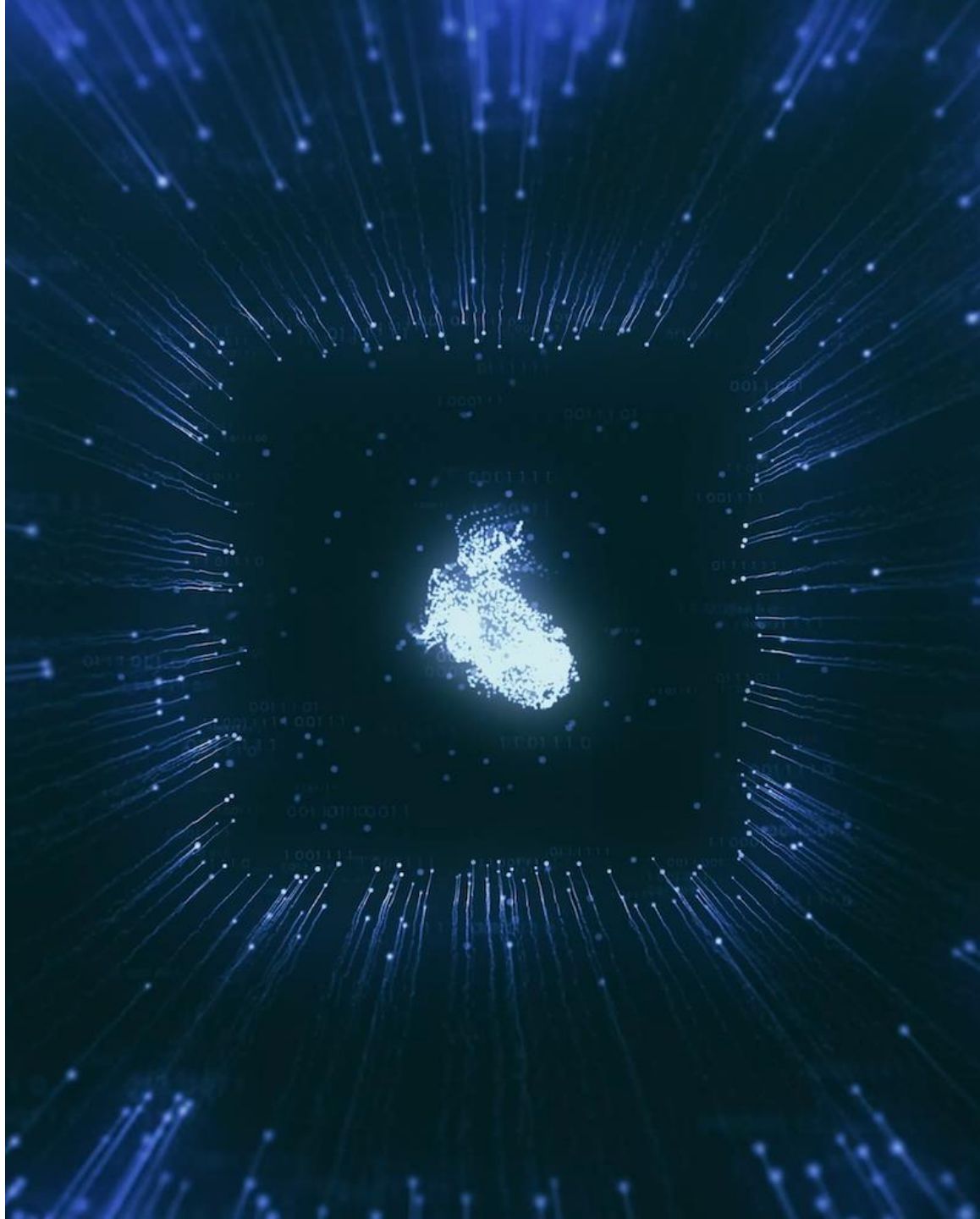
AstraZeneca

How to Incorporate Prior Knowledge

What are Bayesian Statistics and Why Non-Statisticians Should be Interested

Chris Thompson
29JAN2024



What are Bayesian Statistics?

