Boehringer Ingelheim

# When Futility Is Futile

## An economic case for more pragmatism in late phase futility stopping

James Bell - Elderbrook Solutions

Kevin Kunzmann - Boehringer Ingelheim

16 June 2026 | PSI Conference 2026, Belfast

*Life forward*

# Futility Analysis

- Late phase clinical trials are expensive ($100m ++)
- Despite early-phase successes, only 59% average phase III program success rate†

- Idea: Add interim futility analysis!
  - Stop trial if low chance of success

  → Avoid part of trial cost

- But this can be anything from a free lunch to a very expensive mistake

- We argue that, for economic reasons, futility analysis should not be designed with power losses greater than 1%

Boehringer Ingelheim

†Wong CH, Siah KW, Lo AW. Estimation of clinical trial success rates and related parameters. *Biostatistics*. 2018;20(2):273-286

# Marginal Power

- Trial planning requires **marginal operating characteristics**
  - Statistical properties of design itself
  - Explain why design decisions  are made and decision criteria are set

- Futility decision criteria part of trial design, set at planning stage → need impact on marginal operating characteristics

- Futility stopping causes trial failure so reduces both marginal power and type I error
  - Marginal power: power of the designed trial (i.e. including any futility analyses)
  - Regulators require non-binding futility and fix type I error requirements (i.e. cannot increase $\alpha$ to compensate)

- Consequences:
  - Positive: **probability of early stopping → lowered expected trial cost**
  - Negative: **reduces marginal power → reduces expected future revenues**

Boehringer Ingelheim

# The fundamental trade-off

- **Too conservative**: Low stopping chance (even under null hypothesis!)
  - **Low expected savings** – worth it?
  - **Low/no impact on marginal power**
- **Too aggressive**: High stopping chance (even under alternative hypothesis!)
  - Chance of '**false stop**'; wrongly declaring futility
  - **High cost savings** but high expected loss from **forfeiting revenue** from effective drugs
  - **Substantial impact on marginal power**; needs to be adequately represented in CTP

Boehringer Ingelheim

# When is a trial viable?

- We can put together a simple† economic model for a futility analysis
  - **Benefits:**
    - Lowered expected trial costs = Pr*(stop)* * *cost savings when stopping* (**S**)
  - **Costs:**
    - *Loss in expected revenues =* Pr*(incorrect stop)* * *anticipated revenue if successful (**R**)*

- Economic value of a futility analysis is then:
  - $\mathbf{Pr}(\boldsymbol{stop}) \cdot \boldsymbol{S} - \mathbf{Pr}(\boldsymbol{incorrect\ stop}) \cdot \boldsymbol{R}$

- For a futility analysis to add economic value:
  - $\mathbf{Pr}(\boldsymbol{stop}) \cdot \boldsymbol{S} - \mathbf{Pr}(\boldsymbol{incorrect\ stop}) \cdot \boldsymbol{R} > \mathbf{0}$
  - Alternatively: $\boldsymbol{Pr(incorrect\ stop \mid stop)} < \frac{\boldsymbol{S}}{\boldsymbol{R}}$

Boehringer Ingelheim

†More complex approaches are also available, see e.g. Shariati N, Financially Calibrated Risk-Scale:A Proposed Futility Design Framework to Enhance Portfolio-Level Profitability and Performance; presentation at ISCB 2025

# S/R and Trial Costs

- Futility analysis savings often easiest to quantify relative to overall anticipated trial costs (***C***)
  - ***S/C*** is proportion of trial costs recoverable with a futility stop
    - Readily estimated at early trial planning stage by trial teams
    - Linked to (standardised) interim information fraction ($\tau$) : Earlier interim = higher ***S/C***, *lower* $\tau$
    - Many costs are front loaded (trial setup, site openings, patient recruitment etc), so ***S/C < (1- $\tau$)***

- Financial opportunity also can be measured relative to ***C***
  - ***R/C*** measures attractiveness of investment if successful, often estimated & used by finance functions
    - Either available as part of trial feasibility or industry/corporate standard assumptions available
    - Require in Net Present Value terms, i.e. future revenues discounted to same timeframe as trial costs incurred

- $\frac{S/C}{R/C} = S/R$ savings from stopping as a fraction of the financial opportunity

Boehringer Ingelheim

# Statistical Model for Interim Stopping

- To evaluate stopping probabilities, adopt a statistical model with standardised times, effect sizes etc:

