

SMART-Vent: A Bayesian Adaptive Platform Trial for Evaluating Mechanical Ventilation Strategies Using Patient-Centered Endpoints in Heterogeneous ICU Populations

Danila Azzolina

Biostatistics and Clinical Trial Methodology Unit,
Clinical Research Center DEMeTra,
Department of Translational Medical Science,
University of Naples Federico II, Naples, Italy



Background: Challenges with Conventional ICU Trials

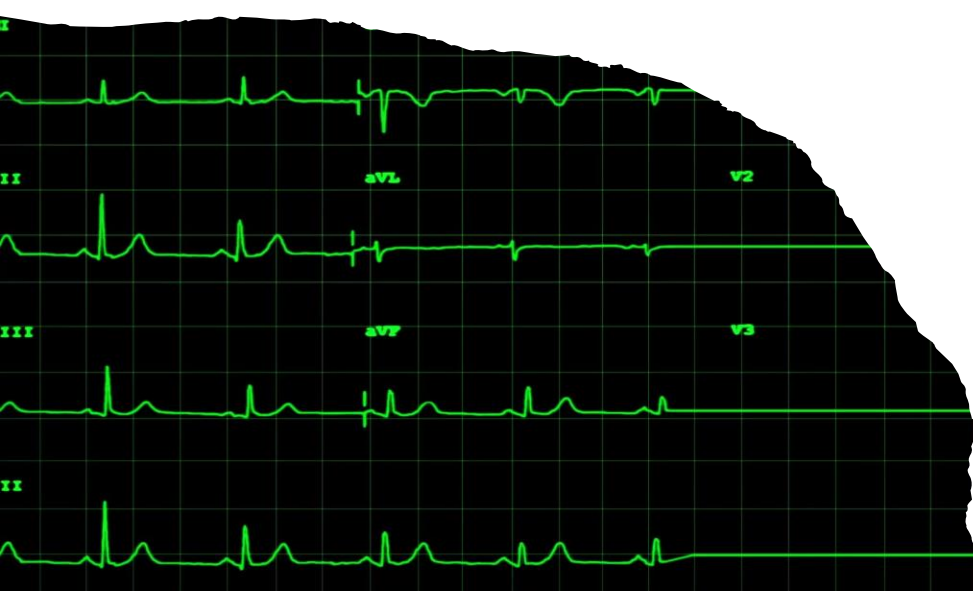
Why a Bayesian Adaptive SMART Platform for ICU Trials?

ICU populations are heterogeneous, making fixed-design RCTs underpowered and slow to yield actionable results (Pocock & Stone, JAMA, 2016).

Static 1:1 randomization fails to learn from accumulating data, while binary endpoints like 28-day mortality overlook recovery trajectories (van der Poll et al., Intensive Care Med, 2021).

Frequent non-response and treatment crossover compromise classical analysis validity (Rochweg et al., Crit Care, 2022).

Study design



Design

- Bayesian adaptive **platform + SMART** framework

Patient Population

- Adult ICU patients on invasive ventilation ≥ 24 h
- Stratified: **ARDS vs non-ARDS**
- Exclusions: chronic ventilation, expected death < 48 h, pregnancy, no consent

Primary Outcome

- **Utility-weighted composite:**
integrates survival, ventilator-free days, and recovery (PROMIS-29 / EQ-5D)