- Bayesian statistics is a field of statistics in which probability is based on
  - Observed data
  - Prior information
- Not a new approach (first paper 1763)
  - Advances in computing have widened its application
- Distinct from Frequentist statistics
- Helpfully intuitive
  - Results describe our current knowledge of the data, model parameters
  - Reflects how most people interpret data

Why Are We Using Bayesian Statistics?

- Prior knowledge
  - Bayes gives a direct and explicit way to incorporate what we already know
- Interpretability
  - Answers the question directly, rather than through a null hypothesis
- Ease of calculation
  - Some statistics are just hard to calculate by alternative methods

CMC Studies Can Benefit From a Bayesian Approach

- Process / Assay Development
- Robustness
- Tech Transfer
- Stability
- Process Capability
- Comparability
- Assay Validation

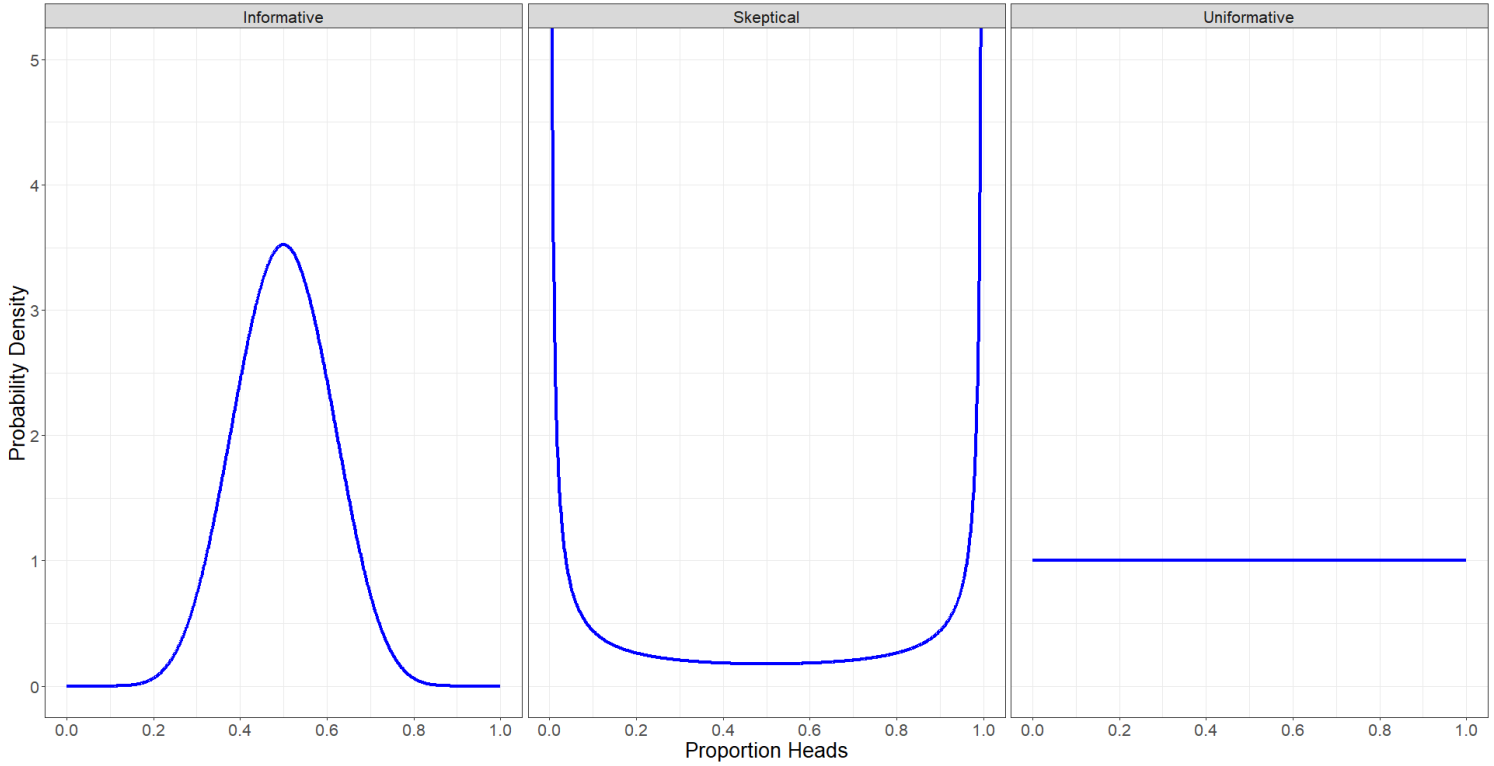
- Many studies where we have prior knowledge
- Including prior knowledge can make study design more efficient
- Newer modalities frequently introduce more complex data
  - Cell therapy
  - mRNA

The Impact of Prior Knowledge

- How much can prior knowledge change our conclusions?
- How do we maintain objectivity?
- How do we justify, defend the use of prior knowledge?
- What happens if my data looks different from my prior knowledge?