- Interim and final test statistic
  - Assume: **bivariate Normal distribution** of interim and final test statistics
  - Z scores at interim ($Z_\tau$) and final analyses ($Z$)
  - $\tau$ = information fraction at interim

$$(Z_\tau, Z) \sim \pi \cdot MVN_2\left(\mu_{1,2}, \Sigma_\tau\right) + (1-\pi) \cdot MVN_2\left((0,0), \Sigma_\tau\right)$$

- $\mu_{1,2} = \left(\Phi^{-1}(1-\beta) + \Phi^{-1}(1-\alpha)\right) \cdot (\sqrt{\tau}, 1)$
- $\Sigma_\tau = \begin{pmatrix} 1 & \sqrt{\tau} \\ \sqrt{\tau} & 1 \end{pmatrix}$

- Assumptions on true treatment effect (a 'prior')
  - Assume probability of **point null and point alternative hypotheses**
    - Frequentist alignment: "Probability the drug does / does not work"

  → **Mixture distribution with $\pi = \Pr(H_1)$**

# Models for Stopping Probabilities

- Assume a Z-statistic boundary ('f'), stop for futility if $Z_\tau < f$
- Stopping probabilities can then be expressed as:
  - $\Pr(stop) = \pi \cdot \Phi(x - \mu_1) + (1 - \pi) \cdot \Phi(x)$
  - $\Pr(incorrect\ stop) =$
    - $\int_{-\infty}^{f} \int_{\Phi^{-1}(1-\alpha)}^{\infty} \pi \cdot \varphi_2\left((z_\tau, z); \mu_{1,2}, \Sigma_\tau\right) dz\, dz_\tau + \int_{-\infty}^{f} \int_{\Phi^{-1}(1-\alpha)}^{\infty} (1 - \pi) \cdot \varphi_2\left((z_\tau, z); (0,0), \Sigma_\tau\right) dz\, dz_\tau$

Loss of marginal power

Loss of type I error; $< \alpha$

Note: $\varphi_2(z; \mu, \Sigma)$= PDF of a bivariate Normal distribution with means $\mu$ and covariance $\Sigma$

# Economic Approaches for Informing Choice of Futility Boundaries

- **Economic viability when $\boldsymbol{Pr(stop) \cdot S - Pr(incorrect\ stop) \cdot R > 0}$**
  - Solve numerically $\frac{S}{R} - \frac{\Pr(incorrect\ stop)}{\Pr(stop)} = 0$ to find maximum viable standardised futility boundary $f_{max}$

- **Economic optimality when $\boldsymbol{Pr(stop) \cdot S - Pr(incorrect\ stop) \cdot R}$ is maximised**
  - Solve numerically $\frac{\delta}{\delta f}\left(\frac{S}{R}.\Pr(stop) - \Pr(incorrect\ stop)\right) = 0$ to find economically optimal futility boundary $f_{opt}$
    - Note: $\frac{S}{R}.\Pr(stop) - \Pr(incorrect\ stop)$ is expected benefit of futility analysis as proportion of expected revenue, *R*

- Corresponding properties can be readily calculated from boundary, e.g. **Power loss**, Pr(stop)

- For phase III trials with 90% power, 1-sided 2.5% type I error, only $\tau$, $\pi$ and $S/R$ determine futility boundary properties

