How statistical tools can help your Proof of Concept studies

Nicolas Bonnet[1]

[1]SANOFI R&D, CSO, Biostatistics, Early Clinical Development

3rd Joint Conference of European Human Pharmacological Societies
Bruxelles, may 2015

Outline

- Costs dominated by late-phase development
- Change in development paradigm
  - Exploratory / confirmatory phase

(Modified from Orloff et al. 2009)

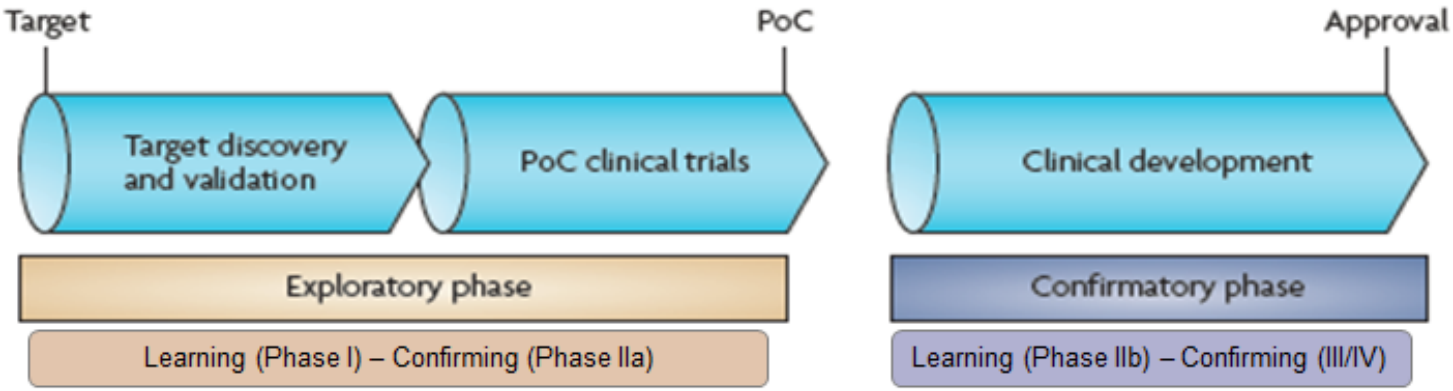
- Strategy
  - Minimize resource investment until PoC is established
  - Confidence to further invest in full development of candidate
  - Confidence to stop a non-promising compound (“go/no-go”)
  - Early development should focus on PoC and dose-finding
    (after the MTD determination)

- Costs dominated by late-phase development
- Change in development paradigm
  - Exploratory / confirmatory phase

(Modified from Orloff et al. 2009)

- Strategy
  - Minimize resource investment until PoC is established
  - Confidence to further invest in full development or candidate
  - Confidence to stop a non-promising compound (fail early)
  - Early development should focus on PoC and MTD
  (after the MTD determination)

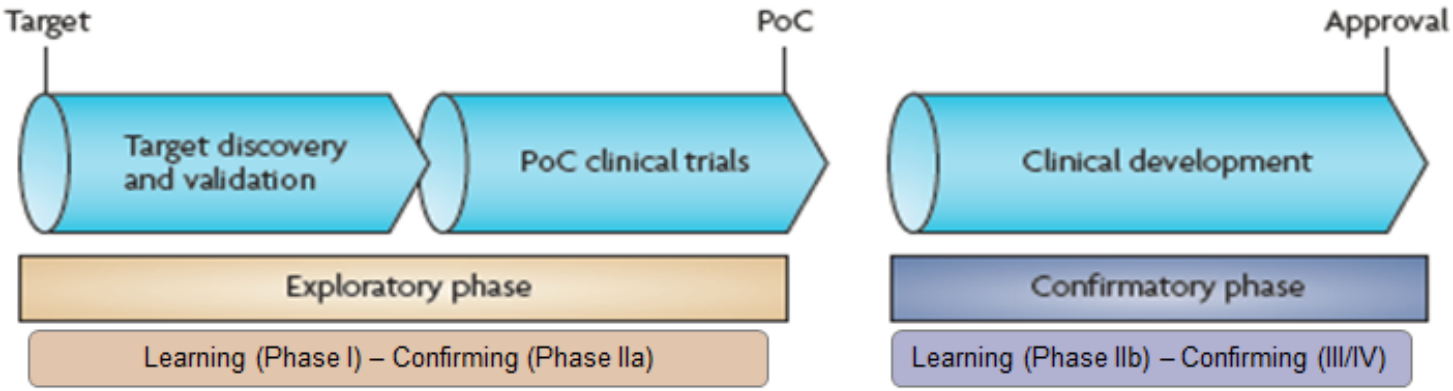
- Costs dominated by late-phase development
- Change in development paradigm
  - Exploratory / confirmatory phase



(Modified from Orloff et al. 2009)

- Strategy
  - Minimize resource investment until PoC is established
  - Confidence to further invest in full-development of candidate
  - Confidence to stop a non-promising compound ("fast-to-fail")
  - Early development should focus on safety or tolerability (after the MTD determination)

- Costs dominated by late-phase development
- Change in development paradigm
  - Exploratory / confirmatory phase

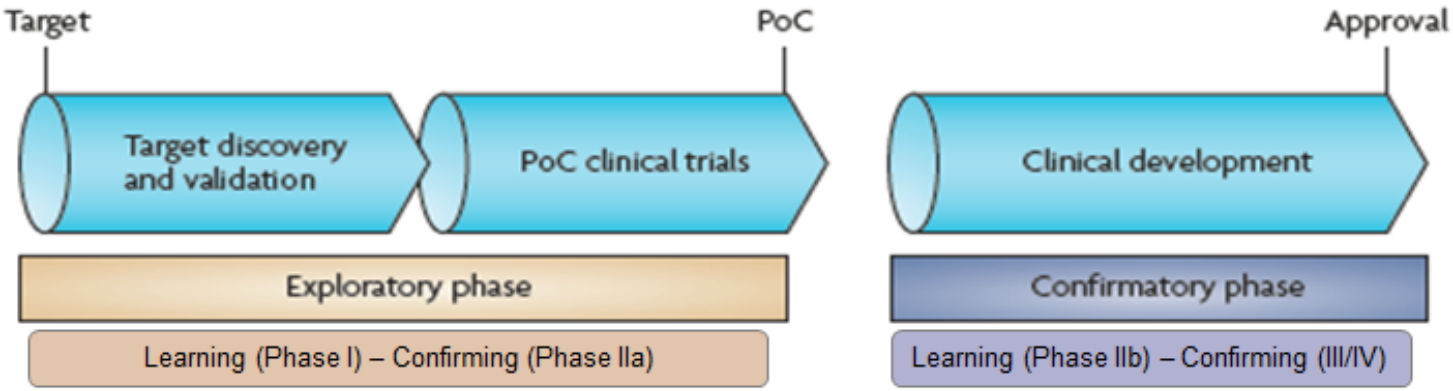


(Modified from Orloff et al. 2009)

- Strategy
  - Minimize resource investment until PoC is established
  - Confidence to further invest in full-development of candidate
  - Confidence to stop a non-promising compound ("fast-to-fail")

  $\Rightarrow$ Early development should focus on PoC as primary goal (after the MTD determination)

- Costs dominated by late-phase development
- Change in development paradigm
  - Exploratory / confirmatory phase

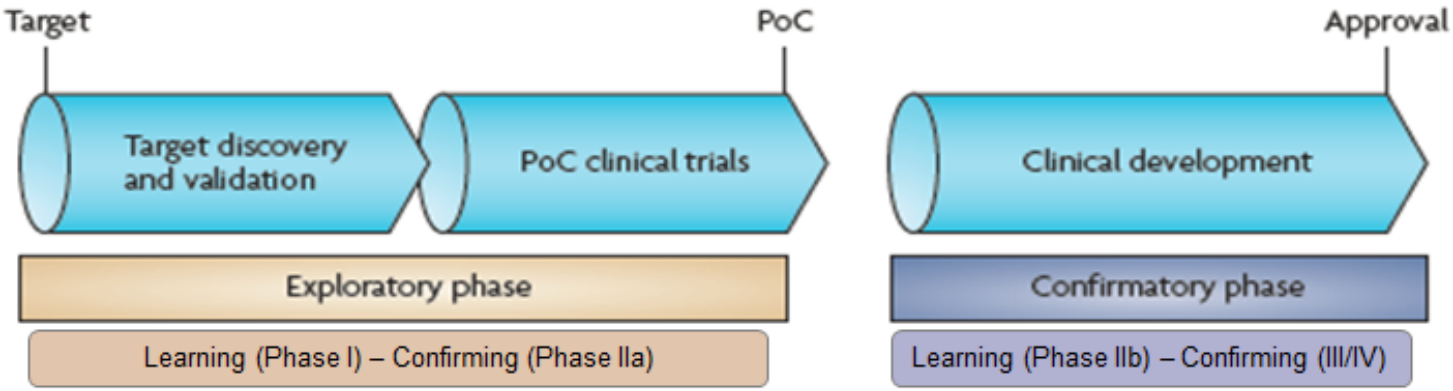


(Modified from Orloff et al. 2009)

- Strategy
  - Minimize resource investment until PoC is established
  - Confidence to further invest in full-development of candidate
  - Confidence to stop a non-promising compound ("fast-to-fail")

  $\Rightarrow$ Early development should focus on PoC as primary goal (after the MTD determination)

- Costs dominated by late-phase development
- Change in development paradigm
  - Exploratory / confirmatory phase

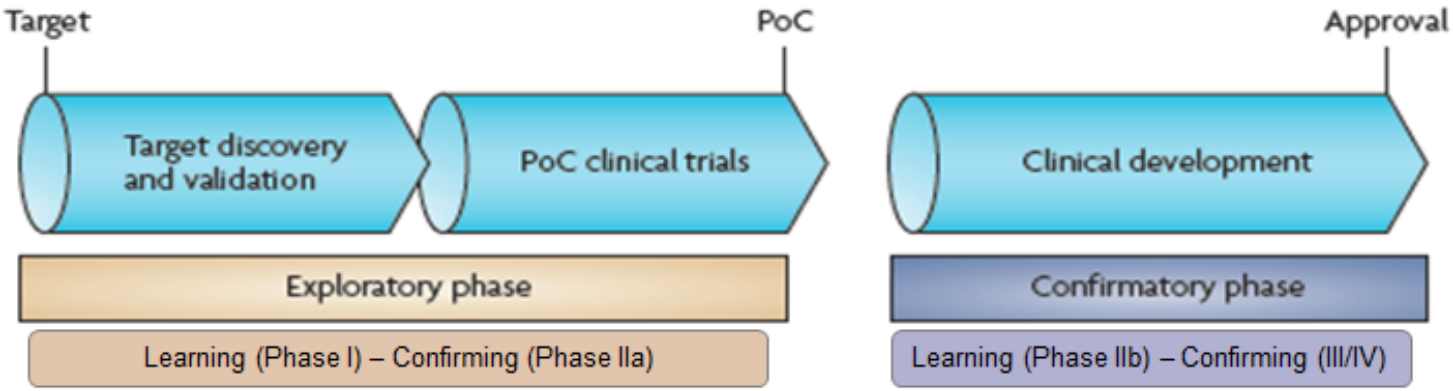


(Modified from Orloff et al. 2009)

- Strategy
  - Minimize resource investment until PoC is established
  - Confidence to further invest in full-development of candidate
  - Confidence to stop a non-promising compound ("fast-to-fail")

$\Rightarrow$ Early development should focus on PoC as primary goal (after the MTD determination)

- Costs dominated by late-phase development
- Change in development paradigm
  - Exploratory / confirmatory phase

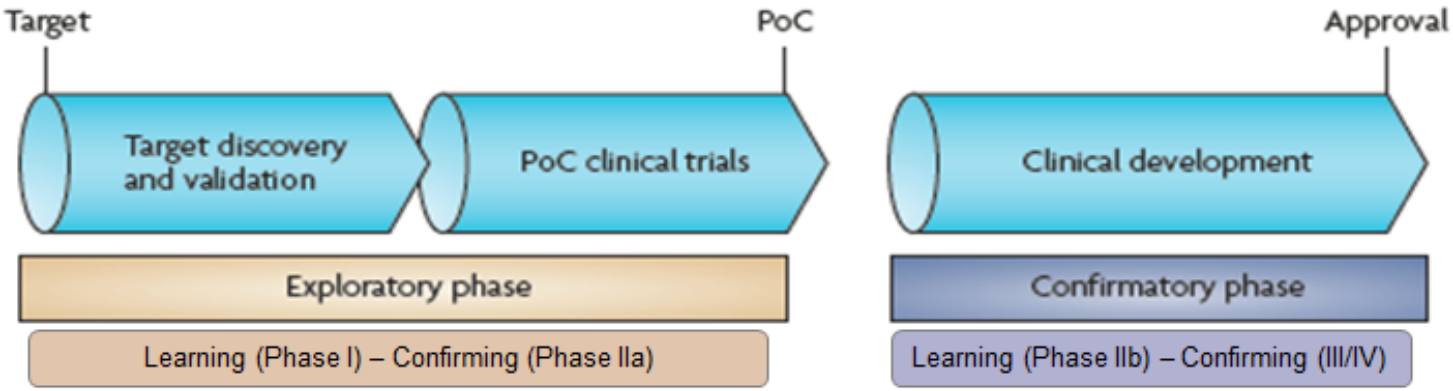


(Modified from Orloff et al. 2009)

- Strategy
  - Minimize resource investment until PoC is established
  - Confidence to further invest in full-development of candidate
  - Confidence to stop a non-promising compound ("fast-to-fail")

$\Rightarrow$ Early development should focus on PoC as primary goal (after the MTD determination)

- Costs dominated by late-phase development
- Change in development paradigm
  - Exploratory / confirmatory phase

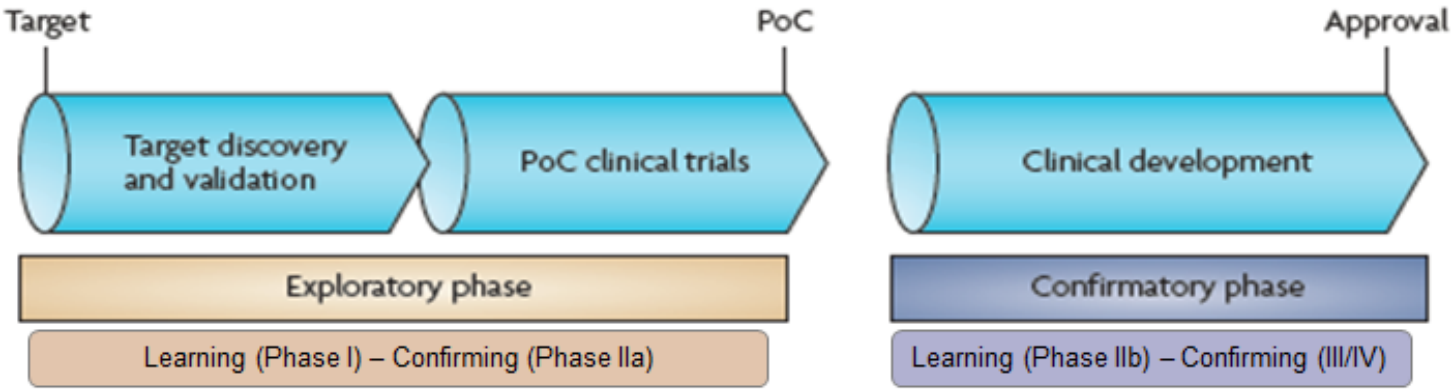


(Modified from Orloff et al. 2009)

- Strategy
  - Minimize resource investment until PoC is established
  - Confidence to further invest in full-development of candidate
  - Confidence to stop a non-promising compound ("fast-to-fail")

$\Rightarrow$ Early development should focus on PoC as primary goal (after the MTD determination)

- Costs dominated by late-phase development
- Change in development paradigm
  - Exploratory / confirmatory phase



(Modified from Orloff et al. 2009)

- Strategy
  - Minimize resource investment until PoC is established
  - Confidence to further invest in full-development of candidate
  - Confidence to stop a non-promising compound ("fast-to-fail")

  $\Rightarrow$ Early development should focus on PoC as primary goal (after the MTD determination)

Proof-of-Concept (Cartwright et al. 2010)

The earliest point in the drug development process at which the weight of evidence suggests that it is "reasonably likely" that the key attributes for success are present and the key causes of failure are absent

- Can take various form depending on the company and the projects
  - Ph IIb dose ranging
  - Ph IIa 2 parallel arms (active MTD vs placebo)
  - Early Proof of Biological Activity ...
- Expectations :
  - Early, time and cost effective demonstration of activity in humans, including an early benefit-risk assessment
  - Increased POS and de-risks further investments
  - PoC study outcome should yield clear "Go/NoGo" decision

- Well-designed PoC studies
  1. Clear predefined "Go/No Go" criteria derived from high but reasonable assumptions
  2. Design adequate to yield
     - High probability of "Go" decision for candidates which truly meet PoC criteria (make this decision reliably)
     - Strong false positive control (stop non promising compound)
- Targeted treatment difference
- Uncertainty
- Type I error/Type II error

- Well-designed PoC studies
  1. Clear predefined "Go/No Go" criteria derived from high but reasonable assumptions
  2. Design adequate to yield
     - High probability of "Go" decision for candidates which truly meet PoC criteria (make this decision reliably)
     - Strong false positive control (stop non promising compound)
- Targeted treatment difference
- Uncertainty
- Type I error/Type II error

- Well-designed PoC studies
  1. Clear predefined "Go/No Go" criteria derived from high but reasonable assumptions
  2. Design adequate to yield
     - High probability of "Go" decision for candidates which truly meet PoC criteria (make this decision reliably)
     - Strong false positive control (stop non promising compound)
- Targeted treatment difference
- Uncertainty
- Type I error/Type II error

- Well-designed PoC studies
  1. Clear predefined "Go/No Go" criteria derived from high but reasonable assumptions
  2. Design adequate to yield
     - High probability of "Go" decision for candidates which truly meet PoC criteria (make this decision reliably)
     - Strong false positive control (stop non promising compound)
- Targeted treatment difference
- Uncertainty
- Type I error/Type II error

- Type I error (α): probability of falsely declaring PoC achieved
  - Type I error is "lost investment" risk
  - Typically set at 5%
  - Need to be controlled in "fast to fail" paradigm
- Type II error (β): probability of falsely declaring PoC failed
  - Type II error is "lost opportunity" risk
  - Typically set at 10%-20% $\rightarrow$ corresponds to power of 80%-90% (i.e. 1-β)
  - A power set at 80% means that we have 20% of chance to have a failed study for a compound that reach the predefined target
    - $\rightarrow$ No more chance later on for the compound!
    - $\rightarrow$ Are we ready to take that risk ?

- Type I error (α): probability of falsely declaring PoC achieved
  - Type I error is "lost investment" risk
  - Typically set at 5%
  - Need to be controlled in "fast to fail" paradigm
- Type II error (β): probability of falsely declaring PoC failed
  - Type II error is "lost opportunity" risk
  - Typically set at 10%-20% $\rightarrow$ corresponds to power of 80%-90% (i.e. 1-β)
  - A power set at 80% means that we have 20% of chance to have a failed study for a compound that reach the predefined target
    - $\rightarrow$ No more chance later on for the compound!
    - $\rightarrow$ Are we ready to take that risk ?

- Type I error (α): probability of falsely declaring PoC achieved
  - Type I error is "lost investment" risk
  - Typically set at 5%
  - Need to be controlled in "fast to fail" paradigm
- Type II error (β): probability of falsely declaring PoC failed
  - Type II error is "lost opportunity" risk
  - Typically set at 10%-20% $\rightarrow$ corresponds to power of 80%-90% (i.e. 1-β)
  - A power set at 80% means that we have 20% of chance to have a failed study for a compound that reach the predefined target
    - $\rightarrow$ No more chance later on for the compound!
    - $\rightarrow$ Are we ready to take that risk?

- Type I error (α): probability of falsely declaring PoC achieved
  - Type I error is "lost investment" risk
  - Typically set at 5%
  - Need to be controlled in "fast to fail" paradigm
- Type II error (β): probability of falsely declaring PoC failed
  - Type II error is "lost opportunity" risk
  - Typically set at 10%-20% $\rightarrow$ corresponds to power of 80%-90% (i.e. 1-β)
  - A power set at 80% means that we have 20% of chance to have a failed study for a compound that reach the predefined target
    - $\rightarrow$ No more chance later on for the compound!
    - $\rightarrow$ Are we ready to take that risk ?

- Type I error (α): probability of falsely declaring PoC achieved
  - Type I error is "lost investment" risk
  - Typically set at 5%
  - Need to be controlled in "fast to fail" paradigm
- Type II error (β): probability of falsely declaring PoC failed
  - Type II error is "lost opportunity" risk
  - Typically set at 10%-20% $\rightarrow$ corresponds to power of 80%-90% (i.e. 1-β)
  - A power set at 80% means that we have 20% of chance to have a failed study for a compound that reach the predefined target
    - No more chance later on for the compound
    - Are we ready to take that risk ?

- Type I error (α): probability of falsely declaring PoC achieved
  - Type I error is "lost investment" risk
  - Typically set at 5%
  - Need to be controlled in "fast to fail" paradigm
- Type II error (β): probability of falsely declaring PoC failed
  - Type II error is "lost opportunity" risk
  - Typically set at 10%-20% $\rightarrow$ corresponds to power of 80%-90% (i.e. 1-β)
  - A power set at 80% means that we have 20% of chance to have a failed study for a compound that reach the predefined target
    - No more chance later on for the compound
    - Are we ready to take that risk ?

- Type I error (α): probability of falsely declaring PoC achieved
  - Type I error is "lost investment" risk
  - Typically set at 5%
  - Need to be controlled in "fast to fail" paradigm
- Type II error (β): probability of falsely declaring PoC failed
  - Type II error is "lost opportunity" risk
  - Typically set at 10%-20% $\rightarrow$ corresponds to power of 80%-90% (i.e. 1-β)
  - A power set at 80% means that we have 20% of chance to have a failed study for a compound that reach the predefined target
    - No more chance later on for the compound
    - Are we ready to take that risk ?

- Type I error (α): probability of falsely declaring PoC achieved
  - Type I error is "lost investment" risk
  - Typically set at 5%
  - Need to be controlled in "fast to fail" paradigm
- Type II error (β): probability of falsely declaring PoC failed
  - Type II error is "lost opportunity" risk
  - Typically set at 10%-20% $\rightarrow$ corresponds to power of 80%-90% (i.e. 1-β)
  - A power set at 80% means that we have 20% of chance to have a failed study for a compound that reach the predefined target
    - No more chance later on for the compound
    - Are we ready to take that risk ?

- Type I error (α): probability of falsely declaring PoC achieved
  - Type I error is "lost investment" risk
  - Typically set at 5%
  - Need to be controlled in "fast to fail" paradigm
- Type II error (β): probability of falsely declaring PoC failed
  - Type II error is "lost opportunity" risk
  - Typically set at 10%-20% $\rightarrow$ corresponds to power of 80%-90% (i.e. 1-β)
  - A power set at 80% means that we have 20% of chance to have a failed study for a compound that reach the predefined target
    - No more chance later on for the compound
    - Are we ready to take that risk ?

- Type I error (α): probability of falsely declaring PoC achieved
  - Type I error is "lost investment" risk
  - Typically set at 5%
  - Need to be controlled in "fast to fail" paradigm
- Type II error (β): probability of falsely declaring PoC failed
  - Type II error is "lost opportunity" risk
  - Typically set at 10%-20% $\rightarrow$ corresponds to power of 80%-90% (i.e. 1-β)
  - A power set at 80% means that we have 20% of chance to have a failed study for a compound that reach the predefined target
    - No more chance later on for the compound
    - Are we ready to take that risk ?

From a global perspective...

- Type I error
  - Next stage cost versus the risk of missing a true opportunity (of revenue on the market)

  $\Rightarrow$ We can relax on α (5% one-sided) or even 10% ?
- Type II error
  - How many compounds do we have as backups in the pipeline ?
  - How many time will we have a Go/NoGo milestone in the development of a given compound?

  $\Rightarrow$ use a portfolio-based approach to optimize PoC[1]

  $\Rightarrow$ we can relax at 80% if we have several backups
- Less regulatory concerns in learning phase
- Under the constraint of a maximal investment (N=100 cap ?)

[1] Mallinckrodt et al. 2012.

From a global perspective...

- Type I error
  - Next stage cost versus the risk of missing a true opportunity (of revenue on the market)

  $\Rightarrow$ We can relax on α (5% one-sided) or even 10% ?
- Type II error
  - How many compounds do we have as backups in the pipeline ?
  - How many time will we have a Go/NoGo milestone in the development of a given compound?

  $\Rightarrow$ use a portfolio-based approach to optimize PoC[1]

  $\Rightarrow$ we can relax at 80% if we have several backups
- Less regulatory concerns in learning phase
- Under the constraint of a maximal investment (N=100 cap ?)

[1] Mallinckrodt et al. 2012.

From a global perspective...

- Type I error
  - Next stage cost versus the risk of missing a true opportunity (of revenue on the market)

  $\Rightarrow$ We can relax on α (5% one-sided) or even 10% ?
- Type II error
  - How many compounds do we have as backups in the pipeline ?
  - How many time will we have a Go/NoGo milestone in the development of a given compound?