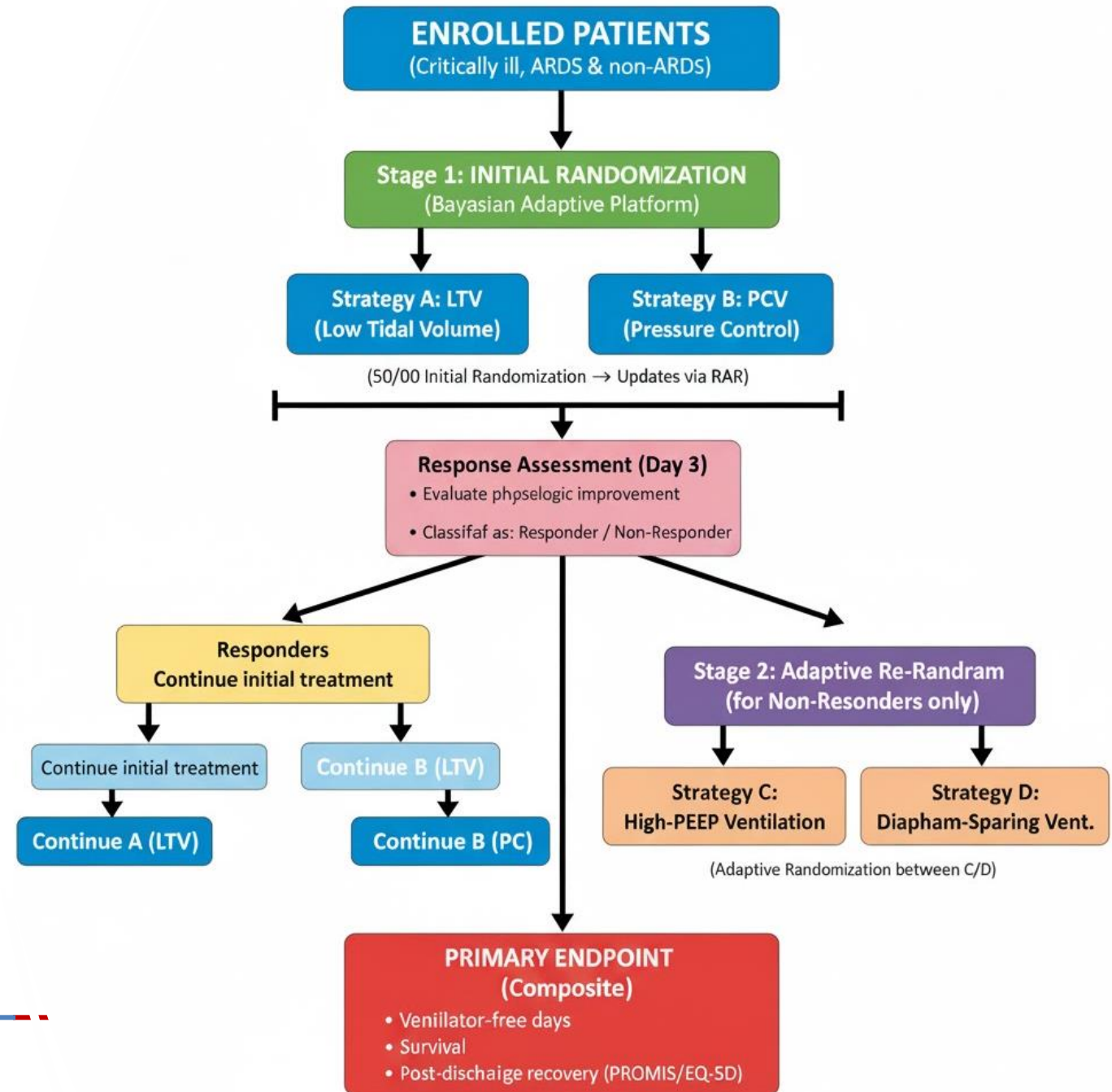
Clinical Context

- Heterogeneous ICU population \rightarrow limited fixed-trial efficiency
- SMART-Vent enables **adaptive learning** and **patient-centered evaluation** of ventilation strategies



SMART-Vent Trial flowchart

- Overall Objective
- Optimize individualized ventilation pathways for critical care patients.



Key Advantages of the SMART-Vent Design

| Feature | Rationale and Reference |
|--|--|
| Response-Adaptive Randomization (RAR) | Allocation probabilities update with accumulating data, improving efficiency and ethical balance (Berry et al., Clin Trials 2010) |
| SMART Logic | Sequential re-randomization reflects ICU decision-making and evaluates full treatment pathways (Kidwell et al., Crit Care Med 2020) |
| Bayesian Hierarchical Borrowing | Partial pooling across ARDS / non-ARDS preserves subgroup inference while sharing information (Schmidli et al., Pharm Stat 2014) |
| Utility-Based Composite Endpoint | Combines ventilator-free days, survival, and recovery (PROMIS-29 / EQ-5D) for patient-centered outcomes (Harhay et al., AJRCCM 2019) |
| Ethical & Operational Gains | More patients benefit during the trial; early stopping and smaller sample sizes enhance feasibility (Saville & Berry, Nat Med 2016) |