Boehringer Ingelheim

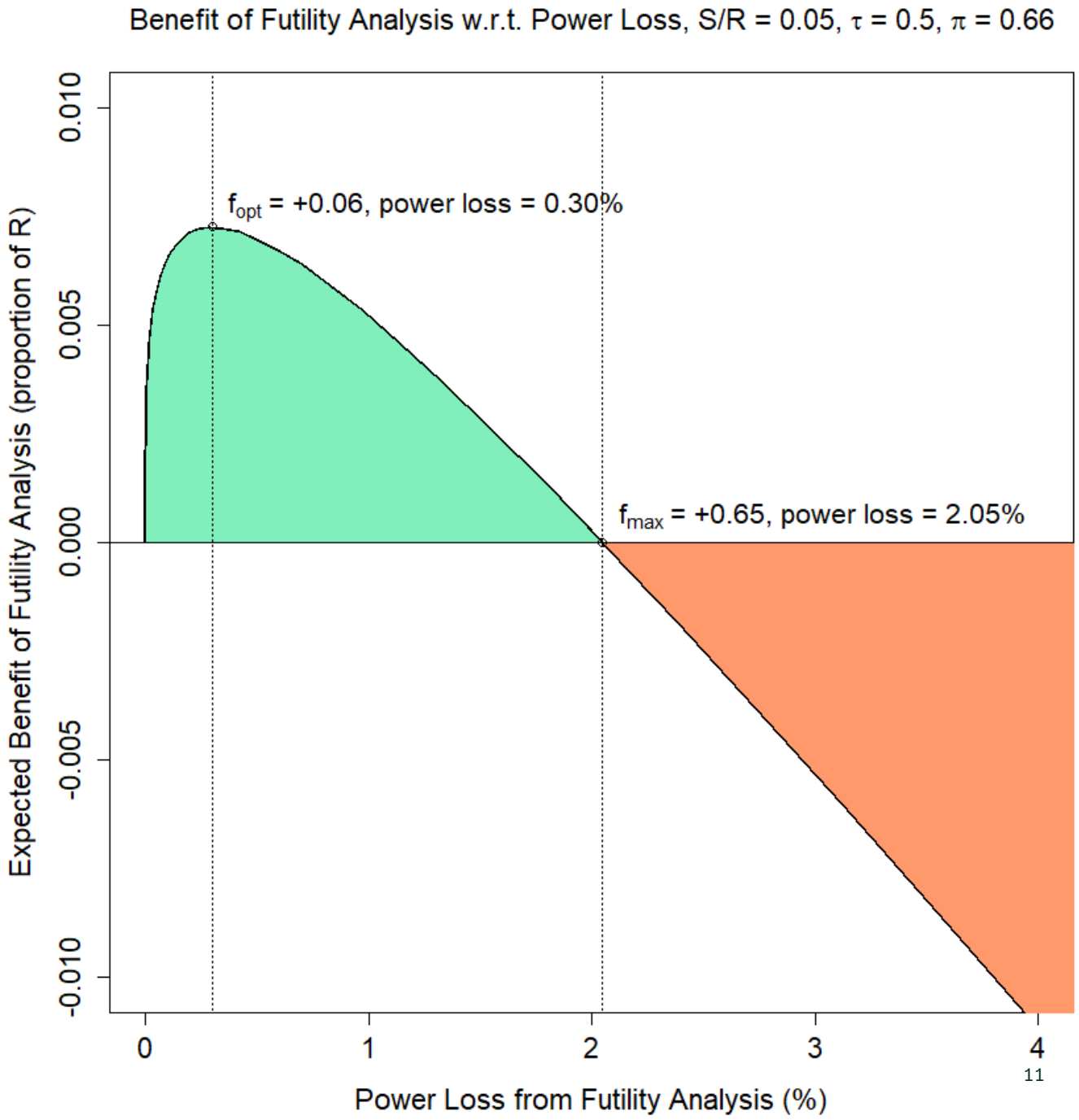
# Economic Value vs Power Loss

- Example plot showing benefit of futility analysis vs power loss from different bounds

- Assumes
  - $\tau = 0.5, \pi = 0.66$
  - $\frac{S}{R} = 0.05,\ e.g.\ S = 0.25,\ R = 5$