  $\Rightarrow$ use a portfolio-based approach to optimize PoC[1]

  $\Rightarrow$ we can relax at 80% if we have several backups
- Less regulatory concerns in learning phase
- Under the constraint of a maximal investment (N=100 cap ?)

[1] Mallinckrodt et al. 2012.

From a global perspective...

- Type I error
  - Next stage cost versus the risk of missing a true opportunity (of revenue on the market)

  $\Rightarrow$ We can relax on α (5% one-sided) or even 10% ?
- Type II error
  - How many compounds do we have as backups in the pipeline ?
  - How many time will we have a Go/NoGo milestone in the development of a given compound?

  $\Rightarrow$ use a portfolio-based approach to optimize PoC[1]

  $\Rightarrow$ we can relax at 80% if we have several backups
- Less regulatory concerns in learning phase
- Under the constraint of a maximal investment (N=100 cap ?)

[1] Mallinckrodt et al. 2012.

From a global perspective...

- Type I error
  - Next stage cost versus the risk of missing a true opportunity (of revenue on the market)

  $\Rightarrow$ We can relax on α (5% one-sided) or even 10% ?
- Type II error
  - How many compounds do we have as backups in the pipeline ?
  - How many time will we have a Go/NoGo milestone in the development of a given compound?

  $\Rightarrow$ use a portfolio-based approach to optimize PoC[1]

  $\Rightarrow$ we can relax at 80% if we have several backups
- Less regulatory concerns in learning phase
- Under the constraint of a maximal investment (N=100 cap ?)

[1] Mallinckrodt et al. 2012.

From a global perspective...

- Type I error
  - Next stage cost versus the risk of missing a true opportunity (of revenue on the market)

  $\Rightarrow$ We can relax on α (5% one-sided) or even 10% ?
- Type II error
  - How many compounds do we have as backups in the pipeline ?
  - How many time will we have a Go/NoGo milestone in the development of a given compound?

  $\Rightarrow$ use a portfolio-based approach to optimize PoC[1]

  $\Rightarrow$ we can relax at 80% if we have several backups
- Less regulatory concerns in learning phase
- Under the constraint of a maximal investment (N=100 cap ?)

[1] Mallinckrodt et al. 2012.

From a global perspective...

- Type I error
  - Next stage cost versus the risk of missing a true opportunity (of revenue on the market)

  $\Rightarrow$ We can relax on α (5% one-sided) or even 10% ?
- Type II error
  - How many compounds do we have as backups in the pipeline ?
  - How many time will we have a Go/NoGo milestone in the development of a given compound?

  $\Rightarrow$ use a portfolio-based approach to optimize PoC[1]

  $\Rightarrow$ we can relax at 80% if we have several backups
- Less regulatory concerns in learning phase
- Under the constraint of a maximal investment (N=100 cap ?)

[1] Mallinckrodt et al. 2012.

From a global perspective...

- Type I error
  - Next stage cost versus the risk of missing a true opportunity (of revenue on the market)
  - $\Rightarrow$ We can relax on α (5% one-sided) or even 10% ?
- Type II error
  - How many compounds do we have as backups in the pipeline ?
  - How many time will we have a Go/NoGo milestone in the development of a given compound?
  - $\Rightarrow$ use a portfolio-based approach to optimize PoC[1]
  - $\Rightarrow$ we can relax at 80% if we have several backups
- Less regulatory concerns in learning phase
- Under the constraint of a maximal investment (N=100 cap ?)

[1] Mallinckrodt et al. 2012.

From a global perspective...

- Type I error
  - Next stage cost versus the risk of missing a true opportunity (of revenue on the market)

  $\Rightarrow$ We can relax on α (5% one-sided) or even 10% ?
- Type II error
  - How many compounds do we have as backups in the pipeline ?
  - How many time will we have a Go/NoGo milestone in the development of a given compound?

  $\Rightarrow$ use a portfolio-based approach to optimize PoC[1]

  $\Rightarrow$ we can relax at 80% if we have several backups
- Less regulatory concerns in learning phase
- Under the constraint of a maximal investment (N=100 cap ?)

[1] Mallinckrodt et al. 2012.

From a global perspective...

- Type I error
  - Next stage cost versus the risk of missing a true opportunity (of revenue on the market)

  $\Rightarrow$ We can relax on α (5% one-sided) or even 10% ?
- Type II error
  - How many compounds do we have as backups in the pipeline ?
  - How many time will we have a Go/NoGo milestone in the development of a given compound?

  $\Rightarrow$ use a portfolio-based approach to optimize PoC[1]

  $\Rightarrow$ we can relax at 80% if we have several backups
- Less regulatory concerns in learning phase
- Under the constraint of a maximal investment (N=100 cap ?)

[1] Mallinckrodt et al. 2012.

A trial case

- After a single and repeated ascending FIM dose
  - We know the maximal tolerated dose (MTD)
- The PoC
  - 2 arms
  - Active vs comparator
  - Relevant effect: $\geq 2 mmHg$
  - SD $= 3.5$
  - Power $= 80\%$
  - Alpha 5% one-sided

A trial case

- After a single and repeated ascending FIM dose
  - We know the maximal tolerated dose (MTD)
- The PoC
  - 2 arms
  - Active vs comparator
  - Relevant effect: $\geq 2mmHg$
  - SD $= 3.5$
  - Power $= 80\%$
  - Alpha 5% one-sided

Classical power

- Once type I error, variability and sample size are choosen, we get the power conditionally on a specified treatment effect

The classical power does not quantify the PoS

It is a conditional PoS on the unknown true trt value, $P(Success|\theta)$

- But the probability of observing such effect ...
- One can sweep across various assumptions on the effect and variability values to assess how robust the power is (for N=39 per group)

θ	SD=3	SD=3.25	SD=3.5	SD=3.75
1.5	70%	64%	59%	54%
2	89%	85%	80%	75%
2.5	97%	95%	93%	89%

- Or integrate over possible θ values

Classical power

- Once type I error, variability and sample size are choosen, we get the power conditionally on a specified treatment effect

The classical power does not quantify the PoS

It is a conditional PoS on the unknown true trt value, $P(Success|\theta)$

- But the probability of observing such effect ...
- One can sweep across various assumptions on the effect and variability values to assess how robust the power is (for N=39 per group)

θ	SD=3	SD=3.25	SD=3.5	SD=3.75
1.5	70%	64%	59%	54%
2	89%	85%	80%	75%
2.5	97%	95%	93%	89%

- Or integrate over possible θ values

Classical power

- Once type I error, variability and sample size are choosen, we get the power conditionally on a specified treatment effect

The classical power does not quantify the PoS

It is a conditional PoS on the unknown true trt value, $P(Success|\theta)$

- But the probability of observing such effect ...
- One can sweep across various assumptions on the effect and variability values to assess how robust the power is (for N=39 per group)

θ	SD=3	SD=3.25	SD=3.5	SD=3.75
1.5	70%	64%	59%	54%
2	89%	85%	80%	75%
2.5	97%	95%	93%	89%

- Or integrate over possible θ values

Classical power

- Once type I error, variability and sample size are choosen, we get the power conditionally on a specified treatment effect

The classical power does not quantify the PoS

It is a conditional PoS on the unknown true trt value, $P(Success|\theta)$

- But the probability of observing such effect ...
- One can sweep across various assumptions on the effect and variability values to assess how robust the power is (for N=39 per group)

θ	SD=3	SD=3.25	SD=3.5	SD=3.75
1.5	70%	64%	59%	54%
2	89%	85%	80%	75%
2.5	97%	95%	93%	89%

- Or integrate over possible θ values

Classical power

- Once type I error, variability and sample size are choosen, we get the power conditionally on a specified treatment effect

The classical power does not quantify the PoS

It is a conditional PoS on the unknown true trt value, $P(Success|\theta)$

- But the probability of observing such effect ...
- One can sweep across various assumptions on the effect and variability values to assess how robust the power is (for N=39 per group)

θ	SD=3	SD=3.25	SD=3.5	SD=3.75
1.5	70%	64%	59%	54%
2	89%	85%	80%	75%
2.5	97%	95%	93%	89%

- Or integrate over possible θ values

Assurance

- By using the unconditional PoS, we take into account the uncertainty around the true underlying treatment effect
  - Define a prior distribution for θ
  - Build a weighted average of the power function with weights from prior distribution

Assurance (Ohagan et al. 2005)

The assurance is the expected power (with respect to the prior)
$P(Success) = E(\pi(\theta)) = \int_\theta P(Success, \theta) d\theta = \int_\theta P(Success|\theta).P(\theta).d\theta$
where $\pi(\theta)$ is the power function

- Quantifies the ability of the trial to achieve a success based on available evidence

Assurance

- By using the unconditional PoS, we take into account the uncertainty around the true underlying treatment effect
  - Define a prior distribution for θ
  - Build a weighted average of the power function with weights from prior distribution

Assurance (Ohagan et al. 2005)

The assurance is the expected power (with respect to the prior)
$P(Success) = E(\pi(\theta)) = \int_\theta P(Success, \theta) d\theta = \int_\theta P(Success|\theta).P(\theta).d\theta$
where $\pi(\theta)$ is the power function

- Quantifies the ability of the trial to achieve a success based on available evidence

Assurance

- By using the unconditional PoS, we take into account the uncertainty around the true underlying treatment effect
  - Define a prior distribution for θ
  - Build a weighted average of the power function with weights from prior distribution

Assurance (Ohagan et al. 2005)

The assurance is the expected power (with respect to the prior)
$P(Success) = E(\pi(\theta)) = \int_\theta P(Success, \theta) d\theta = \int_\theta P(Success|\theta).P(\theta).d\theta$
where $\pi(\theta)$ is the power function

- Quantifies the ability of the trial to achieve a success based on available evidence

Assurance

- By using the unconditional PoS, we take into account the uncertainty around the true underlying treatment effect
  - Define a prior distribution for θ
  - Build a weighted average of the power function with weights from prior distribution

Assurance (Ohagan et al. 2005)

The assurance is the expected power (with respect to the prior)
$P(Success) = E(\pi(\theta)) = \int_\theta P(Success, \theta) d\theta = \int_\theta P(Success|\theta).P(\theta).d\theta$
where $\pi(\theta)$ is the power function

- Quantifies the ability of the trial to achieve a success based on available evidence

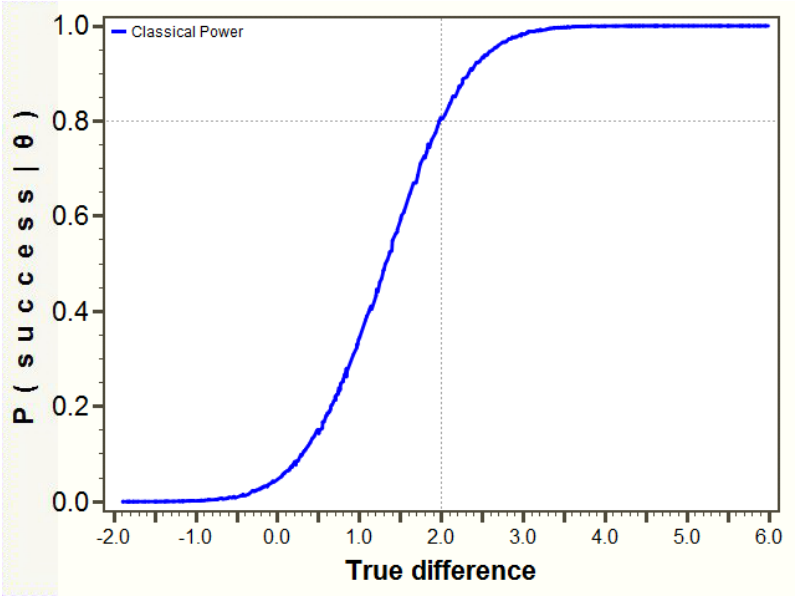
Assurance

- By using the unconditional PoS, we take into account the uncertainty around the true underlying treatment effect
  - Define a prior distribution for θ
  - Build a weighted average of the power function with weights from prior distribution

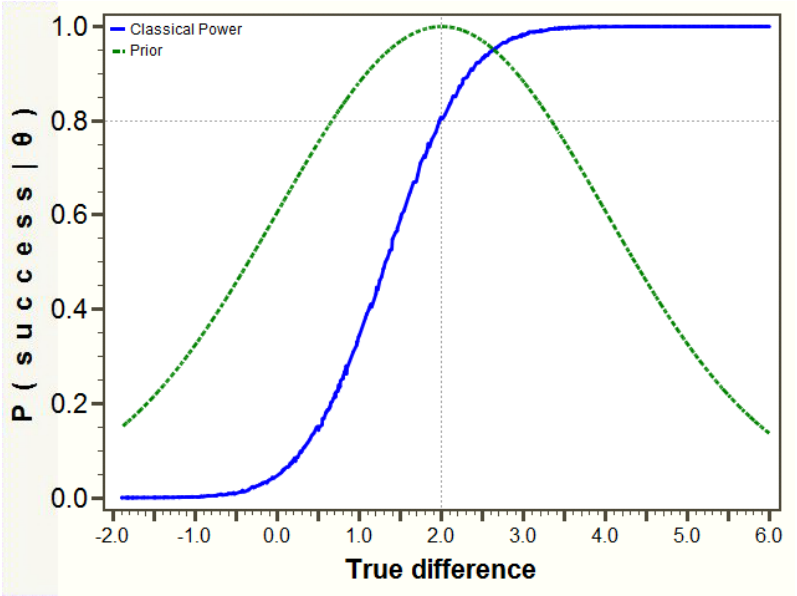
Assurance (Ohagan et al. 2005)

The assurance is the expected power (with respect to the prior)
$P(Success) = E(\pi(\theta)) = \int_\theta P(Success, \theta) d\theta = \int_\theta P(Success|\theta).P(\theta).d\theta$
where $\pi(\theta)$ is the power function

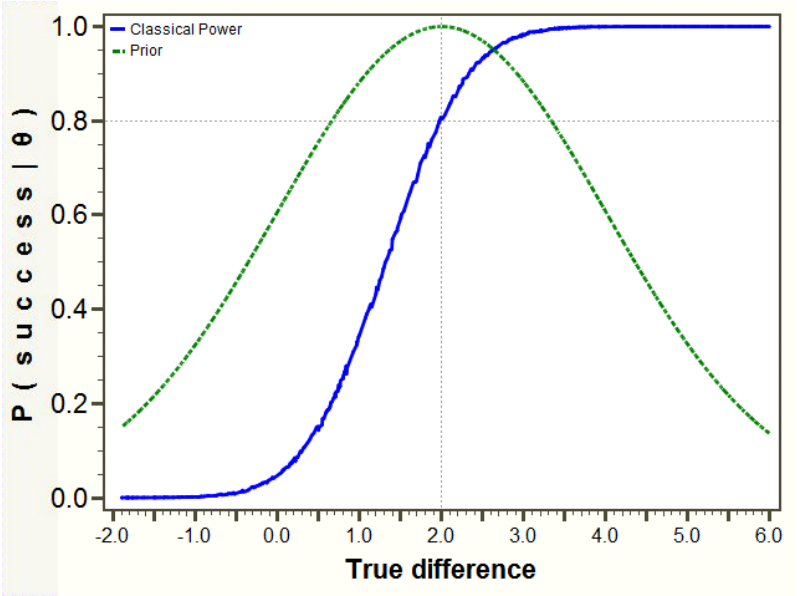
- Quantifies the ability of the trial to achieve a success based on available evidence



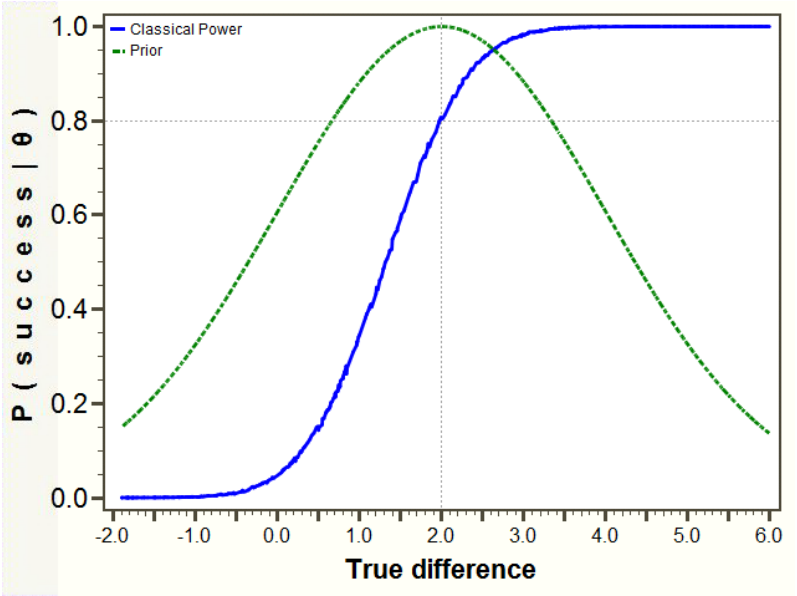
- Assurance = 63%
- More realistic assessment of the chance of success of a drug than classical power
  - From the joint prior distribution, integration over the space of θ and σ^2 is possible
  - A bayesian sample size is also an option along with an expected bayesian power



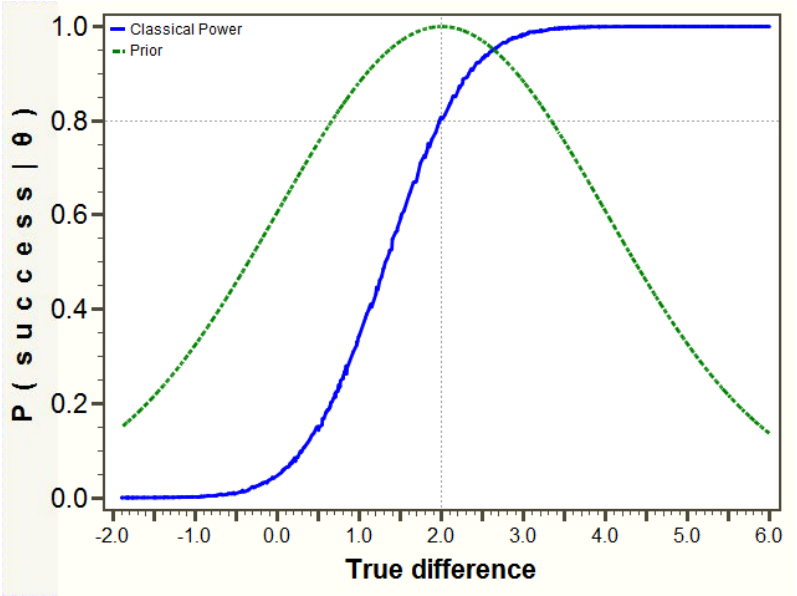
- Assurance = 63%
- More realistic assessment of the chance of success of a drug than classical power
  - From the joint prior distribution, integration over the space of θ and σ^2 is possible
  - A bayesian sample size is also an option along with an expected bayesian power



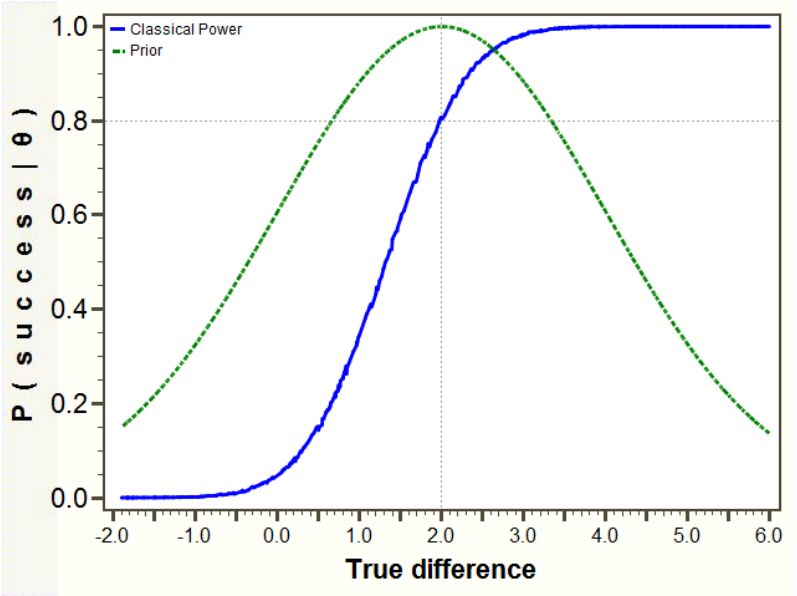
- Assurance = 63%
- More realistic assessment of the chance of success of a drug than classical power
  - From the joint prior distribution, integration over the space of θ and σ^2 is possible
  - A bayesian sample size is also an option along with an expected bayesian power



- Assurance = 63%
- More realistic assessment of the chance of success of a drug than classical power
  - From the joint prior distribution, integration over the space of θ and σ^2 is possible
  - A bayesian sample size is also an option along with an expected bayesian power



- Assurance = 63%
- More realistic assessment of the chance of success of a drug than classical power
  - From the joint prior distribution, integration over the space of θ and σ^2 is possible
  - A bayesian sample size is also an option along with an expected bayesian power



- Assurance = 63%
- More realistic assessment of the chance of success of a drug than classical power
  - From the joint prior distribution, integration over the space of θ and σ^2 is possible
  - A bayesian sample size is also an option along with an expected bayesian power

Moving away from p-value to full posterior distribution

- P-value is not necessary always the answer ...
  - Clinical significant $\neq$ Statistically significant
- Back to the example with a relevant $\theta \geq 2mmHG$

$P(\theta \geq 2) = 5\%$ p-value=5%

$P(\theta \geq 2) = 69\%$ p-value=10%

- Consider the posterior probability : $P(\theta|data) \geq threshold$

Moving away from p-value to full posterior distribution

- P-value is not necessary always the answer ...
  - Clinical significant $\neq$ Statistically significant
- Back to the example with a relevant $\theta \geq 2mmHG$

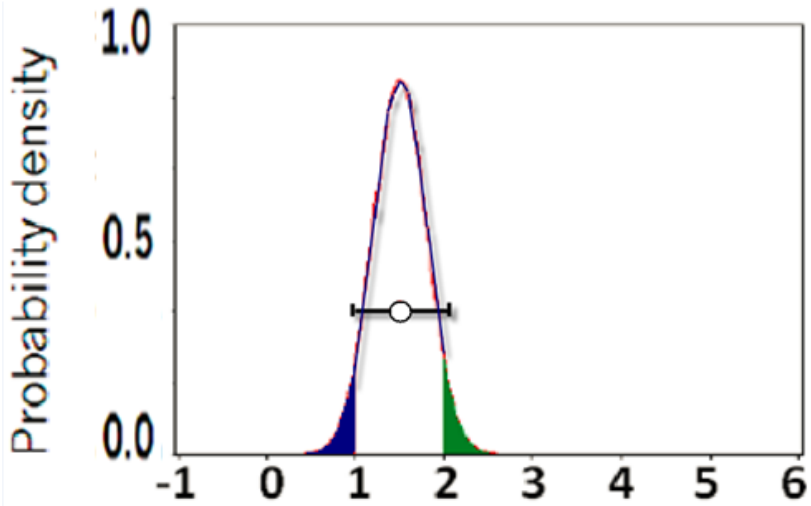
$P(\theta \geq 2) = 5\%$ p-value=5%

$P(\theta \geq 2) = 69\%$ p-value=10%

- Consider the posterior probability : $P(\theta|data) \geq threshold$

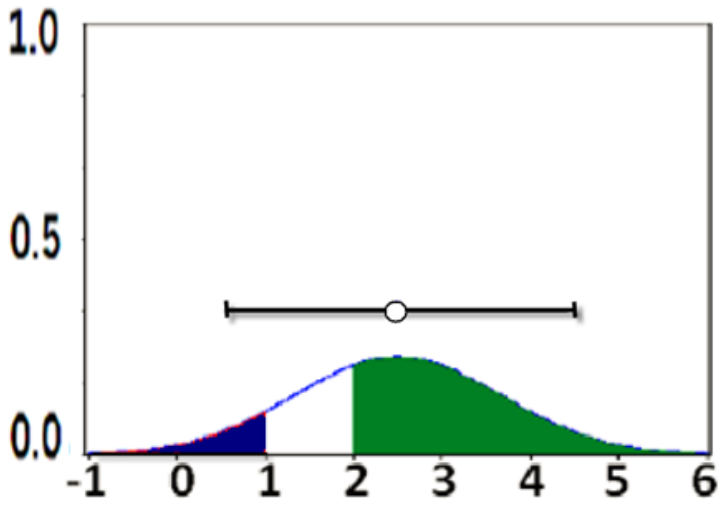
Moving away from p-value to full posterior distribution

- P-value is not necessary always the answer ...
  - Clinical significant $\neq$ Statistically significant
- Back to the example with a relevant $\theta \geq 2mmHG$