Definition for Utility in Critical Care

A *utility* is a **composite, patient-centered numerical index** that integrates multiple outcome domains into a single continuous measure reflecting *overall clinical benefit*:

1. survival,
2. ventilator dependence,
3. health-related quality of life

It represents the **expected value of patient well-being** expressed as:

$$U = \sum_{j=1}^J w_j Y_j$$

Design Weights

- Weights (w_1, w_2, w_3) will be **derived from patient and caregiver preference elicitation** using validated health-utility instruments.
- Preference weights will be estimated from **EQ-5D value sets** based on **time-trade-off (TTO)** and **standard-gamble (SG)** methods and rescaled so that their sum equals 1.
- Typical base values are $w_1 = 0.4, w_2 = 0.4, w_3 = 0.2$, reflecting comparable importance of survival and ventilator independence with moderate weight on recovery quality.
- Sensitivity analyses will evaluate the robustness of trial conclusions under alternative weight configurations.



Model

- Let $s \in \{\text{ARDS, non-ARDS}\}$ denote the patient subgroup and $g \in \mathcal{G}$ a final strategy label, for example $g \in \{A \rightarrow \text{Continue}, B \rightarrow \text{Continue}, A \rightarrow C, A \rightarrow D, B \rightarrow C, B \rightarrow D\}$
- The hierarchical Gaussian model (fixed strategy effects with subgroup-specific random slopes) induces a posterior distribution for each Strategy \times Subgroup mean utility:

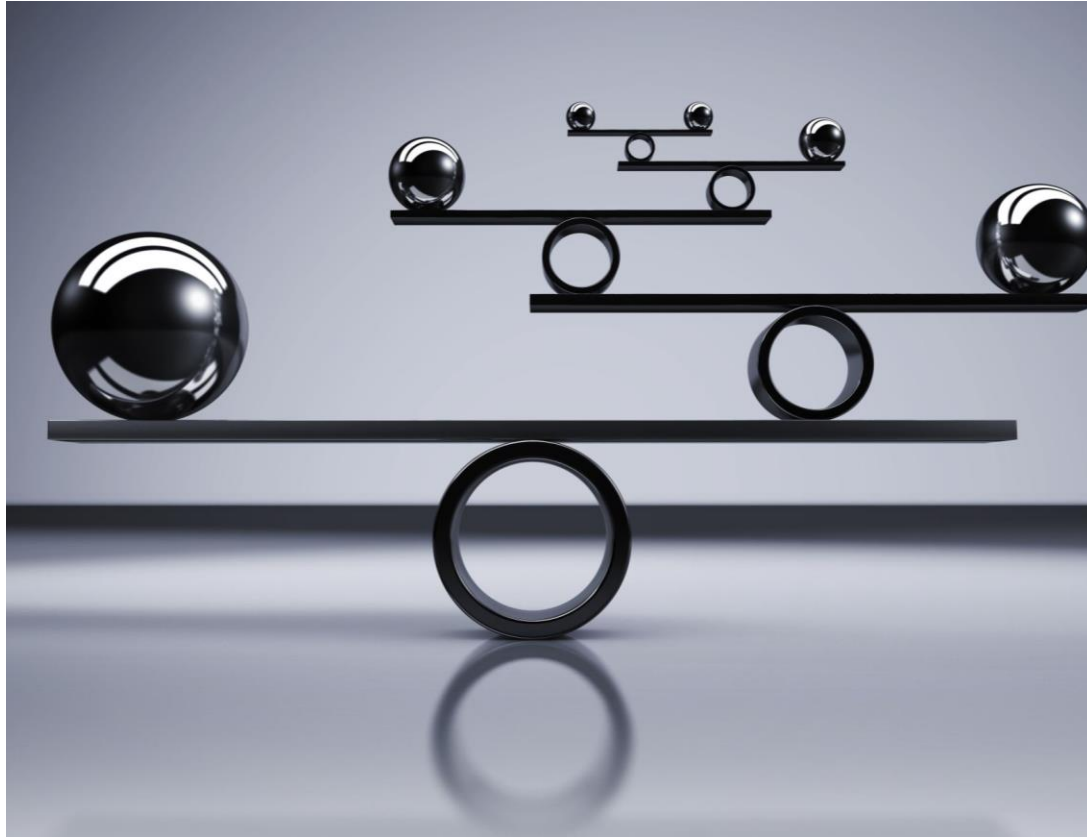
$$U_{i,g,s} = \sum_g \beta_g \text{ Strategy } =g + \sum_s b_{g,s} \text{ Strategy } =g, \text{ Subgroup } =s + \varepsilon_{i,g,s}, \varepsilon_{i,g,s} \sim \mathcal{N}(0, \sigma^2)$$