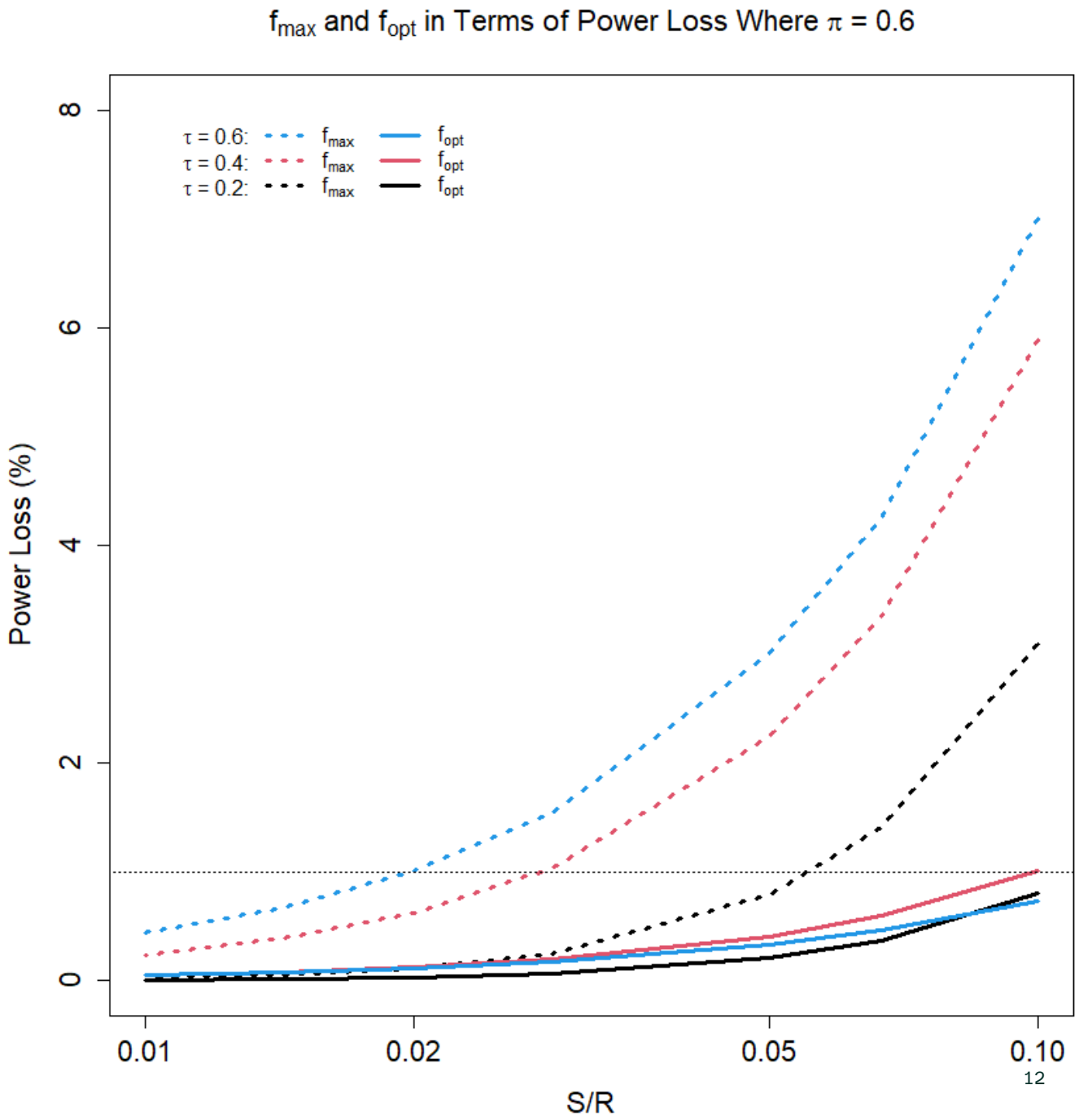
- Benefits optimised at very low power losses (0.3%)
  - 'zero boundary' is essentially optimal
  - Power loss > 2% loses money

- Shape of curve typical:
  - Steep gains in value with minimal power loss
  - Linear decline in value beyond optimal point