$P(\theta \geq 2) = 5\%$

p-value=5%



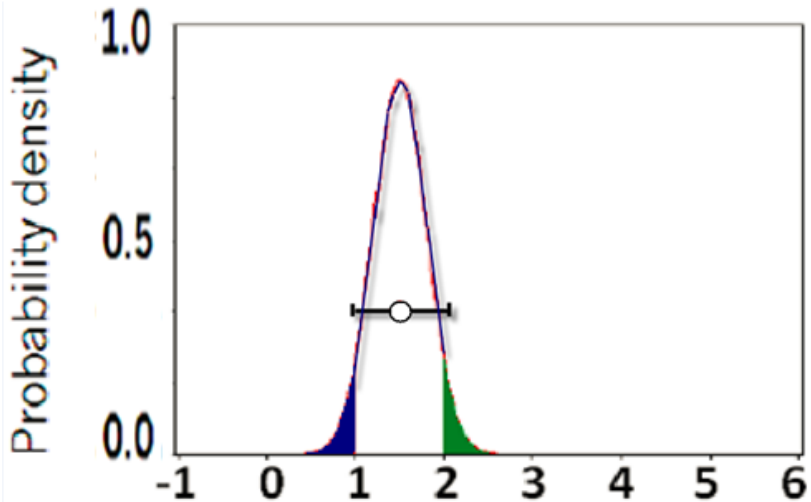
$P(\theta \geq 2) = 69\%$

p-value=10%

- Consider the posterior probability : $P(\theta|data) \geq threshold$

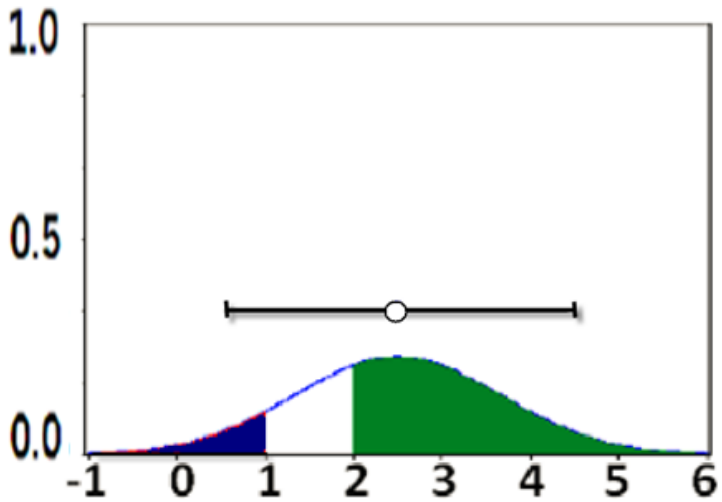
Moving away from p-value to full posterior distribution

- P-value is not necessary always the answer ...
  - Clinical significant $\neq$ Statistically significant
- Back to the example with a relevant $\theta \geq 2mmHG$



$P(\theta \geq 2) = 5\%$

p-value=5%



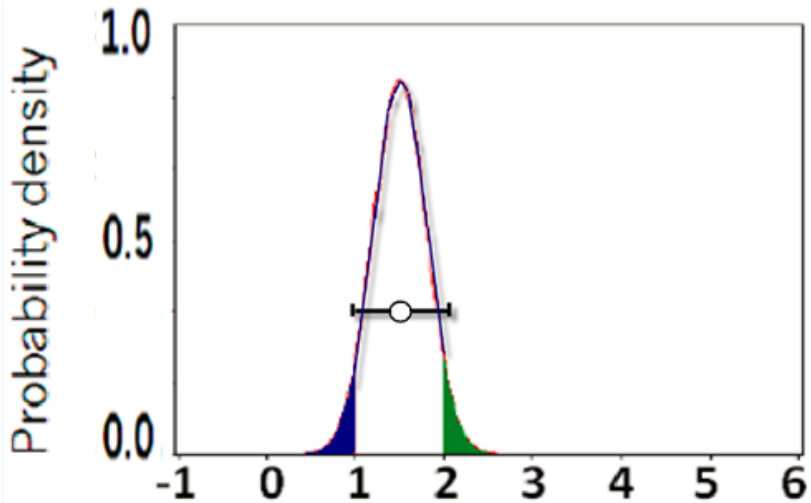
$P(\theta \geq 2) = 69\%$

p-value=10%

- Consider the posterior probability : $P(\theta|data) \geq threshold$

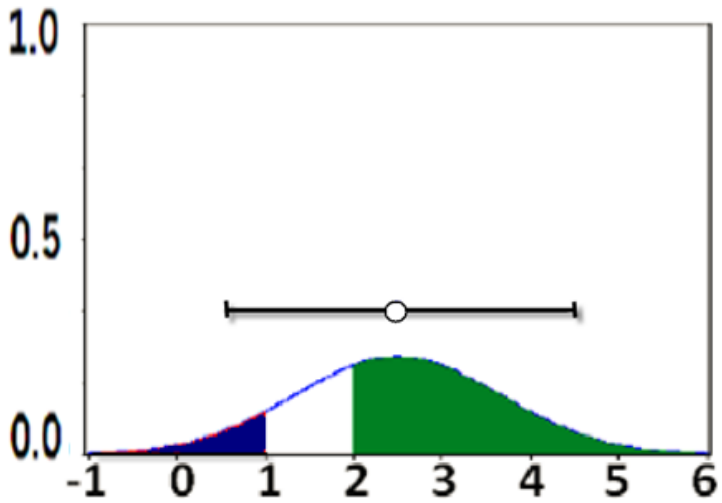
Moving away from p-value to full posterior distribution

- P-value is not necessary always the answer ...
  - Clinical significant $\neq$ Statistically significant
- Back to the example with a relevant $\theta \geq 2mmHG$



$P(\theta \geq 2) = 5\%$

p-value=5%



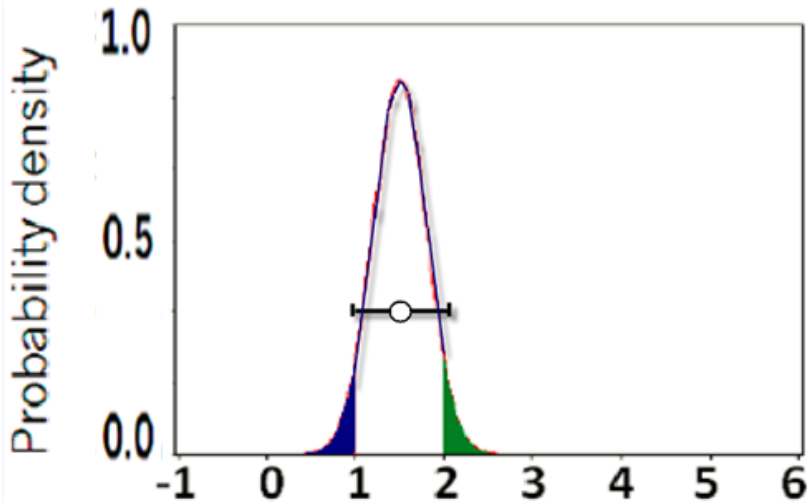
$P(\theta \geq 2) = 69\%$

p-value=10%

- Consider the posterior probability : $P(\theta|data) \geq threshold$

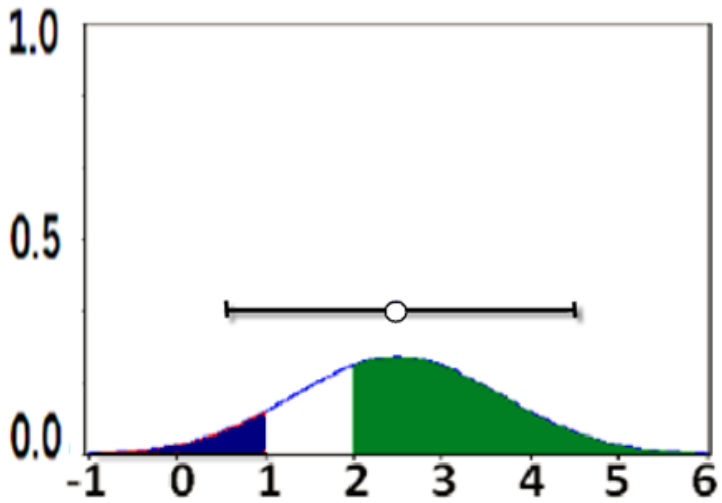
Moving away from p-value to full posterior distribution

- P-value is not necessary always the answer ...
  - Clinical significant $\neq$ Statistically significant
- Back to the example with a relevant $\theta \geq 2mmHG$



$P(\theta \geq 2) = 5\%$

p-value=5%



$P(\theta \geq 2) = 69\%$

p-value=10%

- Consider the posterior probability : $P(\theta|data) \geq threshold$

- When we start a trial, do we really know nothing ?
  - Set-up trial with prior knowledge from past studies or bibliography
    - Expected effect (at least on placebo)
    - Expected variance

  $\Rightarrow$ used for sample size
  - But often not used for the analysis !
- We know *a priori* something
  - The trial will update this knowledge
  - To take a decision

$\Rightarrow$ Let's use bayesian statistics !

- When we start a trial, do we really know nothing ?
  - Set-up trial with prior knowledge from past studies or bibliography
    - Expected effect (at least on placebo)
    - Expected variance

  $\Rightarrow$ used for sample size
  - But often not used for the analysis !
- We know *a priori* something
  - The trial will update this knowledge
  - To take a decision

$\Rightarrow$ Let's use bayesian statistics !

- When we start a trial, do we really know nothing ?
  - Set-up trial with prior knowledge from past studies or bibliography
    - Expected effect (at least on placebo)
    - Expected variance

  $\Rightarrow$ used for sample size
  - But often not used for the analysis !
- We know *a priori* something
  - The trial will update this knowledge
  - To take a decision

$\Rightarrow$ Let's use bayesian statistics !

- When we start a trial, do we really know nothing ?
  - Set-up trial with prior knowledge from past studies or bibliography
    - Expected effect (at least on placebo)
    - Expected variance

  $\Rightarrow$ used for sample size
  - But often not used for the analysis !
- We know *a priori* something
  - The trial will update this knowledge
  - To take a decision

$\Rightarrow$ Let's use bayesian statistics !

- When we start a trial, do we really know nothing ?
  - Set-up trial with prior knowledge from past studies or bibliography
    - Expected effect (at least on placebo)
    - Expected variance

  $\Rightarrow$ used for sample size
  - But often not used for the analysis !
- We know *a priori* something
  - The trial will update this knowledge
  - To take a decision

$\Rightarrow$ Let's use bayesian statistics !

Prior distribution $p(\theta)$ (previous study, litterature,...)

Likelihood $p(y|\theta)$ (from experiment)

Posterior distribution $p(\theta|y)$ (combination)

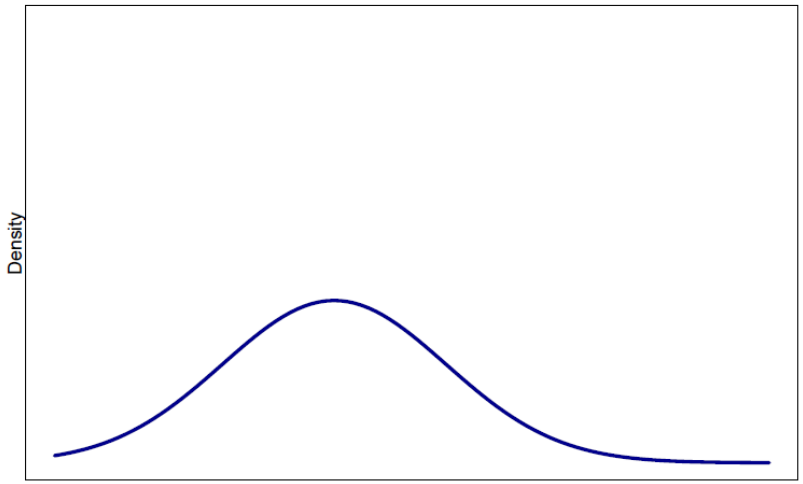
- Posterior distribution for θ is a weighted compromise between prior and likelihood
- The problem is to get the posterior...

Historical Data x Observed Data $\propto$ All Data

Prior distribution $p(\theta)$ (previous study, litterature,...)

Likelihood $p(y|\theta)$ (from experiment)

Posterior distribution $p(\theta|y)$ (combination)



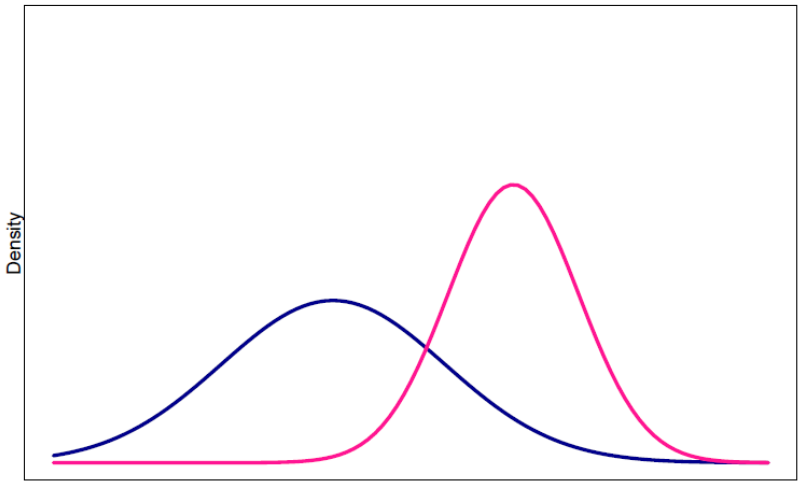
- Posterior distribution for θ is a weighted compromise between prior and likelihood
- The problem is to get the posterior...

Historical Data x Observed Data $\propto$ All Data

Prior distribution $p(\theta)$ (previous study, litterature,...)

Likelihood $p(y|\theta)$ (from experiment)

Posterior distribution $p(\theta|y)$ (combination)



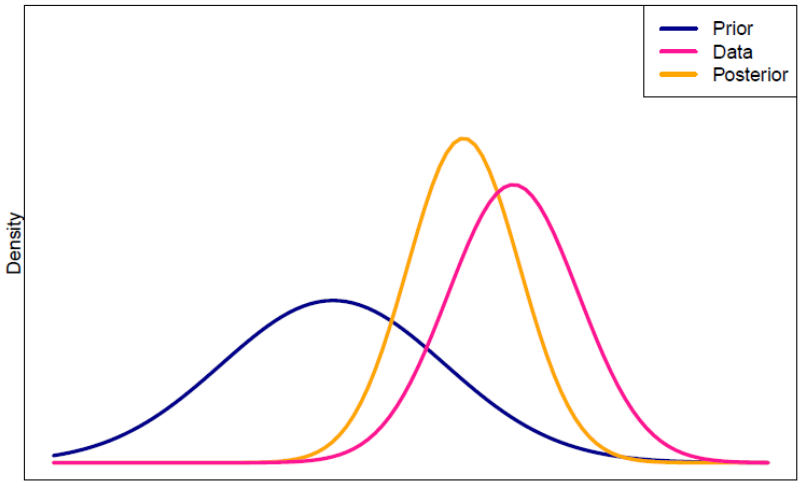
- Posterior distribution for θ is a weighted compromise between prior and likelihood
- The problem is to get the posterior...

Historical Data x Observed Data $\propto$ All Data

Prior distribution $p(\theta)$ (previous study, litterature,...)

Likelihood $p(y|\theta)$ (from experiment)

Posterior distribution $p(\theta|y)$ (combination)



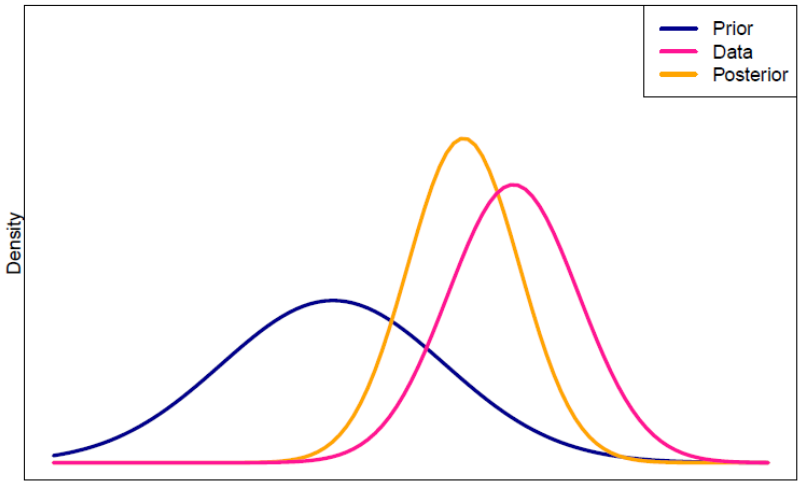
- Posterior distribution for θ is a weighted compromise between prior and likelihood
- The problem is to get the posterior...

Historical Data x Observed Data $\propto$ All Data

Prior distribution $p(\theta)$ (previous study, litterature,...)

Likelihood $p(y|\theta)$ (from experiment)

Posterior distribution $p(\theta|y)$ (combination)



- Posterior distribution for θ is a weighted compromise between prior and likelihood
- The problem is to get the posterior...

Why bayesian ?

Natural way to update our knowledge in a model framework

- Prior distribution includes all prior knowledge before starting
  - Often, little (or nothing) is known: Non-informative or minimally informative prior
  - Adequate prior to be chosen : should vanish rapidly when data accumulate

Informative prior [Plbo] Non-informative (flat) prior [Active]

Exact inference

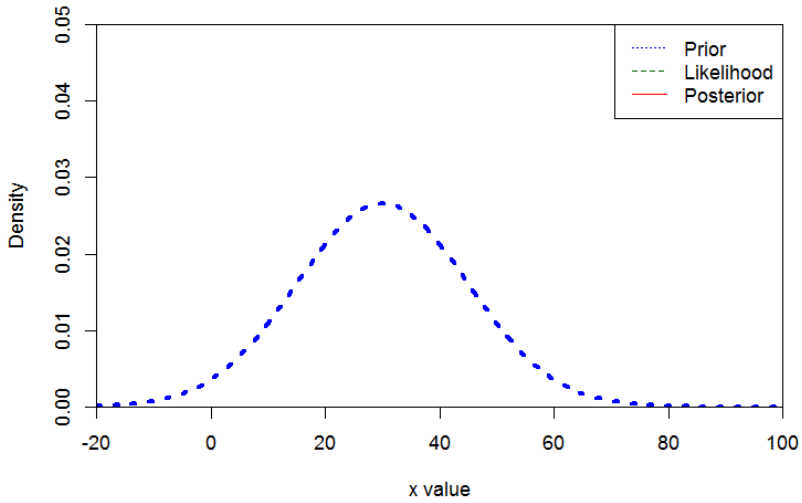
Proper interpretation of credibility interval for efficient decision making

Why bayesian ?

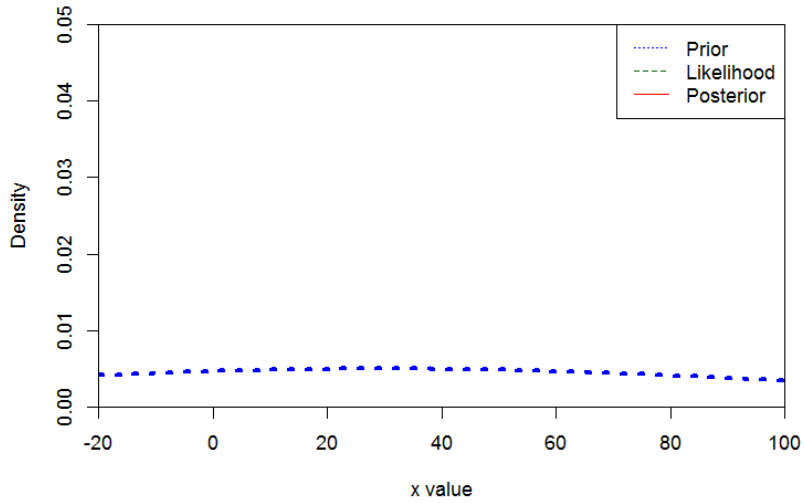
Natural way to update our knowledge in a model framework

- Prior distribution includes all prior knowledge before starting
  - Often, little (or nothing) is known: Non-informative or minimally informative prior
  - Adequate prior to be chosen : should vanish rapidly when data accumulate

Informative prior [Plbo]



Non-informative (flat) prior [Active]

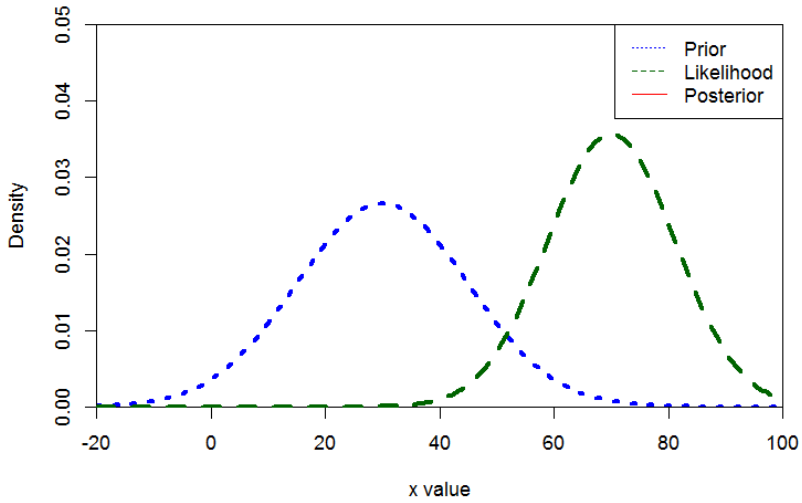


Why bayesian ?

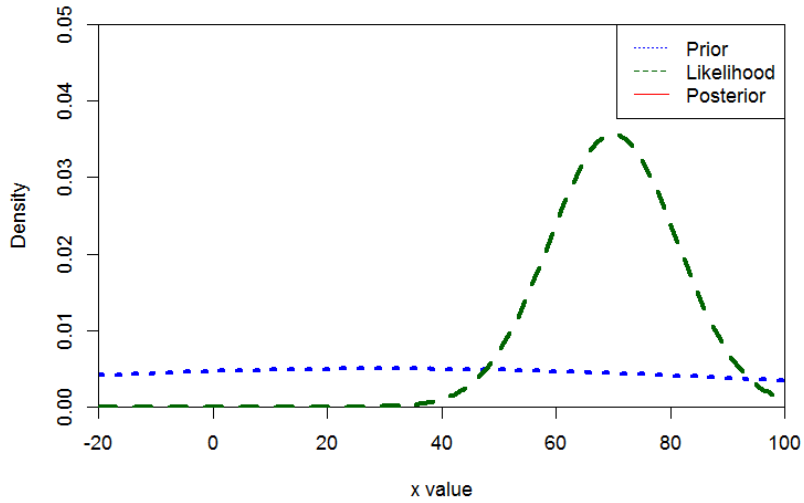
Natural way to update our knowledge in a model framework

- Prior distribution includes all prior knowledge before starting
  - Often, little (or nothing) is known: Non-informative or minimally informative prior
  - Adequate prior to be chosen : should vanish rapidly when data accumulate

Informative prior [Plbo]



Non-informative (flat) prior [Active]

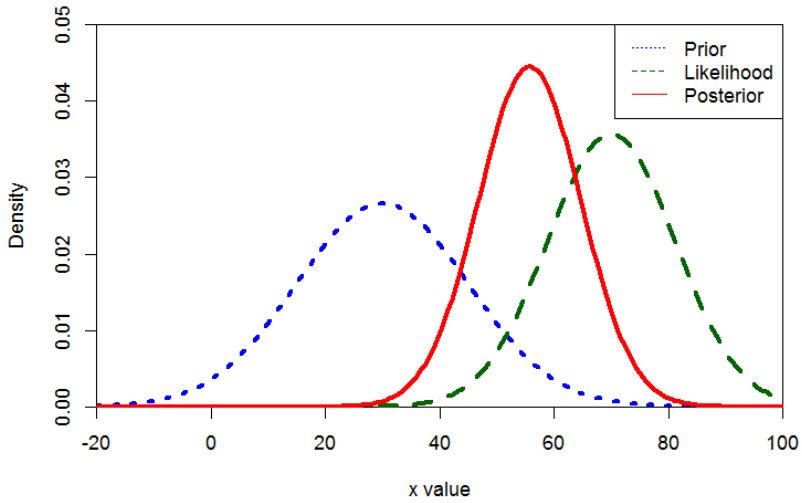


Why bayesian ?

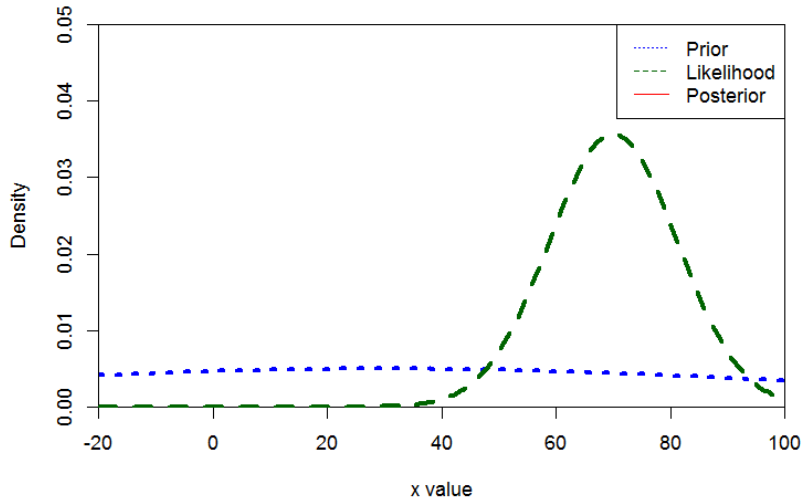
Natural way to update our knowledge in a model framework

- Prior distribution includes all prior knowledge before starting
  - Often, little (or nothing) is known: Non-informative or minimally informative prior
  - Adequate prior to be chosen : should vanish rapidly when data accumulate

Informative prior [Plbo]



Non-informative (flat) prior [Active]

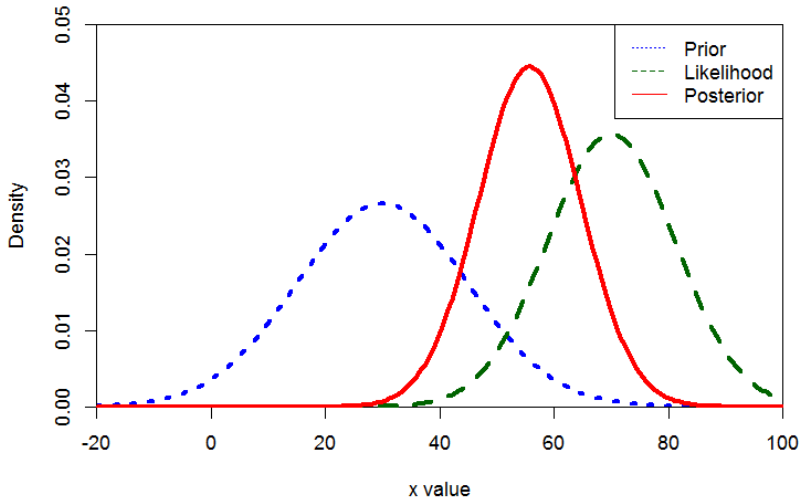


Why bayesian ?