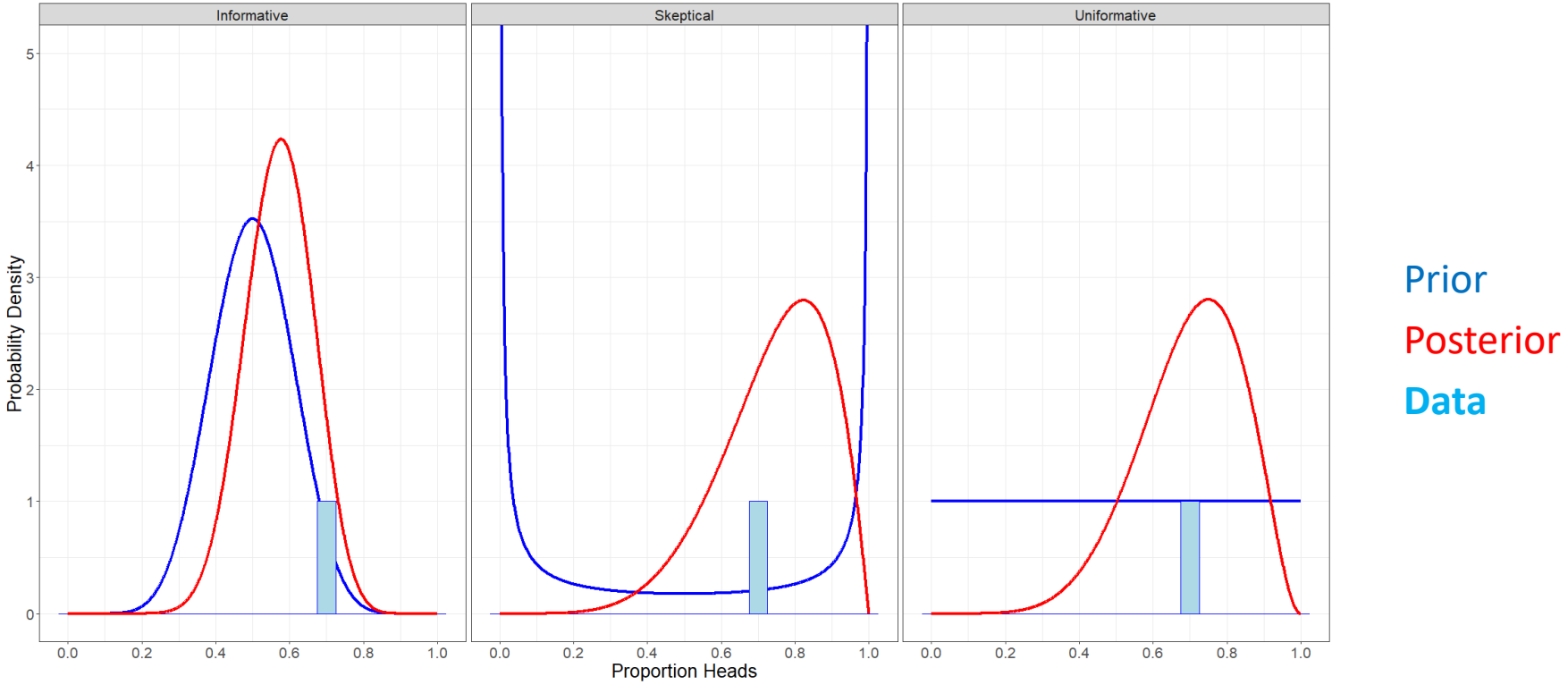
A Simple Example – Flipping a Coin

- Priors on probability of heads, set based on subject matter expertise



A Simple Example – Flipping a Coin

- Result: 7 heads, 3 tails



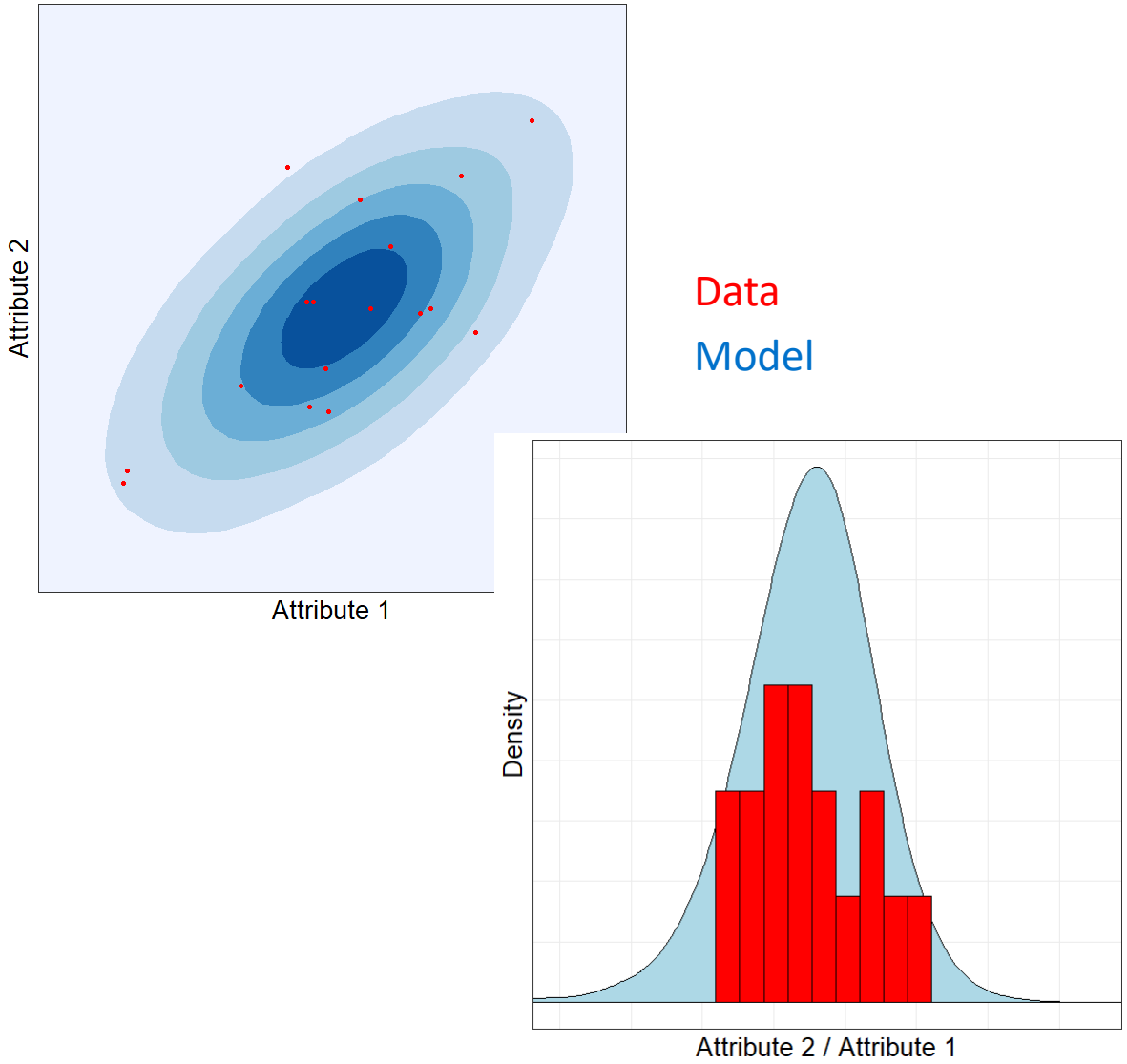
- Data that is consistent with the prior results on a posterior that is more similar to the prior

Case Study: Process Capability

- What is the expected variability of a cell therapy process?
- Compound attributes (ratios of two measured values)
  - Likely correlated with each other