Priors

- **Fixed effects:** “Diffuse Normal around 0 → weak prior, broad learning.”
- **Random SD (τ):** “Half-t → encourages pooling, permits heterogeneity.”
- **Residual SD (σ):** “Half-t with smaller scale → sensible noise control.”
- **Random effect (b):** “Heavy-tailed Normal mixture → shrinkage with robustness.”

| Parameter | Role | Prior Distribution | Scale Used | Interpretation |
|------------|-----------------------------|-----------------------------|--------------|-------------------------------|
| β_g | Fixed (Strategy effect) | Normal(0, 25 ²) | 2.5 × s_y | Broad, weakly informative |
| $b_{g,s}$ | Random (Subgroup deviation) | Normal(0, τ_g) | hierarchical | Allows subgroup heterogeneity |
| τ_g | SD of subgroup effects | Half- $t_3(0, 25)$ | moderate | Shrinkage across subgroups |
| σ^2 | Residual SD | Half- $t_3(0, 10)$ | narrow | Limits extreme outcome noise |

Probability of being best (within subgroup)



- For each draw m and subgroup s , find the best strategy:

$$g_s^{*(m)} = \arg \max_{g \in \mathcal{G}} \mu_{g,s}^{(m)}$$

- Empirical **probability of being best**:

$$p_{\text{best}}(g, s) = \frac{1}{M} \sum_{m=1}^M \mathbb{1} \{ g_s^{*(m)} = g \}$$

Stages

- Stage-1 (A vs B)

$$\mathcal{G}_A^{\text{cont}} = \{A \rightarrow \text{Continue}\}, \mathcal{G}_B^{\text{cont}} = \{B \rightarrow \text{Continue}\}$$

- Stage-2 (C vs D)

$$\mathcal{G}_C^{\text{NR}} = \{A \rightarrow C, B \rightarrow C\}, \mathcal{G}_D^{\text{NR}} = \{A \rightarrow D, B \rightarrow D\}$$



RAR update

- Stage-1 allocation (A vs B):

$$\tilde{\pi}_A^{(t+1)}(s) = \sum_{g \in \mathcal{G}_A^{\text{cont}}} p_{\text{best}}(g, s), \tilde{\pi}_B^{(t+1)}(s) = \sum_{g \in \mathcal{G}_B^{\text{cont}}} p_{\text{best}}(g, s).$$
$$\pi_k^{(t+1)}(s) = \frac{\max\{p_{\min}, \tilde{\pi}_k^{(t+1)}(s)\}}{\max\{p_{\min}, \tilde{\pi}_A^{(t+1)}(s)\} + \max\{p_{\min}, \tilde{\pi}_B^{(t+1)}(s)\}}, k \in \{A, B\}.$$

- Stage-2 allocation (C vs D, for non-responders):

$$\pi_\ell^{(t+1)}(s) = \frac{\max\{p_{\min}, \tilde{\pi}_\ell^{(t+1)}(s)\}}{\max\{p_{\min}, \tilde{\pi}_C^{(t+1)}(s)\} + \max\{p_{\min}, \tilde{\pi}_D^{(t+1)}(s)\}}, \ell \in \{C, D\}$$