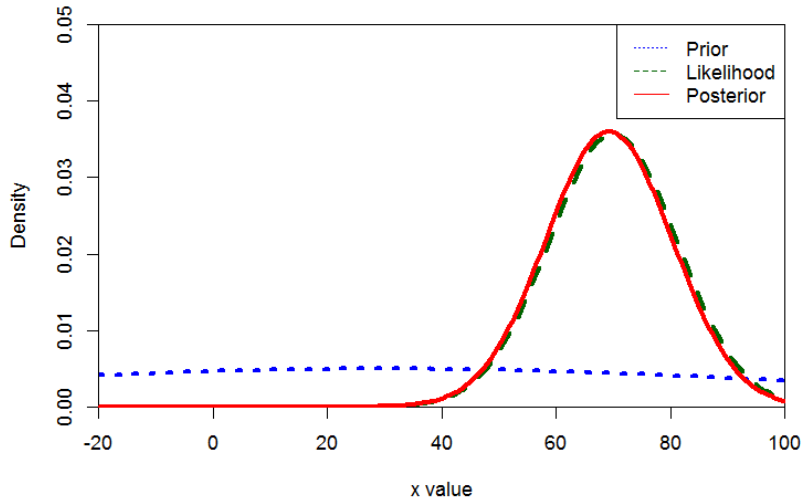
Natural way to update our knowledge in a model framework

- Prior distribution includes all prior knowledge before starting
  - Often, little (or nothing) is known: Non-informative or minimally informative prior
  - Adequate prior to be chosen : should vanish rapidly when data accumulate

Informative prior [Plbo]



Non-informative (flat) prior [Active]



Why bayesian ?

Natural way to update our knowledge in a model framework

- Prior distribution includes all prior knowledge before starting
  - Often, little (or nothing) is known: Non-informative or minimally informative prior
  - Adequate prior to be chosen : should vanish rapidly when data accumulate

Informative prior [Plbo] Non-informative (flat) prior [Active]

Exact inference

Proper interpretation of credibility interval for efficient decision making

Why bayesian ?

Natural way to update our knowledge in a model framework

- Prior distribution includes all prior knowledge before starting
  - Often, little (or nothing) is known: Non-informative or minimally informative prior
  - Adequate prior to be chosen : should vanish rapidly when data accumulate

Informative prior [Plbo] Non-informative (flat) prior [Active]

Exact inference

Proper interpretation of credibility interval for efficient decision making

- 2 arms PoC, $\alpha = 5\%$ one-sided, $\theta = 2$, SD $= 3.5$
- Classical approach
  - Balanced N= 39*2=78 : 80% power
  - Unbalanced N= 52+26 : 78% power
- Bayesian approach integrating informative prior on comparator
- Response on Active : a non-informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0e}, 10^8)$, with huge variability
- Response on Comparator : an informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0c}, \frac{1}{Q_{0c}*\nu})$, Q_{0c} denotes the amount of information (Nber subjects)
- The prior distribution for the precision ν is chosen as a gamma with parameters $(a_0; b_0)$, so that the mean $(\frac{a_0}{b_0} = \frac{1}{SD^2} = \frac{4}{49})$
  For (a_0, b_0) the 3 scenarii are retained (20,245) , (40,490), (200,2450) resulting in a 95% credibility interval for SD of [2.87; 4.48], [3.03; 4.14] and [3.27; 3.76] reflecting the uncertainty we have around SD=3.5.

- 2 arms PoC, $\alpha = 5\%$ one-sided, $\theta = 2$, SD $= 3.5$
- Classical approach
  - Balanced N= 39*2=78 : 80% power
  - Unbalanced N= 52+26 : 78% power
- Bayesian approach integrating informative prior on comparator
- Response on Active : a non-informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0e}, 10^8)$, with huge variability
- Response on Comparator : an informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0c}, \frac{1}{Q_{0c}*\nu})$, Q_{0c} denotes the amount of information (Nber subjects)
- The prior distribution for the precision ν is chosen as a gamma with parameters $(a_0; b_0)$, so that the mean $(\frac{a_0}{b_0} = \frac{1}{SD^2} = \frac{4}{49})$

  For (a_0, b_0) the 3 scenarii are retained (20,245) , (40,490), (200,2450) resulting in a 95% credibility interval for SD of [2.87; 4.48], [3.03; 4.14] and [3.27; 3.76] reflecting the uncertainty we have around SD=3.5.

- 2 arms PoC, $\alpha = 5\%$ one-sided, $\theta = 2$, SD $= 3.5$
- Classical approach
  - Balanced N= 39*2=78 : 80% power
  - Unbalanced N= 52+26 : 78% power
- Bayesian approach integrating informative prior on comparator
- Response on Active : a non-informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0e}, 10^8)$, with huge variability
- Response on Comparator : an informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0c}, \frac{1}{Q_{0c}*\nu})$, Q_{0c} denotes the amount of information (Nber subjects)
- The prior distribution for the precision ν is chosen as a gamma with parameters $(a_0; b_0)$, so that the mean $(\frac{a_0}{b_0} = \frac{1}{SD^2} = \frac{4}{49})$

  For (a_0, b_0) the 3 scenarii are retained (20,245) , (40,490), (200,2450) resulting in a 95% credibility interval for SD of [2.87; 4.48], [3.03; 4.14] and [3.27; 3.76] reflecting the uncertainty we have around SD=3.5.

- 2 arms PoC, $\alpha = 5\%$ one-sided, $\theta = 2$, SD $= 3.5$
- Classical approach
  - Balanced N= 39*2=78 : 80% power
  - Unbalanced N= 52+26 : 78% power
- Bayesian approach integrating informative prior on comparator
- Response on Active : a non-informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0e}, 10^8)$, with huge variability
- Response on Comparator : an informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0c}, \frac{1}{Q_{0c} * \nu})$, Q_{0c} denotes the amount of information (Nber subjects)
- The prior distribution for the precision ν is chosen as a gamma with parameters $(a_0; b_0)$, so that the mean $(\frac{a_0}{b_0} = \frac{1}{SD^2} = \frac{4}{49})$

  For (a_0, b_0) the 3 scenarii are retained (20,245) , (40,490), (200,2450) resulting in a 95% credibility interval for SD of [2.87; 4.48], [3.03; 4.14] and [3.27; 3.76] reflecting the uncertainty we have around SD=3.5.

- 2 arms PoC, $\alpha = 5\%$ one-sided, $\theta = 2$, SD $= 3.5$
- Classical approach
  - Balanced N= 39*2=78 : 80% power
  - Unbalanced N= 52+26 : 78% power
- Bayesian approach integrating informative prior on comparator
- Response on Active : a non-informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0e}, 10^8)$, with huge variability
- Response on Comparator : an informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0c}, \frac{1}{Q_{0c}*\nu})$, Q_{0c} denotes the amount of information (Nber subjects)
- The prior distribution for the precision ν is chosen as a gamma with parameters $(a_0; b_0)$, so that the mean $(\frac{a_0}{b_0} = \frac{1}{SD^2} = \frac{4}{49})$

  For (a_0, b_0) the 3 scenarii are retained (20,245) , (40,490), (200,2450) resulting in a 95% credibility interval for SD of [2.87; 4.48], [3.03; 4.14] and [3.27; 3.76] reflecting the uncertainty we have around SD=3.5.

- 2 arms PoC, $\alpha = 5\%$ one-sided, $\theta = 2$, SD $= 3.5$
- Classical approach
  - Balanced N= 39*2=78 : 80% power
  - Unbalanced N= 52+26 : 78% power
- Bayesian approach integrating informative prior on comparator
- Response on Active : a non-informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0e}, 10^8)$, with huge variability
- Response on Comparator : an informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0c}, \frac{1}{Q_{0c}*\nu})$, Q_{0c} denotes the amount of information (Nber subjects)
- The prior distribution for the precision ν is chosen as a gamma with parameters $(a_0; b_0)$, so that the mean $(\frac{a_0}{b_0} = \frac{1}{SD^2} = \frac{4}{49})$

  For (a_0, b_0) the 3 scenarii are retained (20,245) , (40,490), (200,2450) resulting in a 95% credibility interval for SD of [2.87; 4.48], [3.03; 4.14] and [3.27; 3.76] reflecting the uncertainty we have around SD=3.5.

- 2 arms PoC, $\alpha = 5\%$ one-sided, $\theta = 2$, SD $= 3.5$
- Classical approach
  - Balanced N= 39*2=78 : 80% power
  - Unbalanced N= 52+26 : 78% power
- Bayesian approach integrating informative prior on comparator
- Response on Active : a non-informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0e}, 10^8)$, with huge variability
- Response on Comparator : an informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0c}, \frac{1}{Q_{0c} * \nu})$, Q_{0c} denotes the amount of information (Nber subjects)
- The prior distribution for the precision ν is chosen as a gamma with parameters $(a_0; b_0)$, so that the mean $(\frac{a_0}{b_0} = \frac{1}{SD^2} = \frac{4}{49})$

  For (a_0, b_0) the 3 scenarii are retained (20,245) , (40,490), (200,2450) resulting in a 95% credibility interval for SD of [2.87; 4.48], [3.03; 4.14] and [3.27; 3.76] reflecting the uncertainty we have around SD=3.5.

- 2 arms PoC, $\alpha = 5\%$ one-sided, $\theta = 2$, SD $= 3.5$
- Classical approach
  - Balanced N= 39*2=78 : 80% power
  - Unbalanced N= 52+26 : 78% power
- Bayesian approach integrating informative prior on comparator
- Response on Active : a non-informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0e}, 10^8)$, with huge variability
- Response on Comparator : an informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0c}, \frac{1}{Q_{0c}*\nu})$, Q_{0c} denotes the amount of information (Nber subjects)
- The prior distribution for the precision ν is chosen as a gamma with parameters $(a_0; b_0)$, so that the mean $(\frac{a_0}{b_0} = \frac{1}{SD^2} = \frac{4}{49})$

  For (a_0, b_0) the 3 scenarii are retained (20,245) , (40,490), (200,2450) resulting in a 95% credibility interval for SD of [2.87; 4.48], [3.03; 4.14] and [3.27; 3.76] reflecting the uncertainty we have around SD=3.5.

- 2 arms PoC, $\alpha = 5\%$ one-sided, $\theta = 2$, SD $= 3.5$
- Classical approach
  - Balanced N= 39*2=78 : 80% power
  - Unbalanced N= 52+26 : 78% power
- Bayesian approach integrating informative prior on comparator
- Response on Active : a non-informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0e}, 10^8)$, with huge variability
- Response on Comparator : an informative conditional (knowing ν) normal prior $\mathcal{N}(\mu_{0c}, \frac{1}{Q_{0c}*\nu})$, Q_{0c} denotes the amount of information (Nber subjects)
- The prior distribution for the precision ν is chosen as a gamma with parameters $(a_0; b_0)$, so that the mean $(\frac{a_0}{b_0} = \frac{1}{SD^2} = \frac{4}{49})$

  For (a_0, b_0) the 3 scenarii are retained (20,245) , (40,490), (200,2450) resulting in a 95% credibility interval for SD of [2.87; 4.48], [3.03; 4.14] and [3.27; 3.76] reflecting the uncertainty we have around SD=3.5.

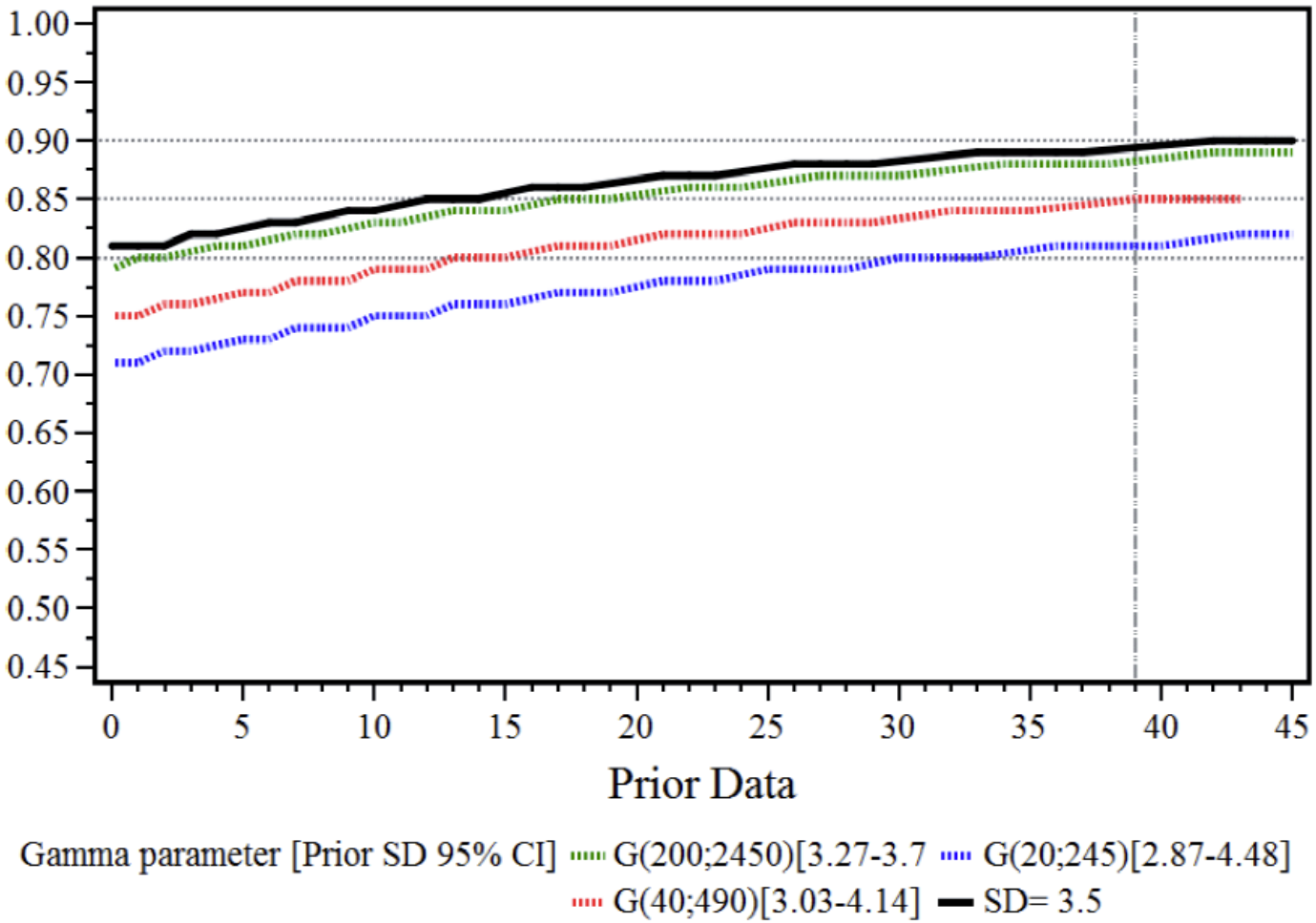


Figure : Probability $P(\tau > \tau^* | data)$ VS prior amount of information (q_{0C}) (with N=39+39; $P(\tau < 0 | data) = 0.95$; θ=2)

Table : Probability of success ($P(\tau < 0|data) = 0.95; \theta = 2$)

SD	Ratio	N_{ACT}	N_C	N_{TOT}	Q_{0c}	PoS
Fixed	**1:1**	39	39	78	-	80%
$\sim \Gamma(40; 490)$	**1:1**	39	39	78	39	85%
95%CI=[3.01-4.14]	**1:2**	52	26	78	26	85%
	1:3	57	19	76	19	80%
					38	87%

- Using bayesian methods leverage the information we have
- Balanced design : PoS increase (even with SD uncertainty)
- Unbalanced design has been proposed
  - to maximize the information on active arm
  - while maintaining good PoS
- Prior elicitation is key : well selected paper
- Many simulations done to assess robustness

Early phase is an exploratory period for learning

- Continuous update of our knowledge
- Altering study components during the course of the trial, based on accumulated data, could save time and streamline the exploration of the parameters space
- The main regulatory focus in exploratory development phases is on safety
  - much flexibility is allowed into trial features related to the assessment of efficacy, for the purpose of sponsor-internal decision-making

⇒ **Adaptive designs**[2] could be more suitable than fixed approach in order to learn and adapt more easily and quickly

[2] PhRMA Working Group, Gallo et al. 2006.

Early phase is an exploratory period for learning

- Continuous update of our knowledge
- Altering study components during the course of the trial, based on accumulated data, could save time and streamline the exploration of the parameters space
- The main regulatory focus in exploratory development phases is on safety
  - much flexibility is allowed into trial features related to the assessment of efficacy, for the purpose of sponsor-internal decision-making

$\Rightarrow$ **Adaptive designs**[2] could be more suitable than fixed approach in order to learn and adapt more easily and quickly

[2] PhRMA Working Group, Gallo et al. 2006.

- Context
  - 2 arms trial
  - Long recruitment due to :
  - Low recruitment rate (difficult patient population)
  - Large number of patients
  - Relatively short biomarker readout
- Motivation
  - Why should we wait until the end of the trial if partial data demonstrate that the trial is being unlikely to achieve its objective ?
  - Unethical

  $\Rightarrow$ **Stop for futility !! (or adaptive randomization design)**
- Which decisions ?
  1. Evidence in favor of null or alternative hypotheses at the interim ?
  2. Evidence in favor of null or alternative hypotheses if additional data are collected ?

- Context
  - 2 arms trial
  - Long recruitment due to :
  - Low recruitment rate (difficult patient population)
  - Large number of patients
  - Relatively short biomarker readout
- Motivation
  - Why should we wait until the end of the trial if partial data demonstrate that the trial is being unlikely to achieve its objective ?
  - Unethical

  $\Rightarrow$ **Stop for futility !! (or adaptive randomization design)**
- Which decisions ?
  1. Evidence in favor of null or alternative hypotheses at the interim ?
  2. Evidence in favor of null or alternative hypotheses if additional data are collected ?

- Context
  - 2 arms trial
  - Long recruitment due to :
  - Low recruitment rate (difficult patient population)
  - Large number of patients
  - Relatively short biomarker readout
- Motivation
  - Why should we wait until the end of the trial if partial data demonstrate that the trial is being unlikely to achieve its objective ?
  - Unethical

  $\Rightarrow$ **Stop for futility !! (or adaptive randomization design)**
- Which decisions ?
  1. Evidence in favor of null or alternative hypotheses at the interim ?
  2. Evidence in favor of null or alternative hypotheses if additional data are collected ?

- Context
  - 2 arms trial
  - Long recruitment due to :
  - Low recruitment rate (difficult patient population)
  - Large number of patients
  - Relatively short biomarker readout
- Motivation
  - Why should we wait until the end of the trial if partial data demonstrate that the trial is being unlikely to achieve its objective ?
  - Unethical

  $\Rightarrow$ **Stop for futility !! (or adaptive randomization design)**
- Which decisions ?
  1. Evidence in favor of null or alternative hypotheses at the interim ?
  2. Evidence in favor of null or alternative hypotheses if additional data are collected ?

Futility tools : evidence at interim

- **P-value**
  - Not the best approach as seen
- **Bayesian posterior distribution**

Table : Case study planned with Ntot= 39*2 = 78 patients

Interim at N	Posterior Distribution Pr(θ>2)	P-value
20	0.50	0.50
30	0.51	0.50
60	0.49	0.50
70	0.46	0.51

- Relative high posterior probability will not stop the trial whereas we have less chance to conclude as the data accumulate
- Posterior drawback in our context
  - Does not quantify the amount of evidence

Futility tools : evidence at interim

- **P-value**
  - Not the best approach as seen
- **Bayesian posterior distribution**

Table : Case study planned with Ntot= 39*2 = 78 patients

Interim at N	Posterior Distribution Pr(θ>2)	P-value
20	0.50	0.50
30	0.51	0.50
60	0.49	0.50
70	0.46	0.51

- Relative high posterior probability will not stop the trial whereas we have less chance to conclude as the data accumulate
- Posterior drawback in our context
  - Does not quantify the amount of remaining uncertainty

Futility tools : evidence at interim

- **P-value**
  - Not the best approach as seen
- **Bayesian posterior distribution**

Table : Case study planned with Ntot= 39*2 = 78 patients

Interim at N	Posterior Distribution Pr(θ>2)	P-value
20	0.50	0.50
30	0.51	0.50
60	0.49	0.50
70	0.46	0.51

- Relative high posterior probability will not stop the trial whereas we have less chance to conclude as the data accumulate
- Posterior drawback in our context
  - Does not quantify the amount of remaining uncertainty

Futility tools : evidence at interim

- **P-value**
  - Not the best approach as seen
- **Bayesian posterior distribution**

Table : Case study planned with Ntot= 39*2 = 78 patients

Interim at N	Posterior Distribution Pr(θ>2)	P-value
20	0.50	0.50
30	0.51	0.50
60	0.49	0.50
70	0.46	0.51

- Relative high posterior probability will not stop the trial whereas we have less chance to conclude as the data accumulate
- Posterior drawback in our context
  - Does not quantify the amount of remaining uncertainty

Futility tools : evidence with remaining data

Conditional Power : $P(FinalSuccess|\mu;\sigma)$

- Take into account for how much data remain to be observed
- But one must assume some values of parameters $\theta = (\mu;\sigma)$ are the truth, which is unknown and provide conditional power estimate under various alternatives
  1. Based upon initial assumptions for $(\mu;\sigma)$
  2. Based upon the observed values at interim (maximum likelihood estimate)
  3. Based upon the null hypothesis ...
  - If initial assumptions were optimistic, (1) will delay the stop
  - If by bad luck, an initial estimate (based on few data) underestimates the trt effect (or overestimates the variance), (2) can results in an erroneous stop of the trial
- Same drawback than classical power : conditionality

$\Rightarrow$ Let's integrate !

Futility tools : evidence with remaining data

Conditional Power : $P(FinalSuccess|\mu;\sigma)$

- Take into account for how much data remain to be observed
- But one must assume some values of parameters $\theta = (\mu;\sigma)$ are the truth, which is unknown and provide conditional power estimate under various alternatives
  1. Based upon initial assumptions for $(\mu;\sigma)$
  2. Based upon the observed values at interim (maximum likelihood estimate)
  3. Based upon the null hypothesis ...
     - If initial assumptions were optimistic, (1) will delay the stop
     - If by bad luck, an initial estimate (based on few data) underestimates the trt effect (or overestimates the variance), (2) can results in an erroneous stop of the trial
- Same drawback than classical power : conditionality

$\Rightarrow$ Let's integrate !

Futility tools : evidence with remaining data

Conditional Power : $P(FinalSuccess|\mu;\sigma)$

- Take into account for how much data remain to be observed
- But one must assume some values of parameters $\theta = (\mu;\sigma)$ are the truth, which is unknown and provide conditional power estimate under various alternatives
  1. Based upon initial assumptions for $(\mu;\sigma)$
  2. Based upon the observed values at interim (maximum likelihood estimate)
  3. Based upon the null hypothesis ...
     - If initial assumptions were optimistic, (1) will delay the stop
     - If by bad luck, an initial estimate (based on few data) underestimates the trt effect (or overestimates the variance), (2) can results in an erroneous stop of the trial
- Same drawback than classical power : conditionality

$\Rightarrow$ Let's integrate !

Futility tools : evidence with remaining data

Conditional Power : $P(FinalSuccess|\mu;\sigma)$

- Take into account for how much data remain to be observed
- But one must assume some values of parameters $\theta = (\mu;\sigma)$ are the truth, which is unknown and provide conditional power estimate under various alternatives
  1. Based upon initial assumptions for $(\mu;\sigma)$
  2. Based upon the observed values at interim (maximum likelihood estimate)
  3. Based upon the null hypothesis ...
     - If initial assumptions were optimistic, (1) will delay the stop
     - If by bad luck, an initial estimate (based on few data) underestimates the trt effect (or overestimates the variance), (2) can results in an erroneous stop of the trial
- Same drawback than classical power : conditionality

$\Rightarrow$ Let's integrate !

