



Panel Discussion: Bayes in Regulatory Science

Nicky Best, GSK

Andy Grieve, King's College London

Sofia Villar, MRC BSU, University of Cambridge

Juan Jose Abellan, EMA

***Moderator:* Bradley P. Carlin, Phase V Trials, Inc.**

Bayesian Biostatistics 2025, Leiden, Netherlands

October 23, 2025

- Regulators typically want strong control of the chance of false positives across *all* adaptations (interims, sample-size changes, arm/endpoint selection, enrichment, etc.), usually demonstrated via a thorough simulation plan. They also typically ask how secondary endpoints and correlated outcomes are handled to avoid hidden Type I error inflation.
- **Nicky**, your paper argues that Type I error control may not be the right yardstick for Bayesian designs, arguing that “Bayesian metrics should be used for Bayesian designs.” Can you share any regulatory reactions you’ve had to trials (run by GSK or others) to these ideas? Are we approaching a “tipping point” where the regulatory community will accept metrics other than strict Type I error?
- **Andy**, you’ve been a “moderate” in this debate in that you’ve argued for hybrid frequentist/Bayesian power calculations, and the use of preposterior distributions in design regardless of whether the analysis will be frequentist or Bayesian. Is the future “hybrid”?

- Adaptive design choices (early stopping, population/endpoint selection, sample-size re-estimation) can bias treatment-effect estimates and intervals. Adaptive tools can boost efficiency, but often complicate interpretation. As such, regulators often look for bias-aware estimators/adjustments, and clear plans for how estimates and CIs will be reported after adaptation.
- **Sofia**, as someone who's been a leader in methods for response-adaptive randomization, what statistical innovations or design principles are helping keep these trials both robust and interpretable?
- **Juanjo and Andy**, where should the field draw the line: are there cases where adaptation risks undermining credibility more than it helps efficiency? What safeguards are essential to ensure trial integrity when interim data are used for adaptation?

- The FDA has been visibly supportive of Bayesian methods for over a decade (first in its guidance for medical devices and more recently with its CID program for drugs), albeit mostly for designs permitting only tightly-controlled borrowing from historical control data. By contrast, EMA has historically been much more cautious, though they have issued guidance on approaches for single arm trials.
 - Now we have seen a draft ICH guidance document on adaptive trials that includes just 2 pages on Bayes, and suggests it should be used only when no frequentist alternative exists.
- **Juanjo**, is there still a chance that this transatlantic gap in attitudes toward Bayes will narrow? Also, what practical challenges arise for sponsors running global trials when regulatory expectations differ, and how should Bayesian statisticians anticipate or bridge those gaps?

- Modern Bayesian adaptive trials typically require simulation to verify decision criteria. FDA expects the adaptation algorithm, decision thresholds, number/timing of interims, min/max sample size, and full simulation reports to be prespecified—and they may ask for the actual code used for simulations/analysis.
- **Panelists**, in complex Bayesian/adaptive trials, how much prespecification is enough to satisfy regulators without stifling innovation? Should full simulation code, prior definitions, and effective sample size calculations be routinely submitted, or is there room for pragmatic compromise? What about the use of (now often AI-assisted) specialty software tools that may not be available (or fully reproducible) to regulators?
 - Looking ahead, how do you envision the balance evolving between methodological rigor, regulatory assurance, and real-world feasibility?