Dynamic Borrowing

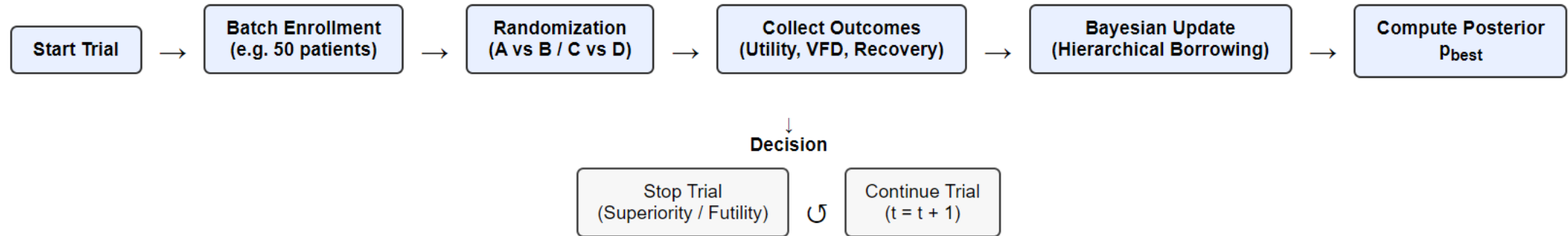
θ_g = global mean

$$\mu_{g,s} \sim \mathcal{N}(\theta_g, \tau_g^2)$$

τ_g controls how strongly subgroups share information.

- As the trial progresses, posterior updating of τ_g drives **dynamic borrowing**:
- If subgroup responses (ARDS vs non-ARDS) are similar $\rightarrow \tau_g$ shrinks, increasing information sharing.
- If subgroup effects diverge $\rightarrow \tau_g$ grows, reducing borrowing.
- Thus, borrowing is **data-driven**, adapting automatically to emerging heterogeneity.

Flowchart



☞ If no stopping rule is met, return to **Batch Enrollment** for the next batch.

Superiority $\max_{g,s} p_{\text{best}}(g, s) > 0.95$

Arm dropping $p_{\text{best}}(g, s) < 0.01$

Simulation Settings and scenarios

- Initial Enrollment: The simulation begins with an initial cohort of 100 randomized equally across all active strategies (A, B, C, D) to ensure unbiased starting estimates.
- Sequential Batch Analyses: After the initial phase, adaptive updates occur in batches of 50 newly enrolled patients.
- Following each batch, the Bayesian hierarchical model is refitted, posterior probabilities are recalculated, and response-adaptive randomization (RAR) probabilities are updated for subsequent allocations.
- Adaptive Evolution: As information accumulates, allocation shifts toward the most promising strategies while maintaining minimal allocation probabilities ($p_{\min}=0.10$) to preserve exploration.

| Parameter | Category | Value |
|-------------------------------|--------------|-------|
| Base (β_0) | — | 65 |
| Strategy effect (β_g) | A | 2 |
| | B | 4 |
| | C | -2 |
| | D | 1 |
| Group effect β_g | ARDS | -5 |
| | non-ARDS | 0 |
| Group* Strategy $b_{g,s}$ | A – ARDS | 0 |
| | A – non-ARDS | 2 |
| | B – ARDS | 0 |
| | B – non-ARDS | -1 |
| | C – ARDS | 0 |
| | C – non-ARDS | 0 |
| | D – ARDS | 0 |
| | D – non-ARDS | 0 |
| Noise (σ) | — | 10 |