Futility tools : evidence with remaining data

Conditional Power : $P(FinalSuccess|\mu;\sigma)$

- Take into account for how much data remain to be observed
- But one must assume some values of parameters $\theta = (\mu;\sigma)$ are the truth, which is unknown and provide conditional power estimate under various alternatives
  1. Based upon initial assumptions for $(\mu;\sigma)$
  2. Based upon the observed values at interim (maximum likelihood estimate)
  3. Based upon the null hypothesis ...
     - If initial assumptions were optimistic, (1) will delay the stop
     - If by bad luck, an initial estimate (based on few data) underestimates the trt effect (or overestimates the variance), (2) can results in an erroneous stop of the trial
- Same drawback than classical power : conditionality

$\Rightarrow$ Let's integrate !

Futility tools : evidence with remaining data

Conditional Power : $P(FinalSuccess|\mu;\sigma)$

- Take into account for how much data remain to be observed
- But one must assume some values of parameters $\theta = (\mu;\sigma)$ are the truth, which is unknown and provide conditional power estimate under various alternatives
  1. Based upon initial assumptions for $(\mu;\sigma)$
  2. Based upon the observed values at interim (maximum likelihood estimate)
  3. Based upon the null hypothesis ...
     - If initial assumptions were optimistic, (1) will delay the stop
     - If by bad luck, an initial estimate (based on few data) underestimates the trt effect (or overestimates the variance), (2) can results in an erroneous stop of the trial
- Same drawback than classical power : conditionality

$\Rightarrow$ Let's integrate !

Futility tools : evidence with remaining data

Conditional Power : $P(FinalSuccess|\mu;\sigma)$

- Take into account for how much data remain to be observed
- But one must assume some values of parameters $\theta = (\mu;\sigma)$ are the truth, which is unknown and provide conditional power estimate under various alternatives
  1. Based upon initial assumptions for $(\mu;\sigma)$
  2. Based upon the observed values at interim (maximum likelihood estimate)
  3. Based upon the null hypothesis ...
     - If initial assumptions were optimistic, (1) will delay the stop
     - If by bad luck, an initial estimate (based on few data) underestimates the trt effect (or overestimates the variance), (2) can results in an erroneous stop of the trial
- Same drawback than classical power : conditionality

$\Rightarrow$ Let's integrate !

Futility tools : evidence with remaining data

Conditional Power : $P(FinalSuccess|\mu;\sigma)$

- Take into account for how much data remain to be observed
- But one must assume some values of parameters $\theta = (\mu;\sigma)$ are the truth, which is unknown and provide conditional power estimate under various alternatives
  1. Based upon initial assumptions for $(\mu;\sigma)$
  2. Based upon the observed values at interim (maximum likelihood estimate)
  3. Based upon the null hypothesis ...
     - If initial assumptions were optimistic, (1) will delay the stop
     - If by bad luck, an initial estimate (based on few data) underestimates the trt effect (or overestimates the variance), (2) can results in an erroneous stop of the trial
- Same drawback than classical power : conditionality

$\Rightarrow$ Let's integrate !

Futility tools : evidence with remaining data

Conditional Power : $P(FinalSuccess|\mu;\sigma)$

- Take into account for how much data remain to be observed
- But one must assume some values of parameters $\theta = (\mu;\sigma)$ are the truth, which is unknown and provide conditional power estimate under various alternatives
  1. Based upon initial assumptions for $(\mu;\sigma)$
  2. Based upon the observed values at interim (maximum likelihood estimate)
  3. Based upon the null hypothesis ...
     - If initial assumptions were optimistic, (1) will delay the stop
     - If by bad luck, an initial estimate (based on few data) underestimates the trt effect (or overestimates the variance), (2) can results in an erroneous stop of the trial
- Same drawback than classical power : conditionality

$\Rightarrow$ Let's integrate !

Futility tools : evidence with remaining data

Bayesian Predictive probability[3]

- $P(Success) = \int_\theta P(Success|\theta).P(\theta|Data_{obs}\ at\ IA)d\theta = \int_\theta P(Success|\theta).P(Data_{obs}\ at\ IA|\theta).P(\theta)d\theta$

Interim at N	**Posterior Distribution Pr(θ>2)**	**Posterior Predictive** Probability
20	0.50	0.3
30	0.51	0.2
60	0.49	0.1
70	**0.46**	**0.05**

- Much better for decision making !
- Take into account
  - the current estimates
  - the amount of data that remains to be observed
- Simulations for thresholds/risks and robustness

[3] Lan et al. 2009.

Futility tools : evidence with remaining data

Bayesian Predictive probability[3]

- $P(Success) = \int_\theta P(Success|\theta).P(\theta|Data_{obs}\ at\ IA)d\theta = \int_\theta P(Success|\theta).P(Data_{obs}\ at\ IA|\theta).P(\theta)d\theta$

Interim at N	**Posterior Distribution Pr(θ>2)**	**Posterior Predictive** Probability
20	0.50	0.3
30	0.51	0.2
60	0.49	0.1
70	**0.46**	**0.05**

- Much better for decision making !
- Take into account
  - the current estimates
  - the amount of data that remains to be observed
- Simulations for thresholds/risks and robustness

[3] Lan et al. 2009.

Futility tools : evidence with remaining data

Bayesian Predictive probability[3]

- $P(Success) = \int_\theta P(Success|\theta).P(\theta|Data_{obs}\ at\ IA)d\theta = \int_\theta P(Success|\theta).P(Data_{obs}\ at\ IA|\theta).P(\theta)d\theta$

Interim at N	**Posterior Distribution Pr(θ>2)**	**Posterior Predictive** Probability
20	0.50	0.3
30	0.51	0.2
60	0.49	0.1
70	**0.46**	**0.05**

- Much better for decision making !
- Take into account
  - the current estimates
  - the amount of data that remains to be observed
- Simulations for thresholds/risks and robustness

[3] Lan et al. 2009.

Futility tools : evidence with remaining data

Bayesian Predictive probability[3]

- $P(Success) = \int_\theta P(Success|\theta).P(\theta|Data_{obs}\ at\ IA)d\theta = \int_\theta P(Success|\theta).P(Data_{obs}\ at\ IA|\theta).P(\theta)d\theta$

Interim at N	**Posterior Distribution Pr(θ>2)**	**Posterior Predictive** Probability
20	0.50	0.3
30	0.51	0.2
60	0.49	0.1
70	**0.46**	**0.05**

- Much better for decision making !
- Take into account
  - the current estimates
  - the amount of data that remains to be observed
- Simulations for thresholds/risks and robustness

[3] Lan et al. 2009.

Futility tools : evidence with remaining data

Bayesian Predictive probability[3]

- $P(Success) = \int_\theta P(Success|\theta).P(\theta|Data_{obs}\ at\ IA)d\theta = \int_\theta P(Success|\theta).P(Data_{obs}\ at\ IA|\theta).P(\theta)d\theta$

Interim at N	**Posterior Distribution Pr(θ>2)**	**Posterior Predictive** Probability
20	0.50	0.3
30	0.51	0.2
60	0.49	0.1
70	**0.46**	**0.05**

- Much better for decision making !
- Take into account
  - the current estimates
  - the amount of data that remains to be observed
- Simulations for thresholds/risks and robustness

[3] Lan et al. 2009.

Futility tools : evidence with remaining data

Bayesian Predictive probability[3]

- $P(Success) = \int_\theta P(Success|\theta).P(\theta|Data_{obs}\ at\ IA)d\theta = \int_\theta P(Success|\theta).P(Data_{obs}\ at\ IA|\theta).P(\theta)d\theta$

Interim at N	**Posterior Distribution Pr(θ>2)**	**Posterior Predictive** Probability
20	0.50	0.3
30	0.51	0.2
60	0.49	0.1
70	**0.46**	**0.05**

- Much better for decision making !
- Take into account
  - the current estimates
  - the amount of data that remains to be observed
- Simulations for thresholds/risks and robustness

[3] Lan et al. 2009.

First Part summary

- Go/NoGo criteria needs to be defined upfront
- Bayesian methods allow proper quantification of the benefit/risk (using clinically relevant threshold and/or cost function)
- Prior information "naturally" included per iterative process of learning
- Cost savings by efficient futility stopping rules (conditional power, posterior or predictive probability)

First Part summary

- Go/NoGo criteria needs to be defined upfront
- Bayesian methods allow proper quantification of the benefit/risk (using clinically relevant threshold and/or cost function)
- Prior information "naturally" included per iterative process of learning
- Cost savings by efficient futility stopping rules (conditional power, posterior or predictive probability)

First Part summary

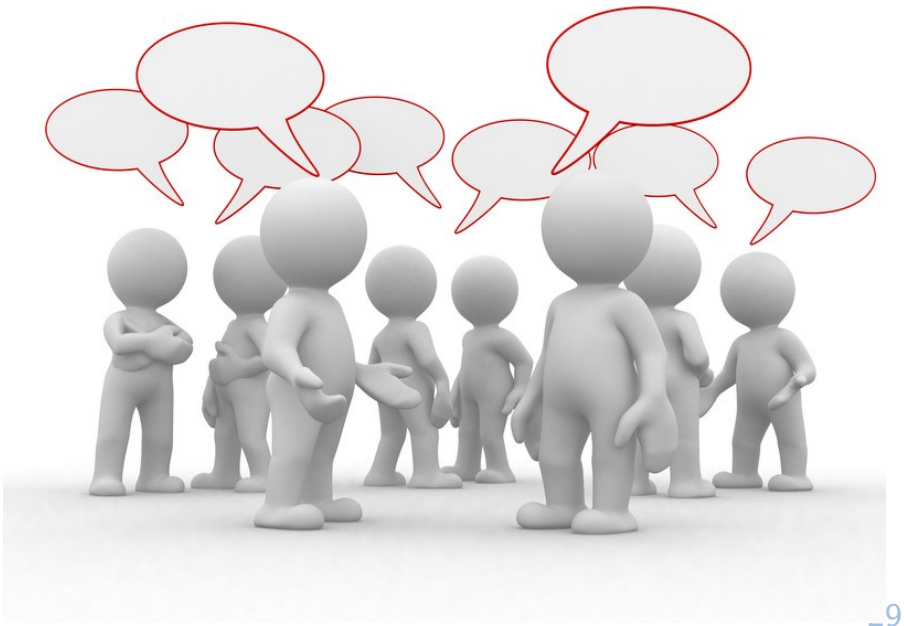
- Go/NoGo criteria needs to be defined upfront
- Bayesian methods allow proper quantification of the benefit/risk (using clinically relevant threshold and/or cost function)
- Prior information "naturally" included per iterative process of learning
- Cost savings by efficient futility stopping rules (conditional power, posterior or predictive probability)

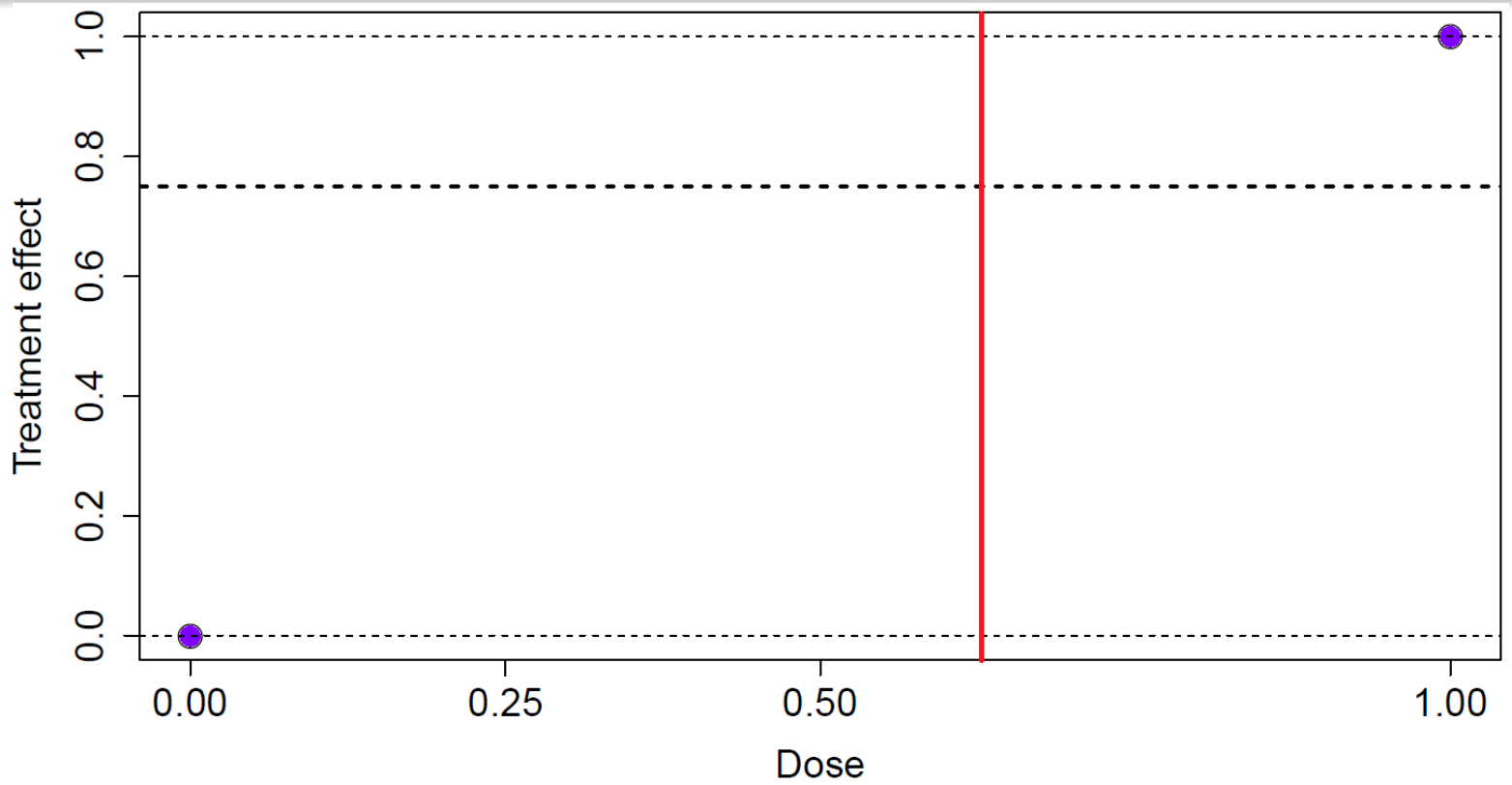
First Part summary

- Go/NoGo criteria needs to be defined upfront
- Bayesian methods allow proper quantification of the benefit/risk (using clinically relevant threshold and/or cost function)
- Prior information "naturally" included per iterative process of learning
- Cost savings by efficient futility stopping rules (conditional power, posterior or predictive probability)

Discussions

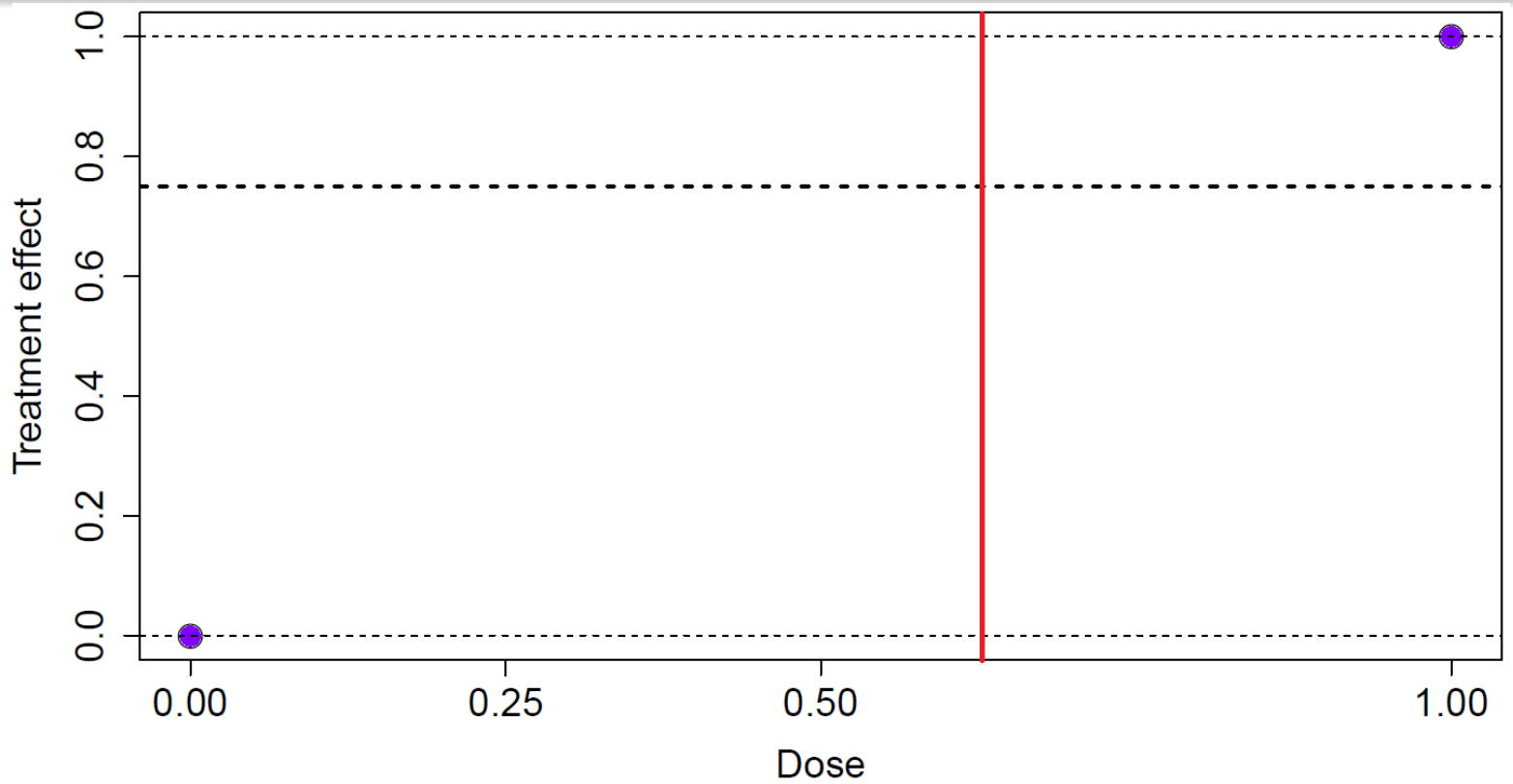
- **What types of PoC designs have you implemented? What were the main benefits or issues?**
- **Who will use (Bayesian) information in future PoC trials? (Yes/ No/ Maybe)**





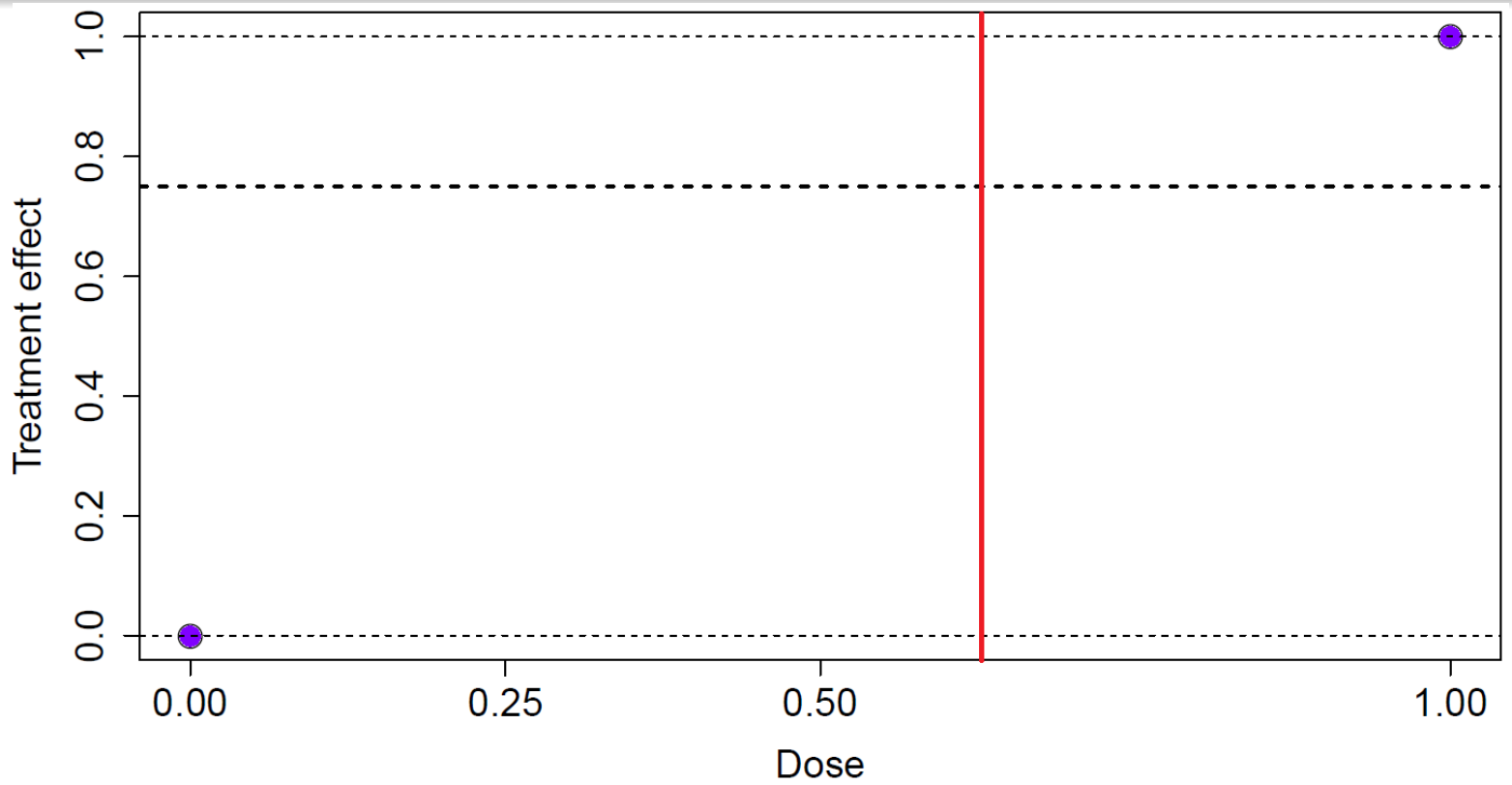
- We target an effect above 0.75
  - Which dose for the next development steps ?
- And if unacceptable safety above dose 0.6 ?

$\Rightarrow$ We can answer with a muti-arms design, getting a rough idea of a "therapeutic window"



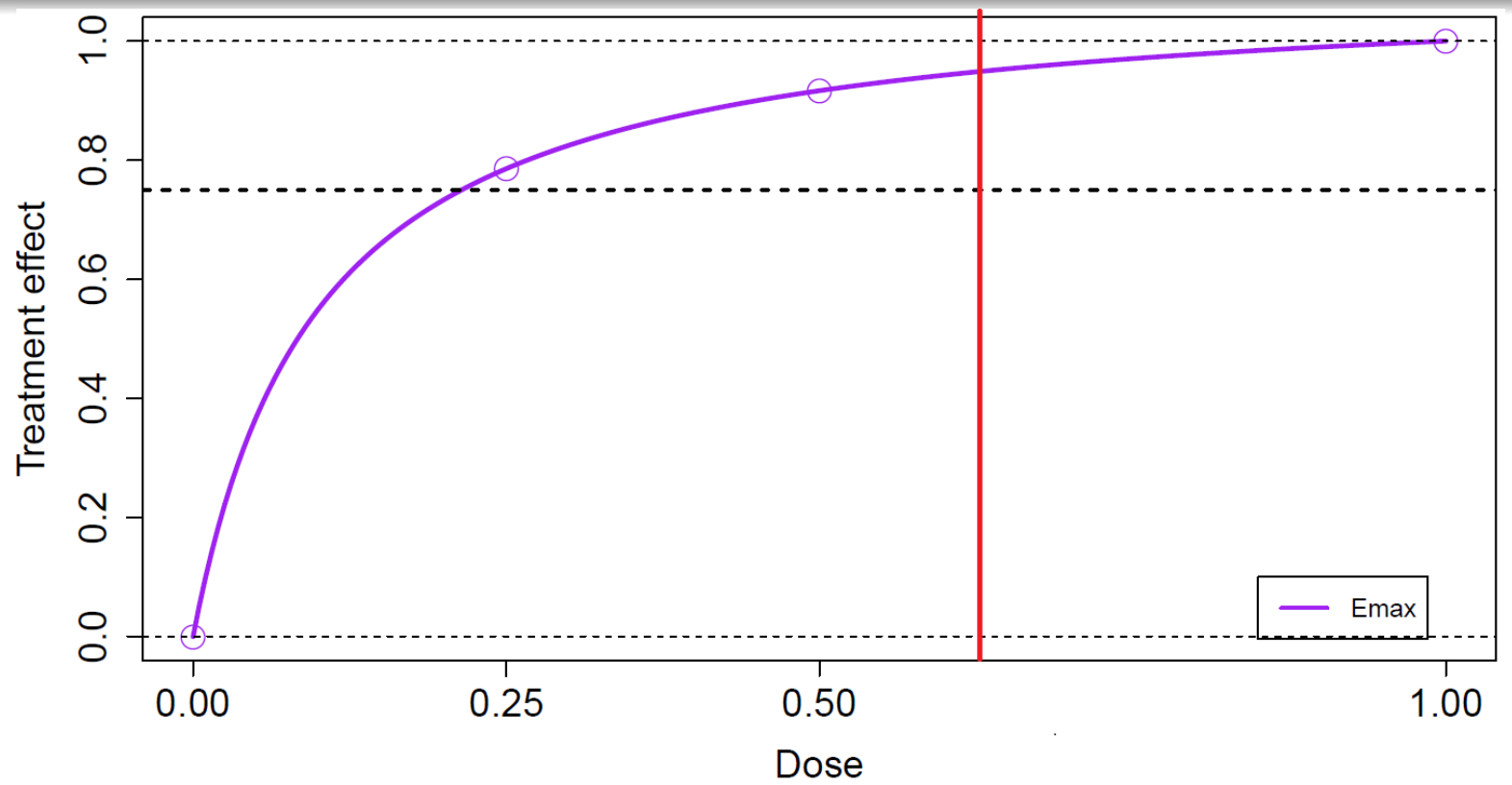
- We target an effect above 0.75
  - Which dose for the next development steps ?
- And if unacceptable safety above dose 0.6 ?

