations for planning clinical trials with time-to-event outcomes. Statistics in medicine. 33. 10.1002/sim.5916.

- Alhussain, Z. and Oakley, J (2020). Assurance for clinical trial design with normally distributed outcomes: eliciting uncertainty about variances. Pharmaceutical Statistics. 19(6) pp. 827-839.

- https://www.psiweb.org/docs/default-source/default-document-library/alessandro-previtali-slides.pdf?sfvrsn=224dedb_0