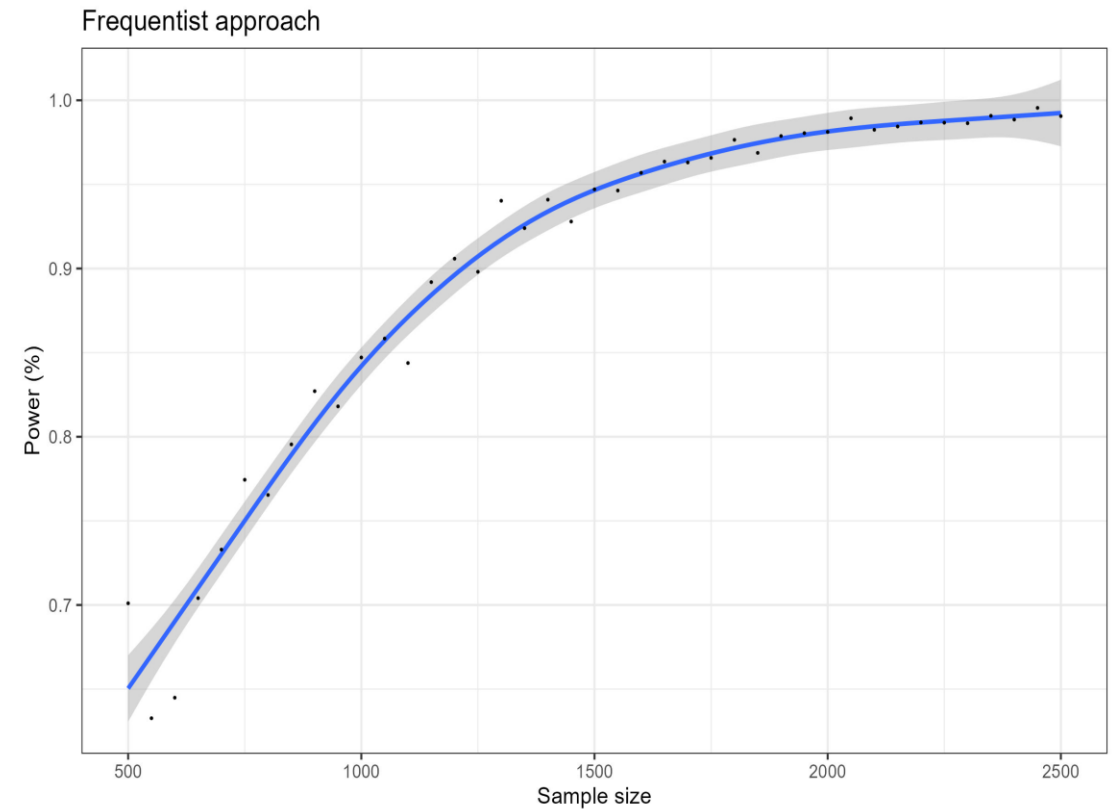
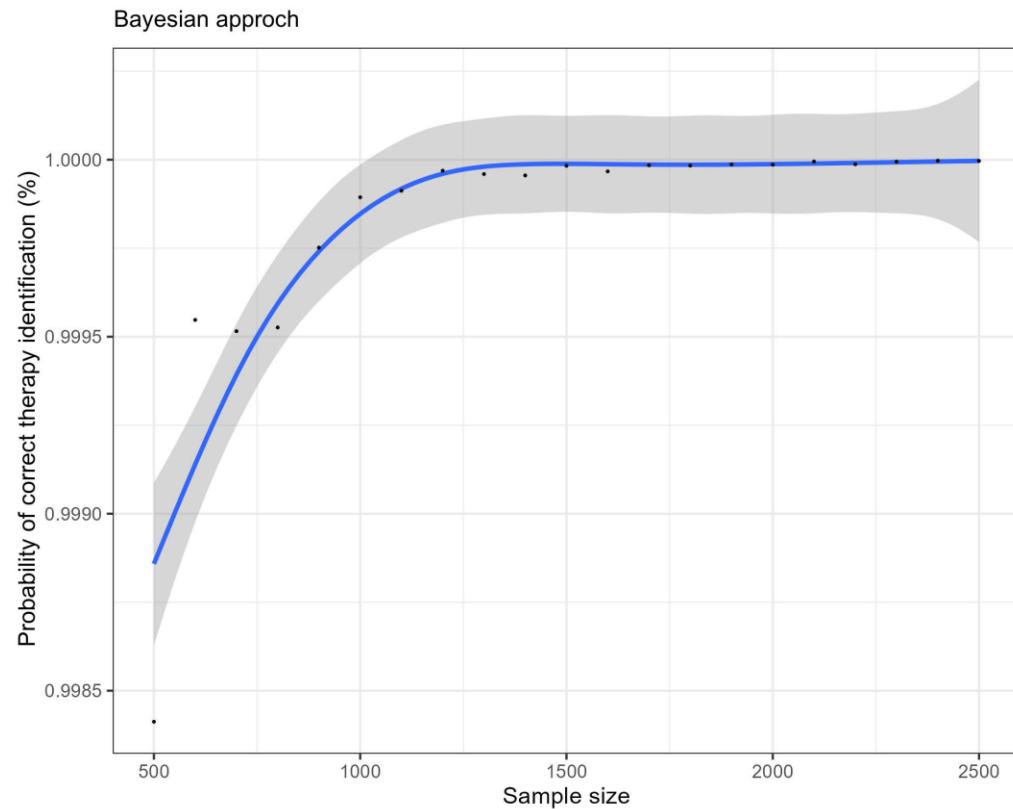
Reference values