$\Rightarrow$ We can answer with a muti-arms design, getting a rough idea of a "therapeutic window"



- We target an effect above 0.75
  - Which dose for the next development steps ?
- And if unacceptable safety above dose 0.6 ?

$\Rightarrow$ We can answer with a muti-arms design, getting a rough idea of a "therapeutic window"



- We target an effect above 0.75
  - Which dose for the next development steps ?
- And if unacceptable safety above dose 0.6 ?

$\Rightarrow$ We can answer with a muti-arms design, getting a rough idea of a "therapeutic window"

Dose reponse is of primary importance from a pharmacological point of view

- For learning and next development steps
- Especially in case of suspected U shape

By doing a 2 arms PoC at MTD...

✓ We go fast and cheap
- Assume little safety issue with MTD in PoC population
- Assume monotonic dose-response

✗ We make strong assumption in choosing a dose which yields the targeted effect

✗ Do not characterize the dose-response curve

✗ We take the risk to conclude on a bad efficacy/risk ratio and so to stop a promising compound

⇒ **Can lead to ambiguity in PoC decision-making**

Dose reponse is of primary importance from a pharmacological point of view

- For learning and next development steps
- Especially in case of suspected U shape

By doing a 2 arms PoC at MTD...

✓ We go fast and cheap
- Assume little safety issue with MTD in PoC population
- Assume monotonic dose-response

✗ We make strong assumption in choosing a dose which yields the targeted effect

✗ Do not characterize the dose-response curve

✗ We take the risk to conclude on a bad efficacy/risk ratio and so to stop a promising compound

⇒ **Can lead to ambiguity in PoC decision-making**

Dose reponse is of primary importance from a pharmacological point of view

- For learning and next development steps
- Especially in case of suspected U shape

By doing a 2 arms PoC at MTD...

✓ We go fast and cheap
- Assume little safety issue with MTD in PoC population
- Assume monotonic dose-response

✗ We make strong assumption in choosing a dose which yields the targeted effect

✗ Do not characterize the dose-response curve

✗ We take the risk to conclude on a bad efficacy/risk ratio and so to stop a promising compound

⇒ **Can lead to ambiguity in PoC decision-making**

Dose reponse is of primary importance from a pharmacological point of view

- For learning and next development steps
- Especially in case of suspected U shape

By doing a 2 arms PoC at MTD...

✓ We go fast and cheap
- Assume little safety issue with MTD in PoC population
- Assume monotonic dose-response

✗ We make strong assumption in choosing a dose which yields the targeted effect

✗ Do not characterize the dose-response curve

✗ We take the risk to conclude on a bad efficacy/risk ratio and so to stop a promising compound

⇒ **Can lead to ambiguity in PoC decision-making**

Dose reponse is of primary importance from a pharmacological point of view

- For learning and next development steps
- Especially in case of suspected U shape

By doing a 2 arms PoC at MTD...

✓ We go fast and cheap
- Assume little safety issue with MTD in PoC population
- Assume monotonic dose-response

✗ We make strong assumption in choosing a dose which yields the targeted effect

✗ Do not characterize the dose-response curve

✗ We take the risk to conclude on a bad efficacy/risk ratio and so to stop a promising compound

⇒ **Can lead to ambiguity in PoC decision-making**

Dose reponse is of primary importance from a pharmacological point of view

- For learning and next development steps
- Especially in case of suspected U shape

By doing a 2 arms PoC at MTD...

✓ We go fast and cheap
- Assume little safety issue with MTD in PoC population
- Assume monotonic dose-response

✗ We make strong assumption in choosing a dose which yields the targeted effect

✗ Do not characterize the dose-response curve

✗ We take the risk to conclude on a bad efficacy/risk ratio and so to stop a promising compound

⇒ **Can lead to ambiguity in PoC decision-making**

Dose reponse is of primary importance from a pharmacological point of view

- For learning and next development steps
- Especially in case of suspected U shape

By doing a 2 arms PoC at MTD...

✓ We go fast and cheap
  - Assume little safety issue with MTD in PoC population
  - Assume monotonic dose-response

✗ We make strong assumption in choosing a dose which yields the targeted effect

✗ Do not characterize the dose-response curve

✗ We take the risk to conclude on a bad efficacy/risk ratio and so to stop a promising compound

⇒ **Can lead to ambiguity in PoC decision-making**

Dose reponse is of primary importance from a pharmacological point of view

- For learning and next development steps
- Especially in case of suspected U shape

By doing a 2 arms PoC at MTD...

✓ We go fast and cheap
- Assume little safety issue with MTD in PoC population
- Assume monotonic dose-response

✗ We make strong assumption in choosing a dose which yields the targeted effect

✗ Do not characterize the dose-response curve

✗ We take the risk to conclude on a bad efficacy/risk ratio and so to stop a promising compound

$\Rightarrow$ Can lead to ambiguity in PoC decision-making

Dose reponse is of primary importance from a pharmacological point of view

- For learning and next development steps
- Especially in case of suspected U shape

By doing a 2 arms PoC at MTD...

✓ We go fast and cheap
- Assume little safety issue with MTD in PoC population
- Assume monotonic dose-response

✗ We make strong assumption in choosing a dose which yields the targeted effect

✗ Do not characterize the dose-response curve

✗ We take the risk to conclude on a bad efficacy/risk ratio and so to stop a promising compound

$\Rightarrow$ Can lead to ambiguity in PoC decision-making

Dose reponse is of primary importance from a pharmacological point of view

- For learning and next development steps
- Especially in case of suspected U shape

By doing a 2 arms PoC at MTD...

✓ We go fast and cheap
  - Assume little safety issue with MTD in PoC population
  - Assume monotonic dose-response

✗ We make strong assumption in choosing a dose which yields the targeted effect

✗ Do not characterize the dose-response curve

✗ We take the risk to conclude on a bad efficacy/risk ratio and so to stop a promising compound

$\Rightarrow$ **Can lead to ambiguity in PoC decision-making**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

⇒ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

⇒ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

⇒ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ **By adding few patients at intermediate doses we leverage the information**

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ By adding few patients at intermediate doses we leverage the information

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ By adding few patients at intermediate doses we leverage the information

Alternative strategy is to include multiple dose levels

- Mitigates risk of false PoC failure due to incorrect dose
- PoC can be tested via several prespecified dose-response candidate model
  - Mitigates risk of false PoC failure due to incorrect dose-response model
  - Analysis adaptive approach
- Provides information to guide subsequent dose selection
  - Better characterize the "therapeutic window"
  - Reduce the risk of selecting the wrong dose
  - Dose-finding still required but can leverage PoC data
- Usually requires modest increase in sample size vs. 2-arm
  - Consider unbalanced randomization
- PoC can be demonstrated by positive dose-response
  - For assumed dose-response models (linear, log-linear, Emax...)

$\Rightarrow$ By adding few patients at intermediate doses we leverage the information

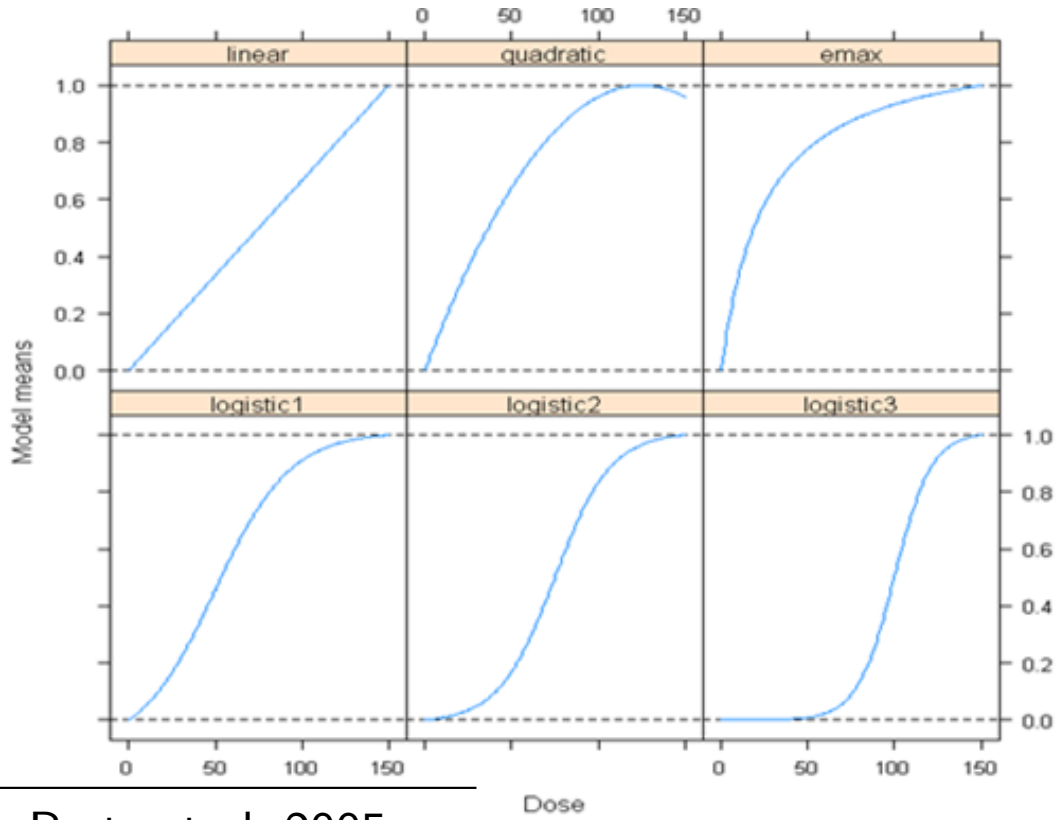
PoC can be demonstrated by positive dose-response relationship or precision objective

From candidate models defined in the protocol[4]

1. Show that a significant dose-effect exist using contrast tests while controlling the family-wise error rate (at 5%)
   - The multiplicity issue is solved by a numerical approach taking into account the correlation between models
   - Retain models above a critical value
2. If yes, fit model on data (AIC $\rightarrow$ weight) and eventually calculate the dose of interest (MED, ED50, ED90) from the selected model(s)
   - By model selection or model averaging (frequentist or bayesian, taking into account the uncertainty of the selection process itself)

[4] MCP-Mod, Bretz et al. 2005.

PoC can be demonstrated by positive dose-response relationship or precision objective
From candidate models defined in the protocol[4]



[4] MCP-Mod, Bretz et al. 2005.

PoC can be demonstrated by positive dose-response relationship or precision objective
From candidate models defined in the protocol[4]

1. Show that a significant dose-effect exist using contrast tests while controlling the family-wise error rate (at 5%)
   - The multiplicity issue is solved by a numerical approach taking into account the correlation between models
   - Retain models above a critical value
2. If yes, fit model on data (AIC $\rightarrow$ weight) and eventually calculate the dose of interest (MED, ED50, ED90) from the selected model(s)
   - By model selection or model averaging (frequentist or bayesian, taking into account the uncertainty of the selection process itself)

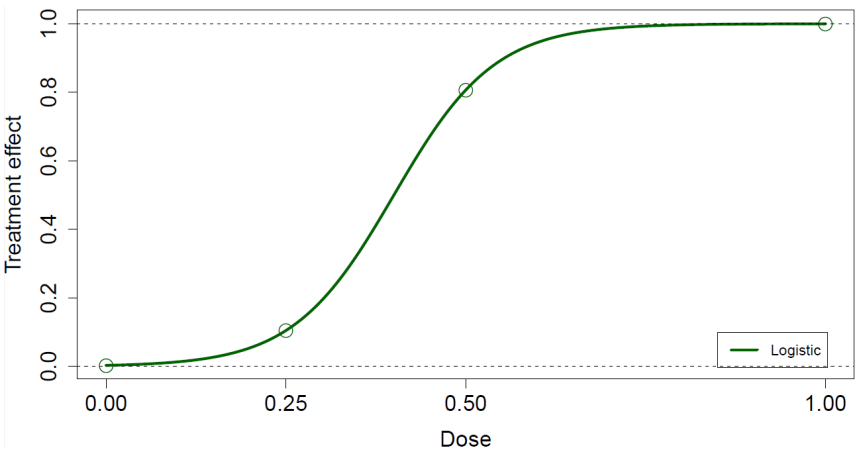
[4] MCP-Mod, Bretz et al. 2005.

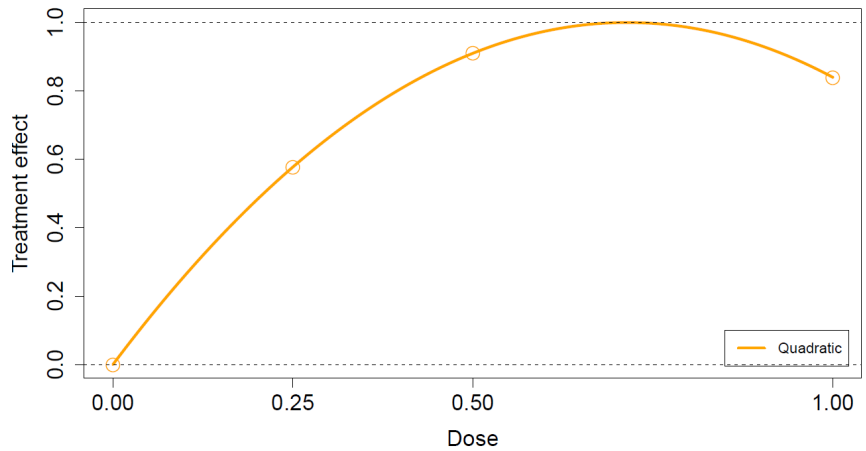
Case study

- Placebo + 3 actives dose $= \{0; 0.25; 0.5; 1\}$; $\alpha = 0.05$ one-sided
- Unbalanced randomization 3:1:1:3
- Maximal ES = 1.67 (SD=0.6)
- 2 arms design (N=18) ANOVA vs 4 arms design (N=24) MCPmod
    - True dose-response : Logistic or quadratic (6% decrease)

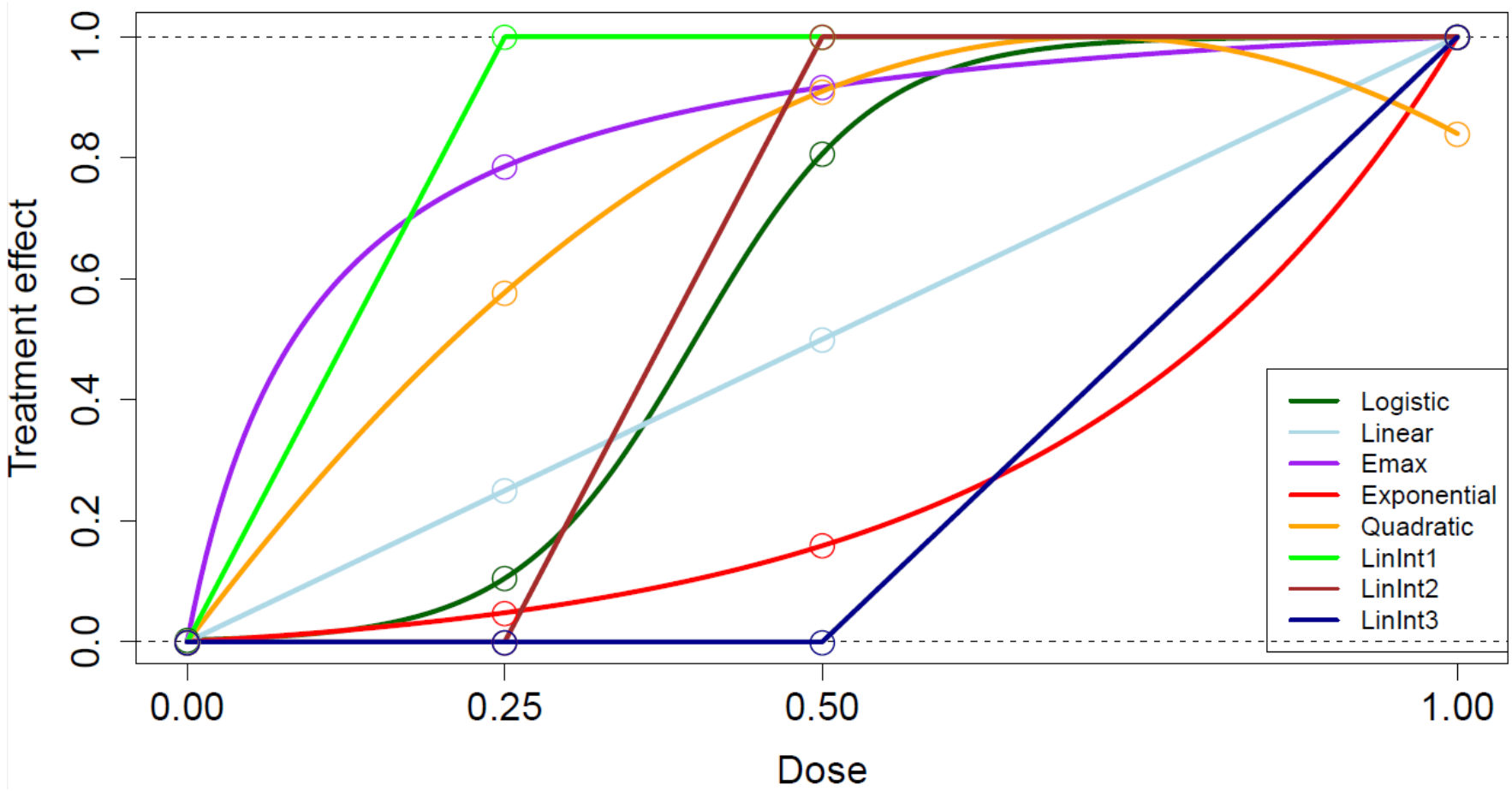
Case study

- Placebo + 3 actives dose $= \{0; 0.25; 0.5; 1\}$; $\alpha = 0.05$ one-sided
- Unbalanced randomization 3:1:1:3
- Maximal ES = 1.67 (SD=0.6)
- 2 arms design (N=18) ANOVA vs 4 arms design (N=24) MCPmod
  - True dose-response : Logistic or quadratic (6% decrease)





Case study : Candidate models



Case study : Results

Table : Power results

Analysis	Sample Size	Design	True Dose response	Power
Anova	18	$\{9; 0; 0; 9\}$	Logistic	95.8%
			Quadratic	88.4%
Anova	24	$\{12; 0; 0; 12\}$	Logistic	98%
			Quadratic	95%
MCP-mod	24	$\{9; 3; 3; 9\}$		92.5% (87.4%-95.6%)

Case study : Conclusions

- Detect a dose-response : evidence of activity represented by a change in clinical response from a change in dose (PoC)
- Estimate the dose-response
  - One can determine if a predefined clinically relevant response could be obtained within the dose-range
  - One can select (interpolate) a targeted dose
- Besides analysis-focused adaptive approach
  - One can choose a design focused adaptive approach, as the General Adaptive Dose Allocation approach (GADA)
- This is just a first step, at the PoC stage
  - need for larger N to obtain good estimates of target doses

Case study : Conclusions

- Detect a dose-response : evidence of activity represented by a change in clinical response from a change in dose (PoC)
- Estimate the dose-response
  - One can determine if a predefined clinically relevant response could be obtained within the dose-range
  - One can select (interpolate) a targeted dose
- Besides analysis-focused adaptive approach
  - One can choose a design focused adaptive approach, as the General Adaptive Dose Allocation approach (GADA)
- This is just a first step, at the PoC stage
  - need for larger N to obtain good estimates of target doses

Case study : Conclusions

- Detect a dose-response : evidence of activity represented by a change in clinical response from a change in dose (PoC)
- Estimate the dose-response
  - One can determine if a predefined clinically relevant response could be obtained within the dose-range
  - One can select (interpolate) a targeted dose
- Besides analysis-focused adaptive approach
  - One can choose a design focused adaptive approach, as the General Adaptive Dose Allocation approach (GADA)
- This is just a first step, at the PoC stage
  - need for larger N to obtain good estimates of target doses

Case study : Conclusions

- Detect a dose-response : evidence of activity represented by a change in clinical response from a change in dose (PoC)
- Estimate the dose-response
  - One can determine if a predefined clinically relevant response could be obtained within the dose-range
  - One can select (interpolate) a targeted dose
- Besides analysis-focused adaptive approach
  - One can choose a design focused adaptive approach, as the General Adaptive Dose Allocation approach (GADA)
- This is just a first step, at the PoC stage
  - need for larger N to obtain good estimates of target doses

Case study : Conclusions

- Detect a dose-response : evidence of activity represented by a change in clinical response from a change in dose (PoC)
- Estimate the dose-response
  - One can determine if a predefined clinically relevant response could be obtained within the dose-range
  - One can select (interpolate) a targeted dose
- Besides analysis-focused adaptive approach
  - One can choose a design focused adaptive approach, as the General Adaptive Dose Allocation approach (GADA)
- This is just a first step, at the PoC stage
  - need for larger N to obtain good estimates of target doses

Case study : Conclusions

- Detect a dose-response : evidence of activity represented by a change in clinical response from a change in dose (PoC)
- Estimate the dose-response
  - One can determine if a predefined clinically relevant response could be obtained within the dose-range
  - One can select (interpolate) a targeted dose
- Besides analysis-focused adaptive approach
  - One can choose a design focused adaptive approach, as the General Adaptive Dose Allocation approach (GADA)
- This is just a first step, at the PoC stage
  - need for larger N to obtain good estimates of target doses

Case study : Conclusions

- Detect a dose-response : evidence of activity represented by a change in clinical response from a change in dose (PoC)
- Estimate the dose-response
  - One can determine if a predefined clinically relevant response could be obtained within the dose-range
  - One can select (interpolate) a targeted dose
- Besides analysis-focused adaptive approach
  - One can choose a design focused adaptive approach, as the General Adaptive Dose Allocation approach (GADA)
- This is just a first step, at the PoC stage
  - need for larger N to obtain good estimates of target doses

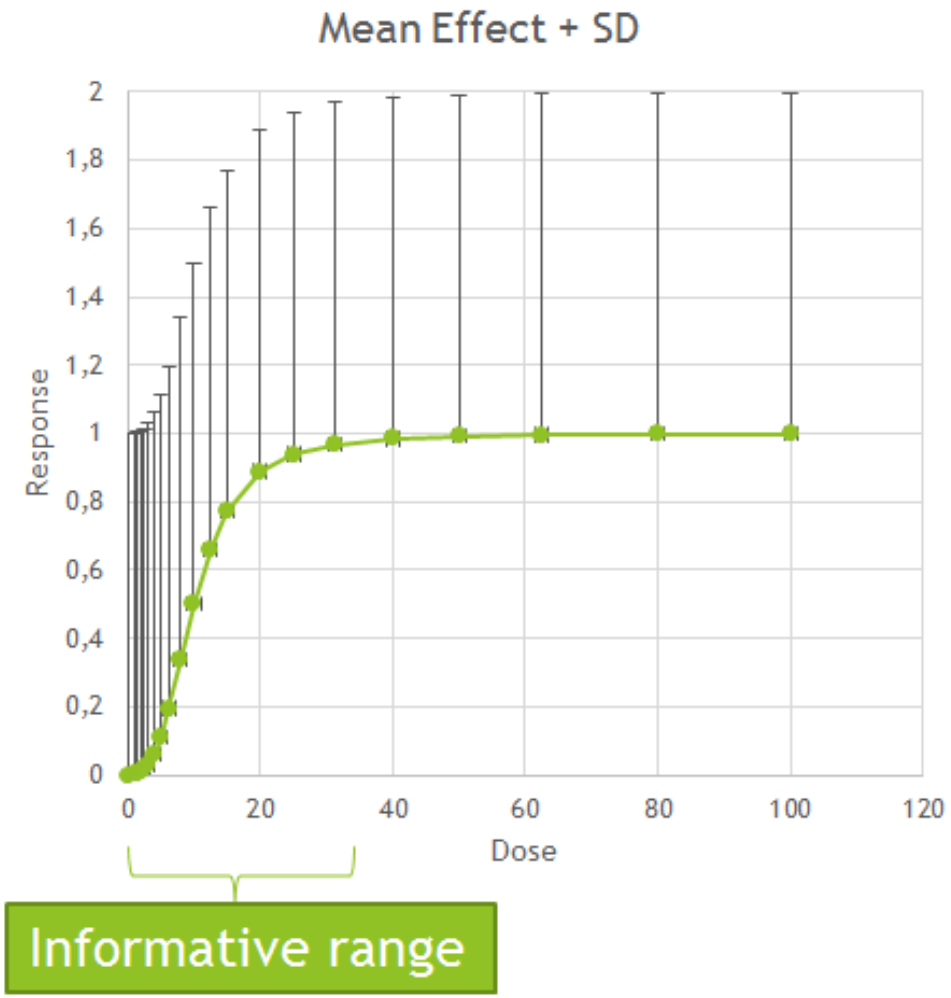
Case study : Conclusions

- Detect a dose-response : evidence of activity represented by a change in clinical response from a change in dose (PoC)
- Estimate the dose-response
  - One can determine if a predefined clinically relevant response could be obtained within the dose-range
  - One can select (interpolate) a targeted dose
- Besides analysis-focused adaptive approach
  - One can choose a design focused adaptive approach, as the General Adaptive Dose Allocation approach (GADA)
- This is just a first step, at the PoC stage
  - need for larger N to obtain good estimates of target doses