- Most methods to assess variability of this type are not trivial and are approximate
- With a Bayesian method we fit a model and get the full distribution for both attributes
  - And we can add prior knowledge if available

Process Capability Output



- Using a Bayesian model the multivariate distribution could be sampled
- We get the full joint distribution of all of the model parameters
- We can calculate our ratio distribution
- We can answer questions about the data
  - What is the probability of being above x?
- Method works generally
  - Not dependent on levels of attributes
  - Or degree of correlation

Case Study: Process Development

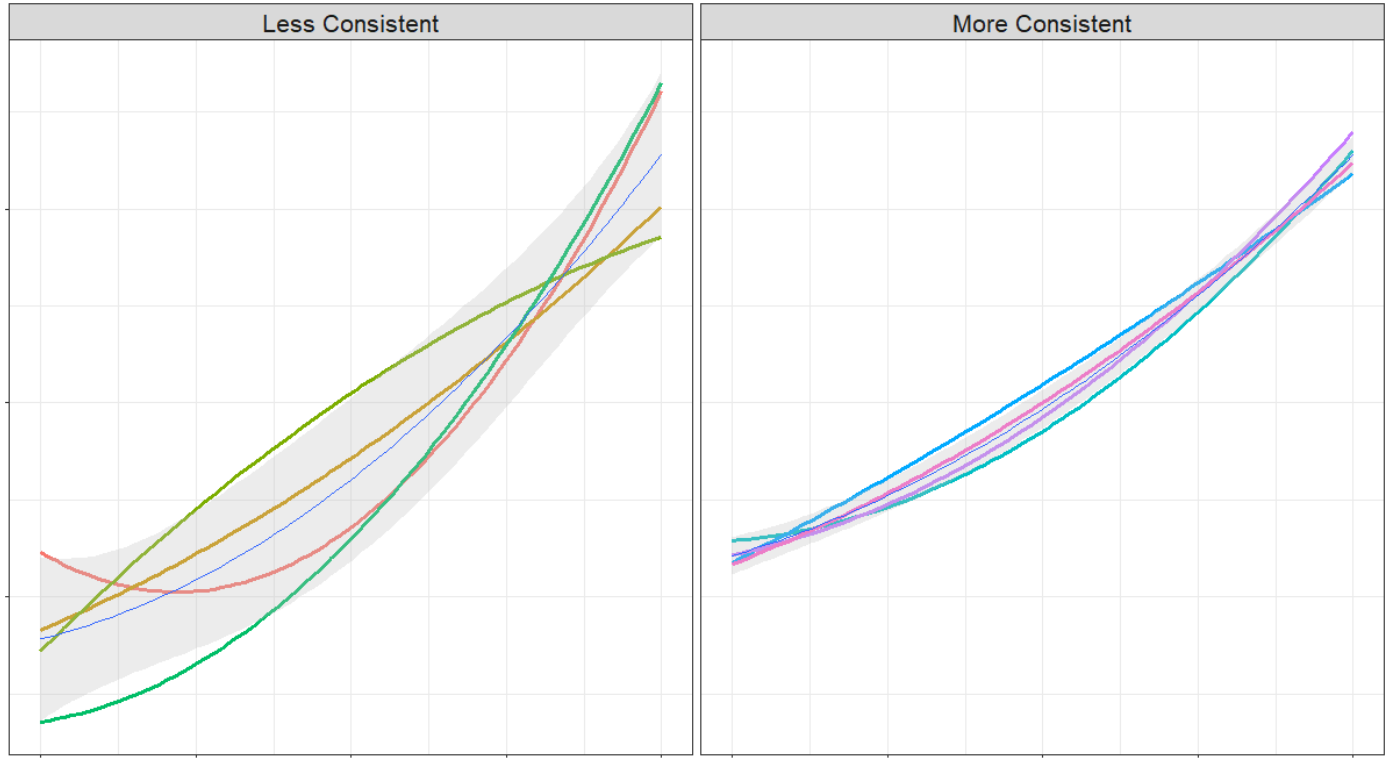
- Results of process characterization studies are similar for highly homologous biologics expressed in clones generated from the same host strain under platform conditions

- Leveraging the data across a platform may allow more focused characterization experiments
  - New studies designed to confirm impact observed previously directed