| Parameter Type | Range | Clinical Interpretation | Supporting References |
|--|---------------------------------|-----------------------------------|---|
| Base Utility ($\beta_0 = 65$) | EQ-5D ≈ 0.65 – 0.75 | Typical ICU survivor health state | Herridge NEJM 2011; Needham AJRCCM 2013 |
| Strategy effects β_g (± 2 – 4) | 2–4 pts (≈ 3 – 5%) | Clinically relevant difference | PROSEVA NEJM 2013; Tonelli ICM 2020 |
| Subgroup penalty (ARDS = -5) | 5 pts (≈ 7 – 8%) | Lower recovery in ARDS | Bellani JAMA 2016 |
| Noise ($\sigma = 10$) | — | Inter-patient variability | Dinglas CCM 2018 |
| $b_{g,s}$ | ± 2 | Small plausible modifiers | Harhay AJRCCM 2019 |

Operative characteristics (I)

Figure 1: Probability of correctly identifying the optimal treatment strategy within at least one patient subgroup, as a function of the total sample size.

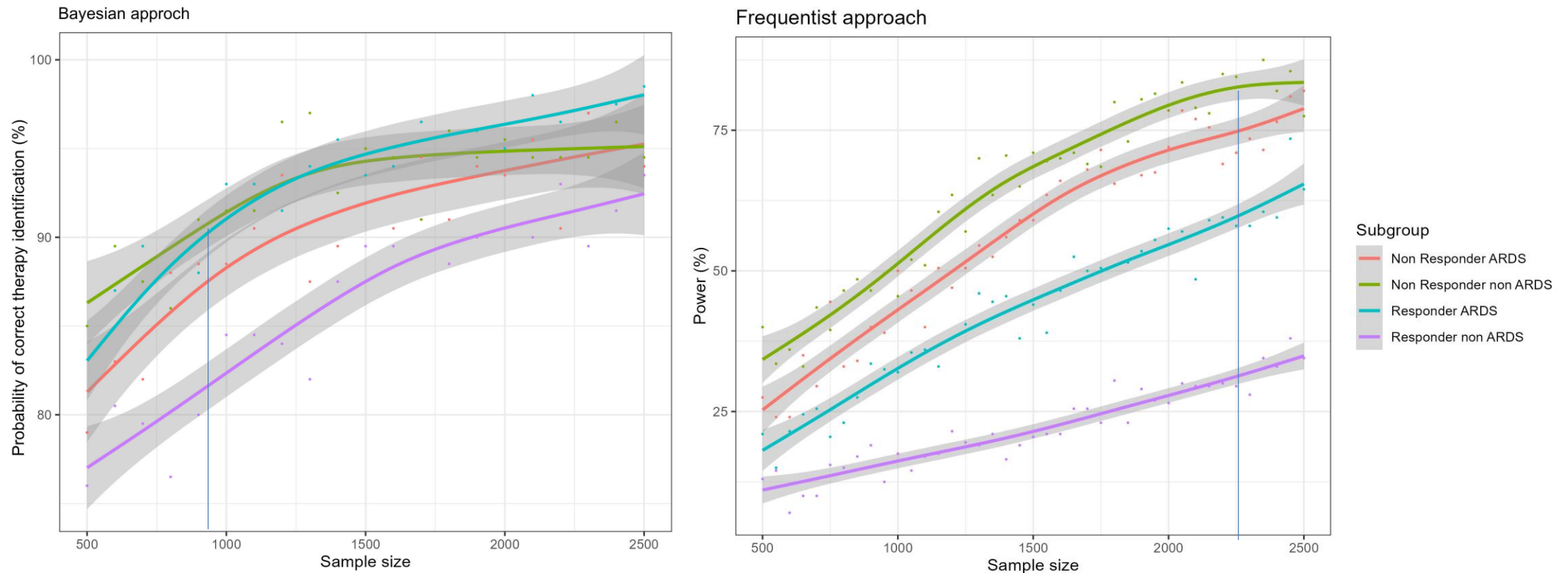
The comparison is made between two analytical frameworks: the SMART (Sequential Multiple Assignment Randomized Trial) Bayesian approach and the classical frequentist approach.