Case study from ClinBAY

- Compound at end of phase I
  - Maximum dose of 100mg, determined based on safety and PK
- Phase Ib trial goals
  - Proof of concept
  - Dose finding for phase II
- Adaptive dose-ranging versus control

True Dose-Response Relationship



- True Emax model
  - $E_0 = 0$
  - Emax=1
  - ED50=10 mg
  - Hill coefficient = 3
- Signal /Noise ratio = 1
- Placebo + 21 available doses from 1 to 100 mg
- Model-based adaptive dosing finding strategy

Trial Objectives and Bayesian Analysis

Trial Objectives

- Drug effect
  - Test Emax >0
- Dose response
  - Estimate Emax model precisely
- Dose selection
  - Find dose with relevant effect : ED50

Priors

- Informative on E_0: based on historical data
  - Mean = 0, SD = 0.6
- No information on other parameters

Trial Objectives and Bayesian Analysis

Trial Objectives

- Drug effect
  - Test Emax >0
- Dose response
  - Estimate Emax model precisely
- Dose selection
  - Find dose with relevant effect : ED50

Priors

- Informative on E_0: based on historical data
  - Mean = 0, SD = 0.6
- No information on other parameters

Adaptive Strategy

- Enroll subjects in small successive cohorts & perform dose-response modeling to select doses after each cohort
- First cohort: N=8
  - Placebo (N=2)
  - 3 dose levels : 1, 10, 100mg (N=2/dose)
- Next cohorts (N=8):
  - Placebo (N=2)
  - Flexible Dose Levels between 1 and 100mg : Target = minimum dose $>$ ED50 (N=6)
- Decision criteria & stopping rules
  - $Pr(Emax > 0) > 95\%$: drug effect
  - Precision of ED50: $CV < 30\%$
  - Stop after a maximum of 6 cohorts (N=48)

Adaptive Strategy

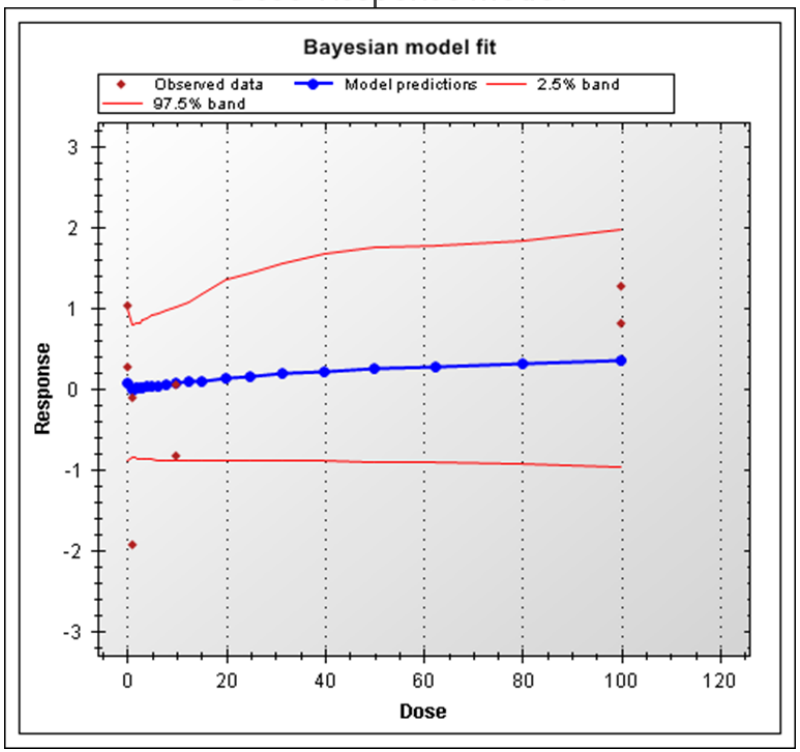
- Enroll subjects in small successive cohorts & perform dose-response modeling to select doses after each cohort
- First cohort: N=8
  - Placebo (N=2)
  - 3 dose levels : 1, 10, 100mg (N=2/dose)
- Next cohorts (N=8):
  - Placebo (N=2)
  - Flexible Dose Levels between 1 and 100mg : Target = minimum dose $>$ ED50 (N=6)
- Decision criteria & stopping rules
  - $Pr(Emax > 0) > 95\%$: drug effect
  - Precision of ED50: $CV < 30\%$
  - Stop after a maximum of 6 cohorts (N=48)

Adaptive Strategy

- Enroll subjects in small successive cohorts & perform dose-response modeling to select doses after each cohort
- First cohort: N=8
  - Placebo (N=2)
  - 3 dose levels : 1, 10, 100mg (N=2/dose)
- Next cohorts (N=8):
  - Placebo (N=2)
  - Flexible Dose Levels between 1 and 100mg : Target = minimum dose $>$ ED50 (N=6)
- Decision criteria & stopping rules
  - $Pr(Emax > 0) > 95\%$: drug effect
  - Precision of ED50: $CV < 30\%$
  - Stop after a maximum of 6 cohorts (N=48)

Adaptive Strategy : After Cohort 1 (N=8)

Dose-Response model

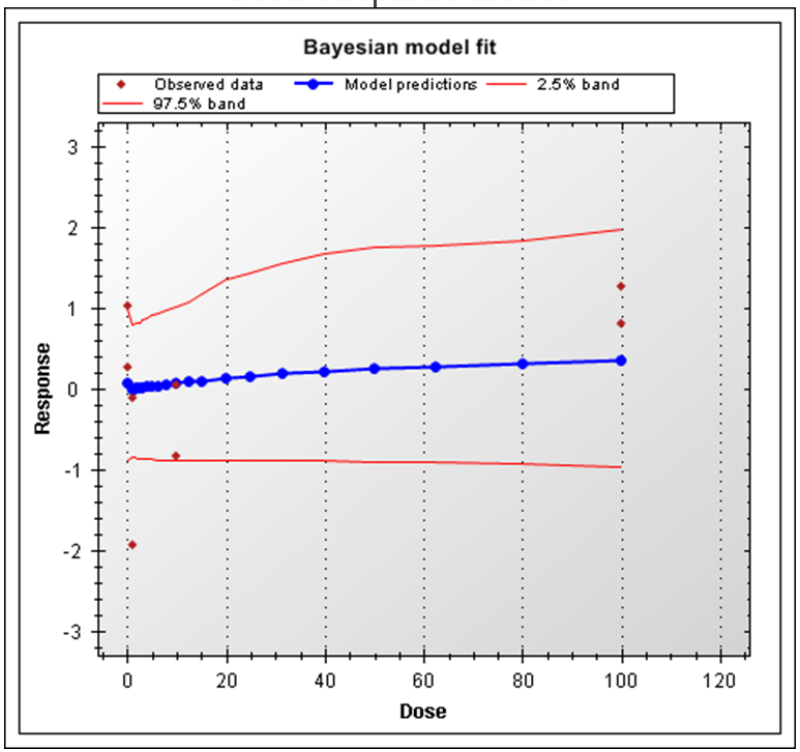


✗ Pr(Emax>0)=56%

⇒ Continue to cohort 2

Adaptive Strategy : After Cohort 1 (N=8)

Dose-Response model

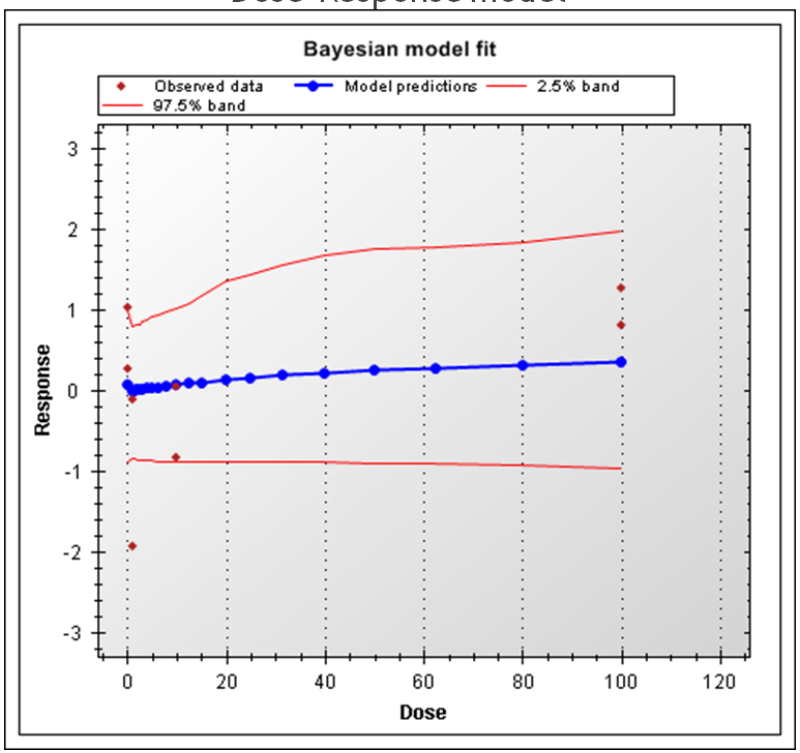


✗ Pr(Emax>0)=56%

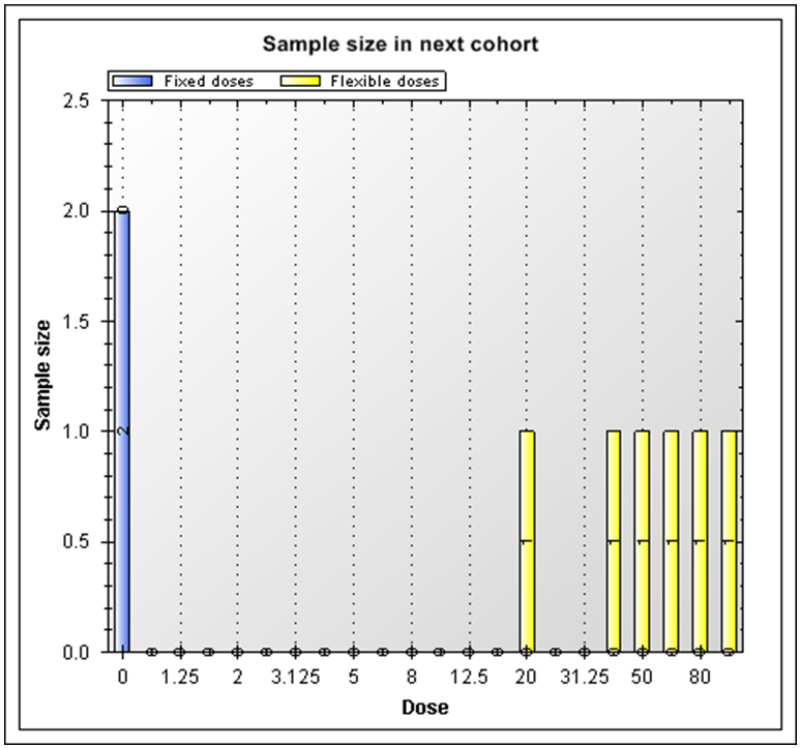
⇒ Continue to cohort 2

Adaptive Strategy : After Cohort 1 (N=8)

Dose-Response model



Choice of doses for cohort 2

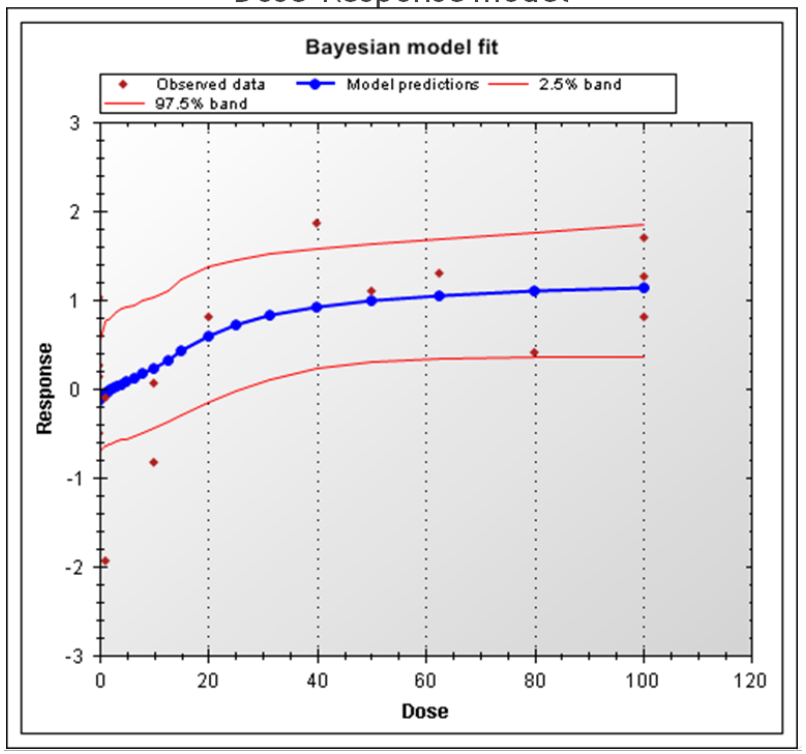


✗ Pr(Emax>0)=56%

⇒ Continue to cohort 2

Adaptive Strategy : After Cohort 2 (N=16)

Dose-Response model



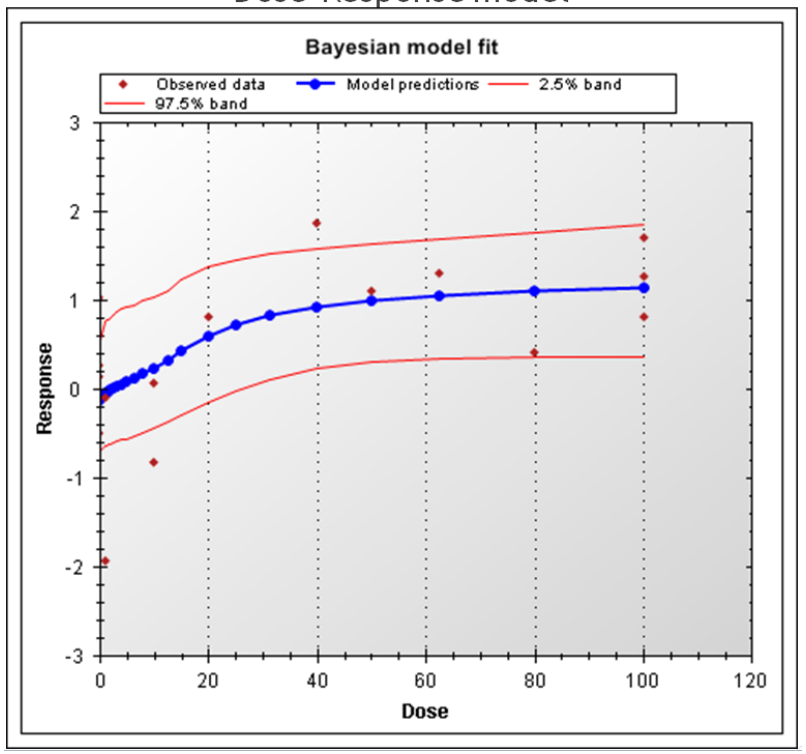
✓ Pr(Emax>0)=99%: Proof of concept achieved !

✗ CV(ED50)=83%

⇒ Continue to cohort 3

Adaptive Strategy : After Cohort 2 (N=16)

Dose-Response model



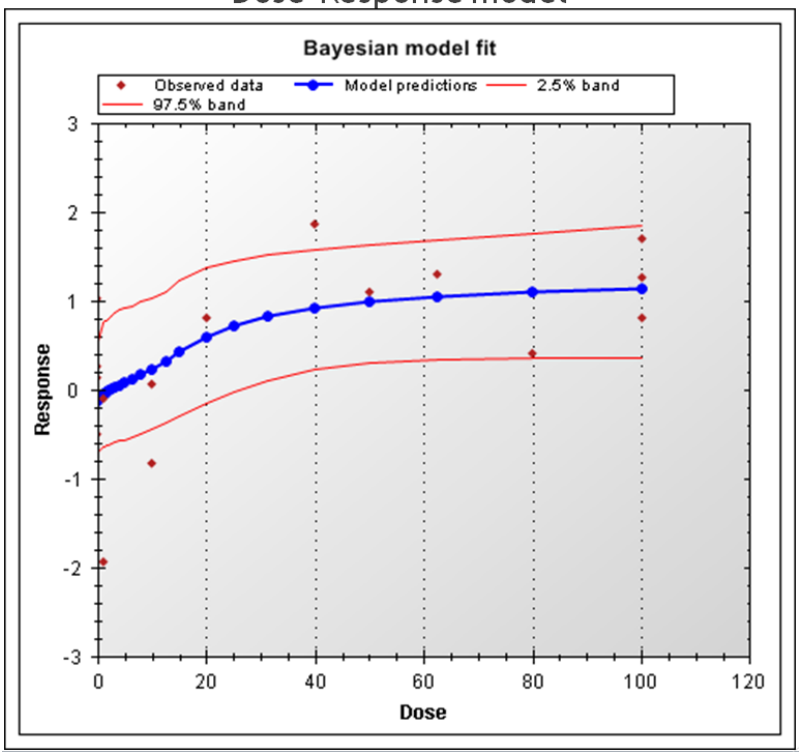
✓ Pr(Emax>0)=99%: Proof of concept achieved !

✗ CV(ED50)=83%

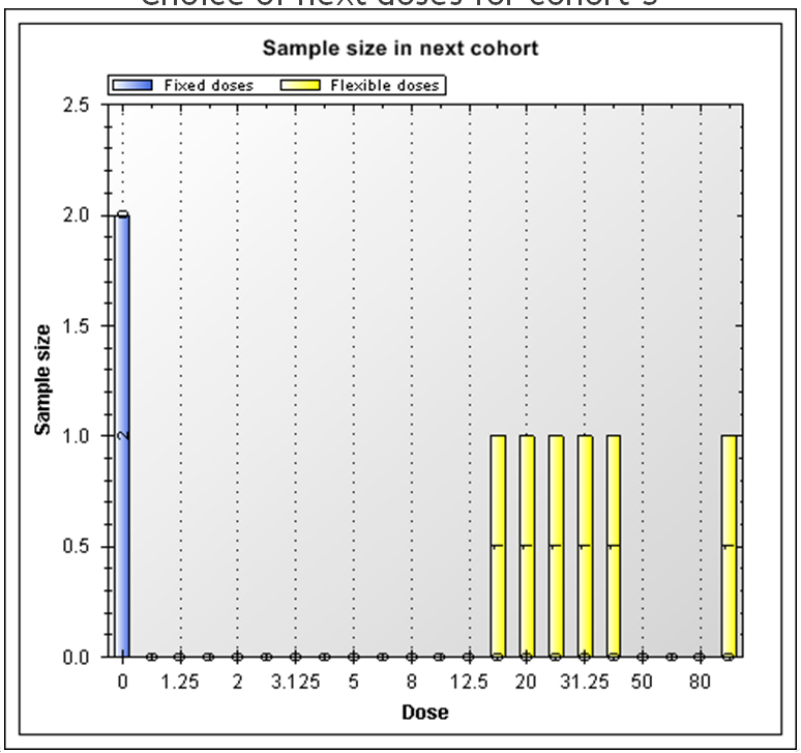
⇒ Continue to cohort 3

Adaptive Strategy : After Cohort 2 (N=16)

Dose-Response model



Choice of next doses for cohort 3



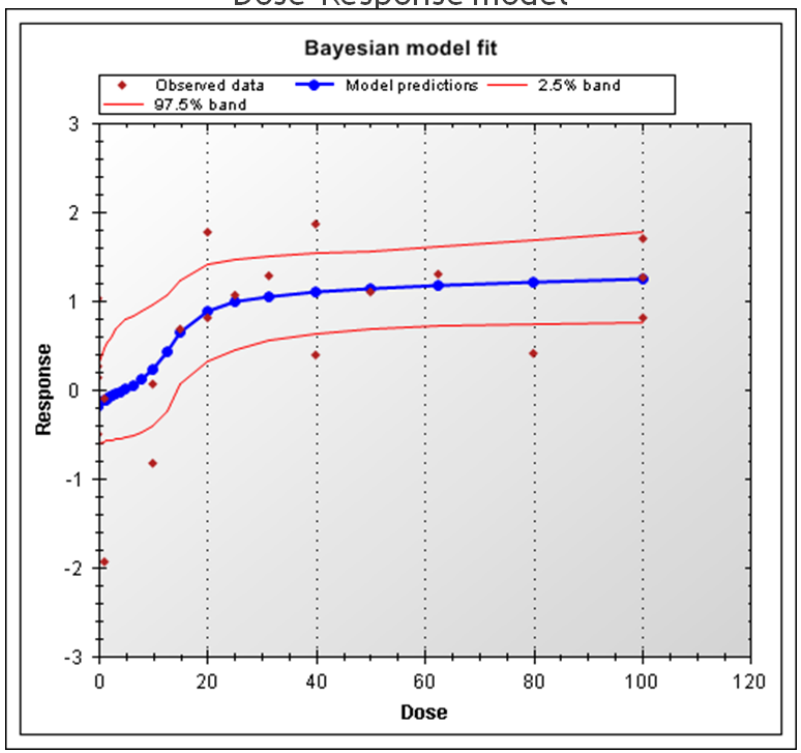
✓ Pr(Emax>0)=99%: Proof of concept achieved !

✗ CV(ED50)=83%

⇒ Continue to cohort 3

Adaptive Strategy : After Cohort 3 (N=24)

Dose-Response model



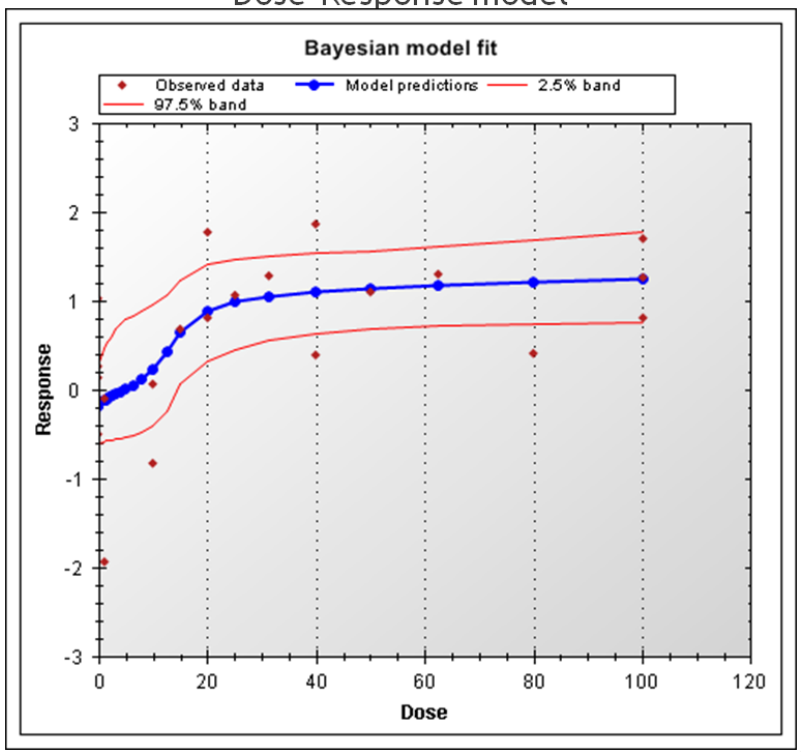
✓ Pr(Emax>0)=99%: Proof of concept achieved !

✗ CV(ED50)=49%

⇒ Continue to cohort 4

Adaptive Strategy : After Cohort 3 (N=24)

Dose-Response model



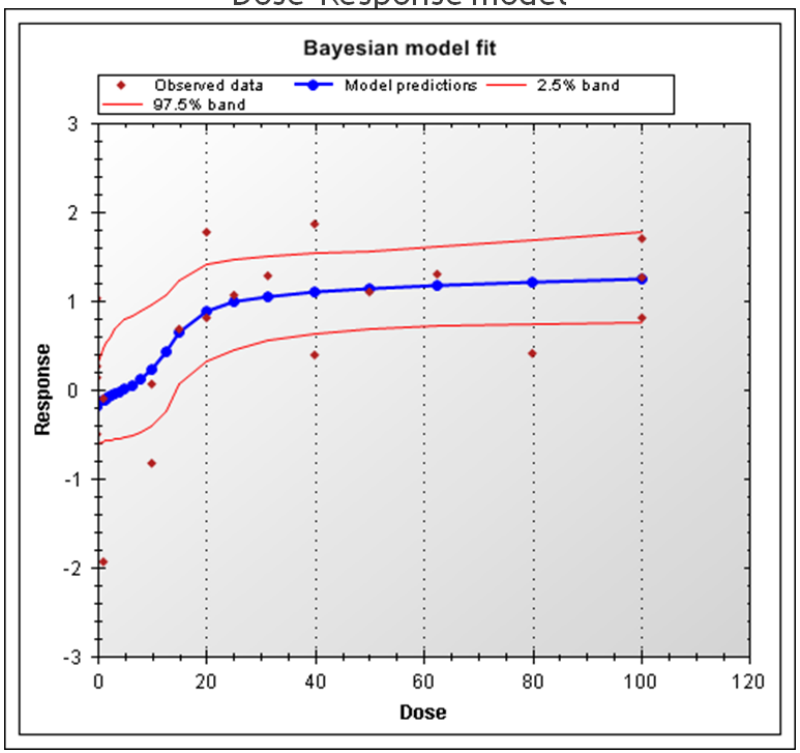
✓ Pr(Emax>0)=99%: Proof of concept achieved !