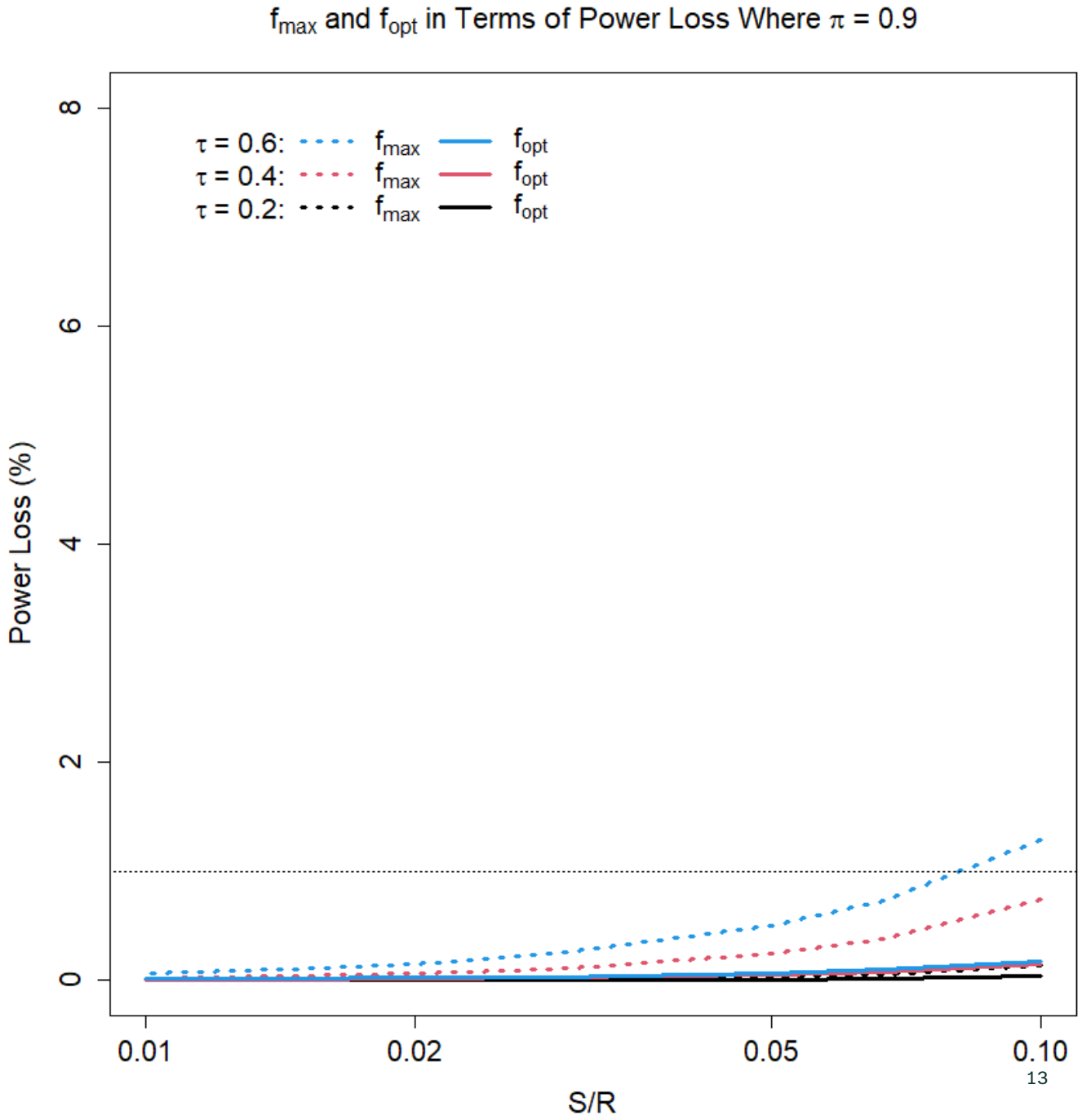
# Survey of Parameter Space

- Can screen many potential designs for reasonable futility boundaries and power loss
- $f_{opt}, f_{max}$ and power losses are highly sensitive to *S/R* and $\pi$
- $\tau$ has less effect…
  - … and raising $\tau$ also lowers *S/R*
- For high-risk phase III trials ($\pi = 0.6$), power losses of 2-4% may remain economically viable if relative savings are high.
- However, optimal futility analyses still result in 0.01% to 1% power loss



# Survey of Parameter Space

- For low-risk phase III trials ($\pi = 0.9$), power losses above 1% are not economically viable
- Optimal futility power loss is typically < 0.2%



# The 1% Rule of Thumb

- **"Futility boundaries for phase III trials that lose *more* than 1% marginal power are hard to justify economically."**
- **Justification**
  - **Broad applicability** - Almost all parameter combinations produce power loss < 1%
  - **Safe** - Limits expected economic loss if compound works
  - **Efficient** - Conservative stopping captures most of the potential economic benefits
  - **Simple** - Small impact on power avoids need to adjust power / sample size calculations
- **Caveats:**
  - 1% is an upper limit, not a target
    - 1% is often above $f_{max}$; may need much lower (e.g. 0.1%) for economic optimality
  - Overpowered trials or large effect requirements may still benefit from more aggressive futility stopping
  - Might not apply when *a priori* confidence in compound is very low (speculative phase III)

Boehringer Ingelheim

# Conclusions

- Futility analyses are important for derisking late phase trials
- Economic perspective can help to navigate the fundamental trade off between cost saving and preserving expected revenue
- In typical phase III trials power loss from futility analysis should be less than 1%
- Allows simpler trial design process by considering futility analysis **after** core trial design is aligned, due to negligible impact on power

Boehringer Ingelheim