Operative characteristics (II)

Figure 2: Probability of correctly identifying the optimal treatment strategy for each patient subgroup separately, as a function of the total sample size.

Please note that the sample size shown on the x-axis refers to the total number of patients, not to the size of the individual subgroups.



Operative characteristics (III)

Table 2: Probability of correctly identifying the optimal treatment strategy within each subgroup, according to both the SMART (Bayesian) and classical frequentist approaches.

| Sample size | Group | Power (%) | Probability of correct therapy identification(%) |
|-------------|------------------------|-----------|--|
| 500 | Responder ARDS | 21.0 | 79.5 |
| | Responder non ARDS | 13.0 | 71.3 |
| | Non Responder ARDS | 27.5 | 82 |
| | Non Responder non ARDS | 40.0 | 80 |
| | Cumulative | 70.0 | 99.8 |
| 1000 | Responder ARDS | 32.0 | 91.5 |
| | Responder non ARDS | 17.5 | 81.5 |
| | Non Responder ARDS | 50.0 | 90 |
| | Non Responder non ARDS | 45.5 | 90 |
| | Cumulative | 84.7 | 99.9 |
| 1500 | Responder ARDS | 44.0 | 92 |
| | Responder non ARDS | 20.5 | 87.5 |
| | Non Responder ARDS | 59.0 | 94 |
| | Non Responder non ARDS | 71.0 | 94 |
| | Cumulative | 94.7 | 99.9 |
| 2000 | Responder ARDS | 57.5 | 94 |
| | Responder non ARDS | 26.5 | 88.5 |
| | Non Responder ARDS | 72.0 | 94.5 |
| | Non Responder non ARDS | 78.5 | 94.5 |
| | Cumulative | 98.1 | 99.9 |
| 2500 | Responder ARDS | 64.5 | 98,5 |
| | Responder non ARDS | 34.5 | 92 |
| | Non Responder ARDS | 82.0 | 94.5 |
| | Non Responder non ARDS | 77.5 | 93.5 |
| | Cumulative | 99.1 | 99.9 |

Implications



Improved Efficiency and Power

- Hierarchical Bayesian borrowing increases empirical power by sharing information across subgroups (ARDS vs non-ARDS).
- Dynamic borrowing reduces variance of treatment effect estimates without inflating Type I error.
- Response-Adaptive Randomization (RAR) concentrates allocation toward better-performing strategies, increasing precision for promising arms.

Control of False Discoveries

- Bayesian posterior-based decision thresholds (e.g., $P(\text{best}) > 0.95$) maintain low FDR even with multiple adaptive comparisons.
- Hierarchical priors stabilize subgroup estimates, preventing spurious findings in small strata.

Ethical and Operational Gains

- More patients are exposed to effective ventilation strategies during the trial.
- Early futility and superiority rules shorten trial duration and reduce sample size.
- Adaptivity enhances feasibility in heterogeneous ICU populations with variable accrual rates.

Interpretability and Clinical Translation

- Utility-weighted outcomes integrate survival, ventilation days, and recovery → patient-centered inference.
- Posterior probabilities directly support decision-making, not only hypothesis testing.



Future research developments



Computational Complexity: Bayesian adaptive updates (RAR + hierarchical borrowing) require intensive posterior sampling and real-time computation infrastructure.



Regulatory Acceptance: Despite growing support, Bayesian platform and SMART designs still face challenges in regulatory review and harmonization across agencies.



Model Dependence: Inference depends on prior assumptions (e.g., $\tau(g)$, σ) and choice of utility weights; mis-specification may affect borrowing strength and posterior decisions.



Operational Demands: Requires continuous data flow, timely monitoring, and coordination between clinical sites and data scientists.



Generalizability: Borrowing structure validated for ARDS vs non-ARDS; extension to other ICU subgroups needs calibration and external validation.