✗ CV(ED50)=49%

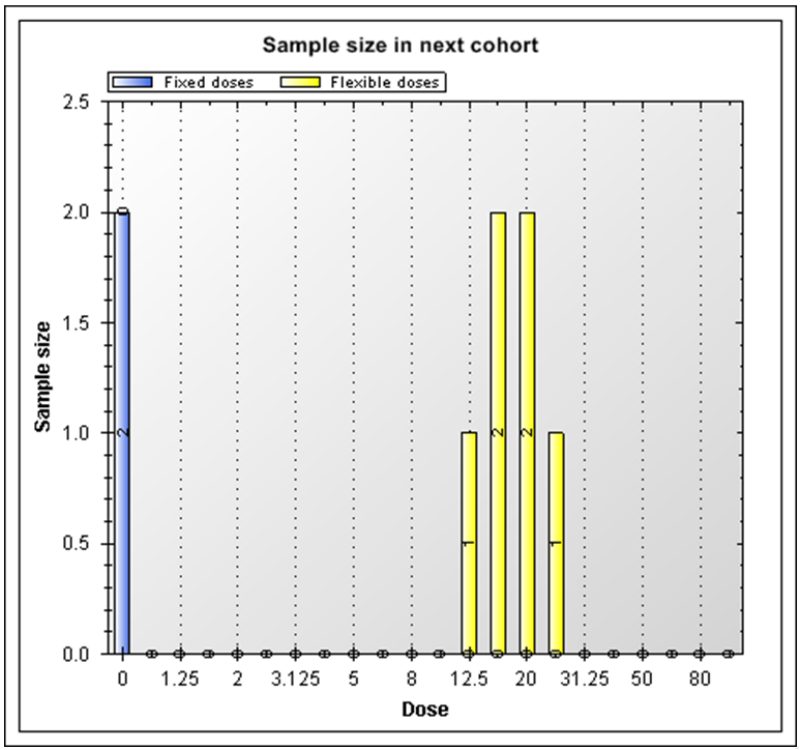
⇒ Continue to cohort 4

Adaptive Strategy : After Cohort 3 (N=24)

Dose-Response model



Choice of next doses for cohort 4



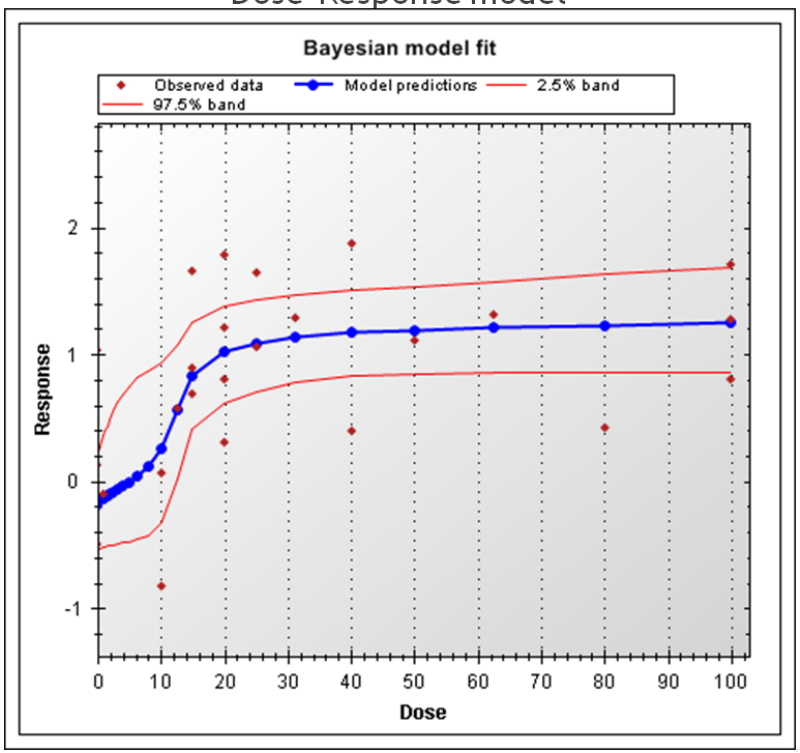
✓ Pr(Emax>0)=99%: Proof of concept achieved !

✗ CV(ED50)=49%

⇒ Continue to cohort 4

Adaptive Strategy : After Cohort 4 (N=32)

Dose-Response model



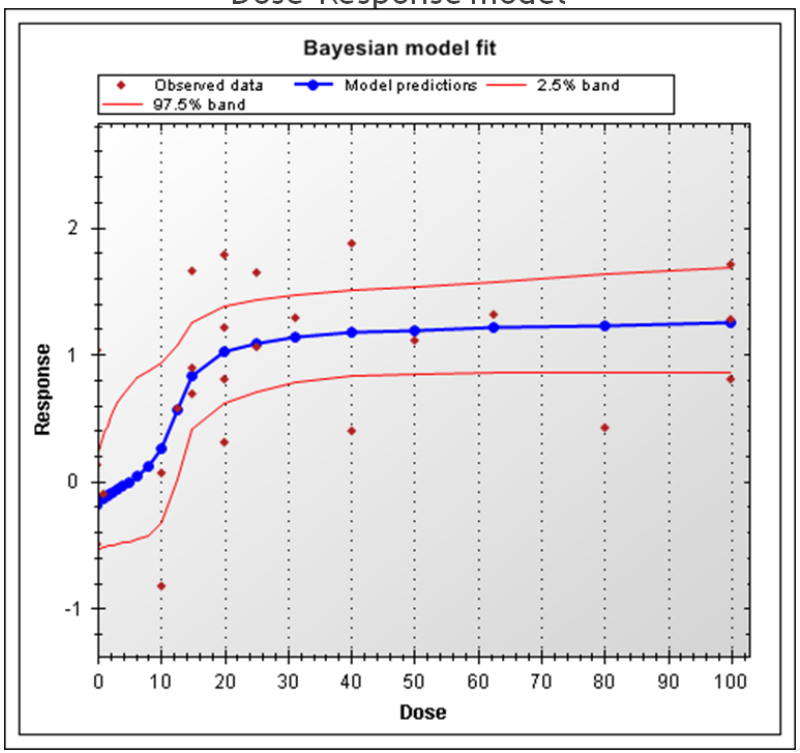
✓ Pr(Emax>0)=99%: Proof of concept achieved !

✓ CV(ED50)=30%

⇒ Stop trial after 32 subjects

Adaptive Strategy : After Cohort 4 (N=32)

Dose-Response model



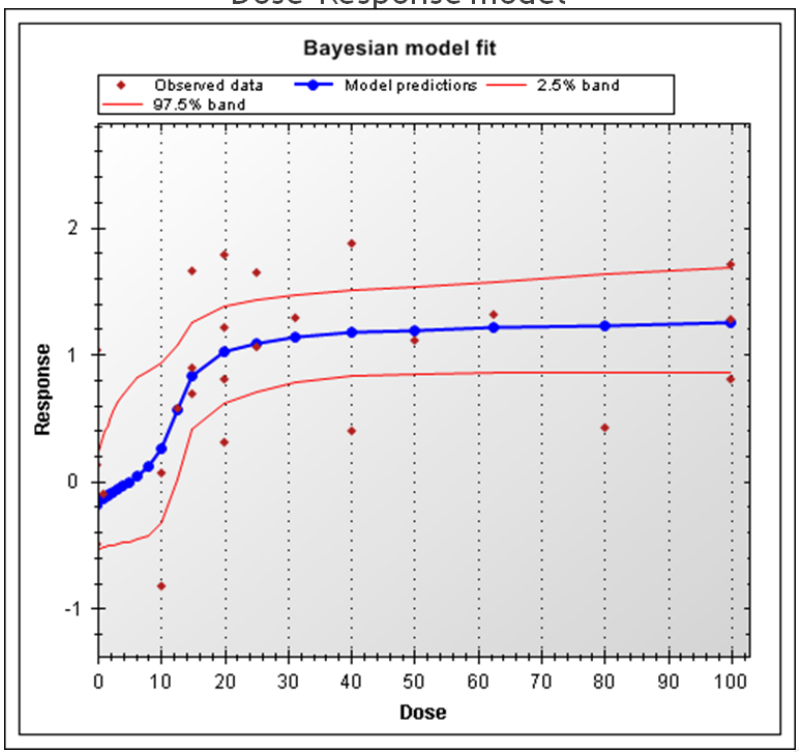
✓ Pr(Emax>0)=99%: Proof of concept achieved !

✓ CV(ED50)=30%

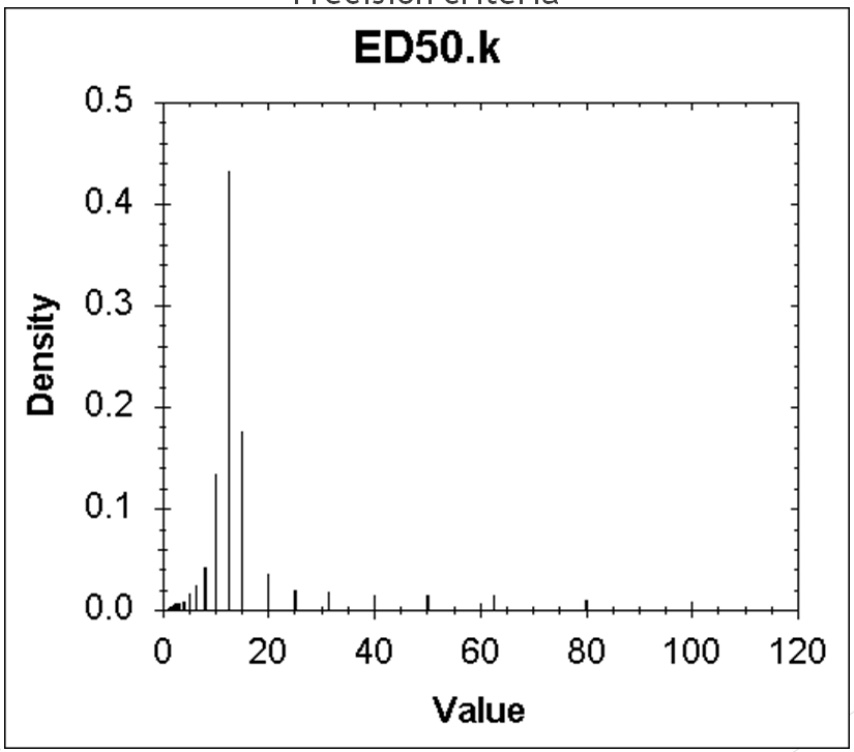
⇒ Stop trial after 32 subjects

Adaptive Strategy : After Cohort 4 (N=32)

Dose-Response model



Precision criteria



✓ Pr(Emax>0)=99%: Proof of concept achieved !

✓ CV(ED50)=30%

⇒ Stop trial after 32 subjects

Adaptive Strategy : Results

Emax model estimated with 32 subjects

- $E_0 = -0.2(0.2)$
- $E_{max} = 1.5(0.5)$
- $ED_{50} = 12.5mg(CV = 30\%)$
- $Hill = 4.5(6.6)$

Conclusion

- PoC & dose finding with limited sample size
- Adaptive dose selection toward informative range

Adaptive Strategy : Results

Emax model estimated with 32 subjects

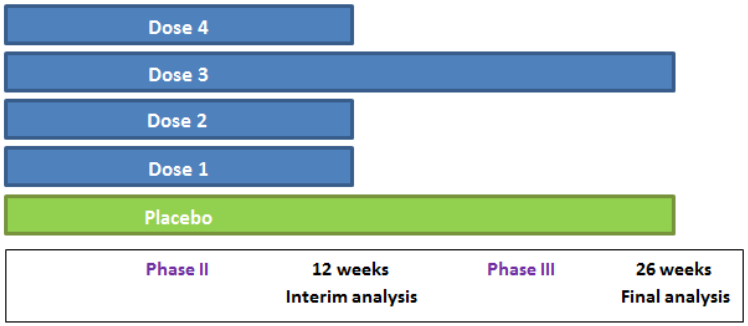
- $E_0 = -0.2(0.2)$
- $E_{max} = 1.5(0.5)$
- $ED_{50} = 12.5mg(CV = 30\%)$
- $Hill = 4.5(6.6)$

Conclusion

- PoC & dose finding with limited sample size
- Adaptive dose selection toward informative range

Seamless design (Gallo et al. 2006)

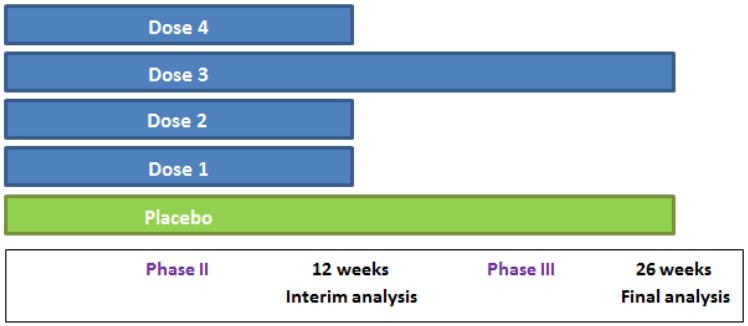
Design that addresses within a single trial objectives that are normally achieved through separate trials of clinical development



- Advantages
  - Reduce execution timelines
  - Increase final analysis with first stage data (if inferentially)
- Drawbacks
  - Longer protocol preparation; important planning
  - Limited/no sponsor involvment in dose-selection (IDMC), especially for confirmatory trials $\Rightarrow$ less flexible than traditional design

Seamless design (Gallo et al. 2006)

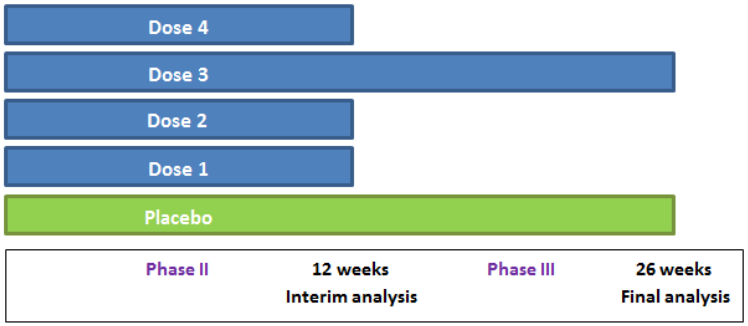
Design that addresses within a single trial objectives that are normally achieved through separate trials of clinical development



- Advantages
  - Reduce execution timelines
  - Increase final analysis with first stage data (if inferentially)
- Drawbacks
  - Longer protocol preparation; important planning
  - Limited/no sponsor involvment in dose-selection (IDMC), especially for confirmatory trials $\Rightarrow$ less flexible than traditional design

Seamless design (Gallo et al. 2006)

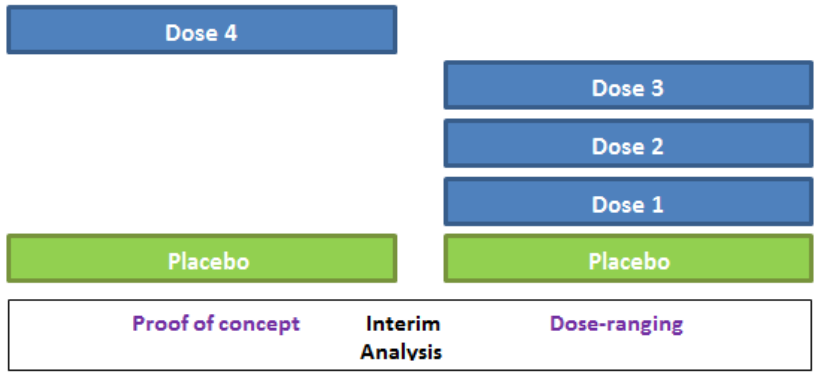
Design that addresses within a single trial objectives that are normally achieved through separate trials of clinical development



- Advantages
  - Reduce execution timelines
  - Increase final analysis with first stage data (if inferentially)
- Drawbacks
  - Longer protocol preparation; important planning
  - Limited/no sponsor involvment in dose-selection (IDMC), especially for confirmatory trials $\Rightarrow$ less flexible than traditional design

Seamless PoC design

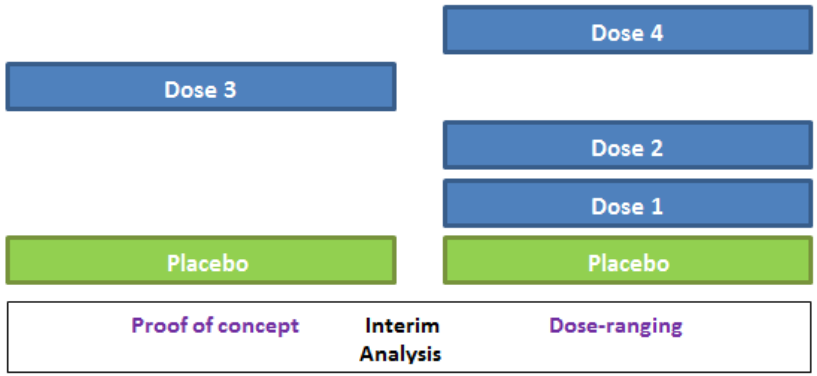
Figure : Stage 1=PoC (unbalanced)/Stage 2=dose-ranging or dose expansion



- Operationally seamless
  - More flexible : possibility to learn and adapt
  - Interim analysis can include sample size reassessment
- PoC can integrate historical data to improve decision making
- PoC can integrate futility
- 2 arms drawback $\Rightarrow$ a medium dose can be added at stage 1

Seamless PoC design

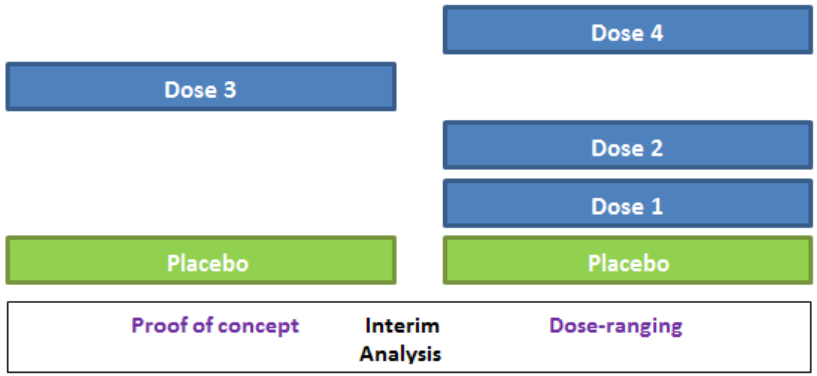
Figure : Stage 1=PoC (unbalanced)/Stage 2=dose-ranging or dose expansion



- Operationally seamless
  - More flexible : possibility to learn and adapt
  - Interim analysis can include sample size reassessment
- PoC can integrate historical data to improve decision making
- PoC can integrate futility
- 2 arms drawback $\Rightarrow$ a medium dose can be added at stage 1

Seamless PoC design

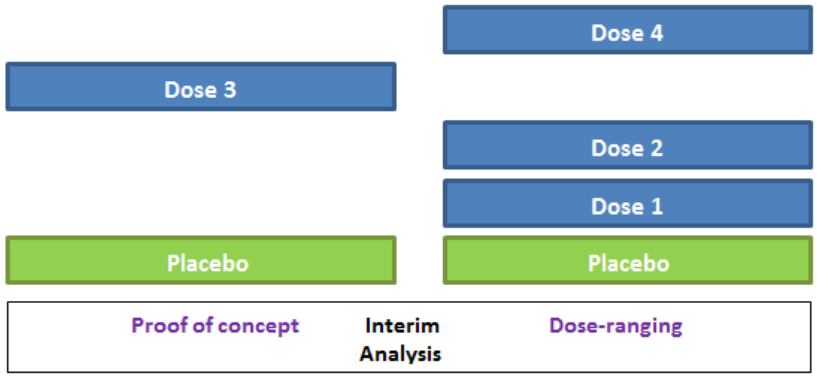
Figure : Stage 1=PoC (unbalanced)/Stage 2=dose-ranging or dose expansion



- Operationally seamless
  - More flexible : possibility to learn and adapt
  - Interim analysis can include sample size reassessment
- PoC can integrate historical data to improve decision making
- PoC can integrate futility
- 2 arms drawback $\Rightarrow$ a medium dose can be added at stage 1

Seamless PoC design

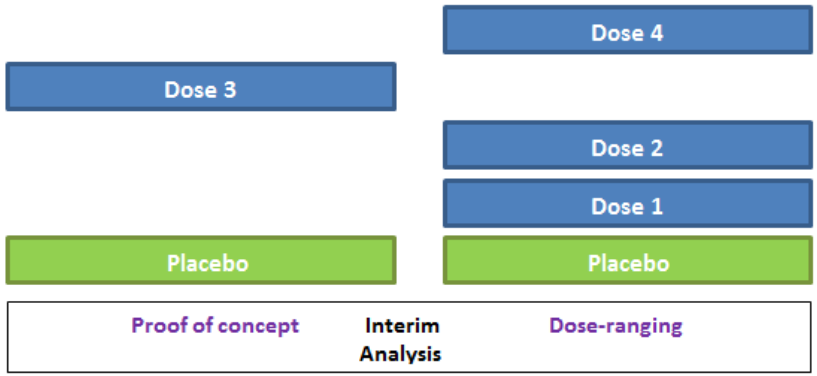
Figure : Stage 1=PoC (unbalanced)/Stage 2=dose-ranging or dose expansion



- Operationally seamless
  - More flexible : possibility to learn and adapt
  - Interim analysis can include sample size reassessment
- PoC can integrate historical data to improve decision making
- PoC can integrate futility
- 2 arms drawback ⇒ a medium dose can be added at stage 1

Seamless PoC design

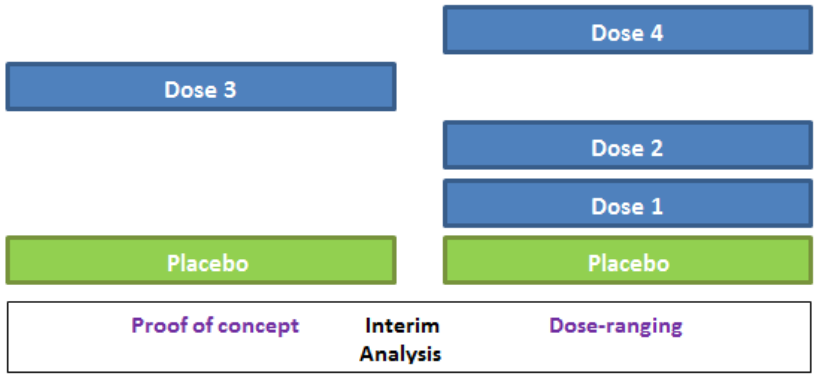
Figure : Stage 1=PoC (unbalanced)/Stage 2=dose-ranging or dose expansion



- Operationally seamless
  - More flexible : possibility to learn and adapt
  - Interim analysis can include sample size reassessment
- PoC can integrate historical data to improve decision making
- PoC can integrate futility
- 2 arms drawback ⇒ a medium dose can be added at stage 1

Seamless PoC design

Figure : Stage 1=PoC (unbalanced)/Stage 2=dose-ranging or dose expansion



- Operationally seamless
  - More flexible : possibility to learn and adapt
  - Interim analysis can include sample size reassessment
- PoC can integrate historical data to improve decision making
- PoC can integrate futility
- 2 arms drawback $\Rightarrow$ a medium dose can be added at stage 1

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to ↗ PoS or ↘ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to ↗ PoS or ↘ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to ↗ PoS or ↘ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to ↗ PoS or ↘ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to ↗ PoS or ↘ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to ↗ PoS or ↘ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to $\nearrow$ PoS or $\searrow$ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to $\nearrow$ PoS or $\searrow$ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to $\nearrow$ PoS or $\searrow$ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to $\nearrow$ PoS or $\searrow$ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

The Toolbox

- Quantifying the risks (Type I/II errors; PoS) is critical
  - Assurance can be useful for managing stakeholder expectations
  - Move from p-value to full posterior distribution
  - Early phase is a learning phase, dominated by estimation
- A bayesian integrated approach
  - Full posterior distribution derivation
  - Proper historical data integration (to $\nearrow$ PoS or $\searrow$ sample size)
  - Futility rule (posterior/predictive vs p-value/conditional power)
- Multi-arms PoC : consider adaptive analysis/designs
  - MCPmod, dose-ranging, seamless...
- Work upfront : Team discussions and simulations

Discussions

- **What types of dose-finding designs have you implemented? Were they fixed or adaptive ?**
- **What were the main advantages / disadvantages ?**