- Data from large response surface studies were used to compare different experimental designs
  - A hierarchical model

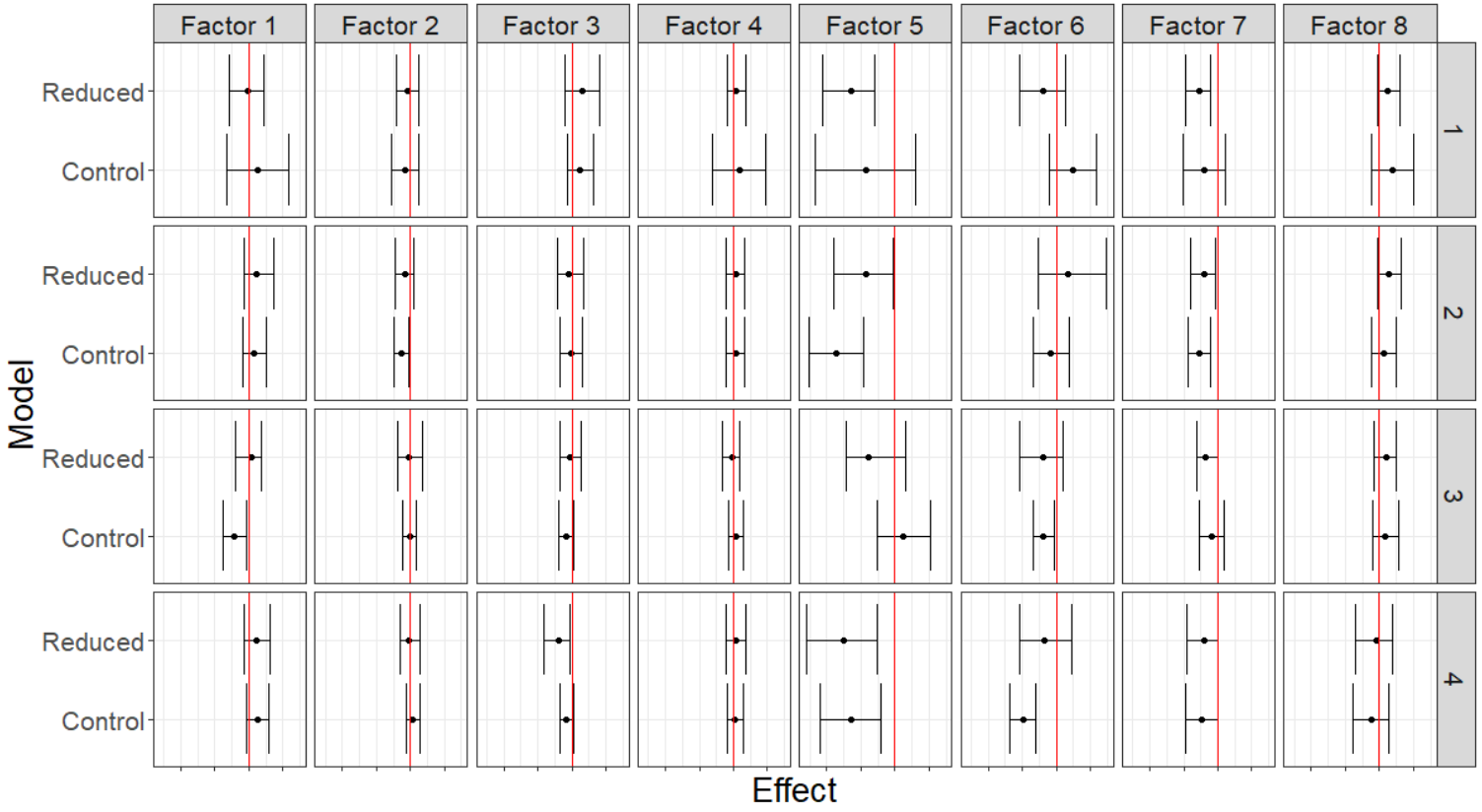
Hierarchical Models

- Hierarchical models allow the molecules to be different
  - More consistency increases expectation that the future will also be consistent



Process Characterization Data

- Incorporating data from across molecules, a reduced DoE gives comparable results to a full size study using data from a single molecule



Additional factors not shown for clarity

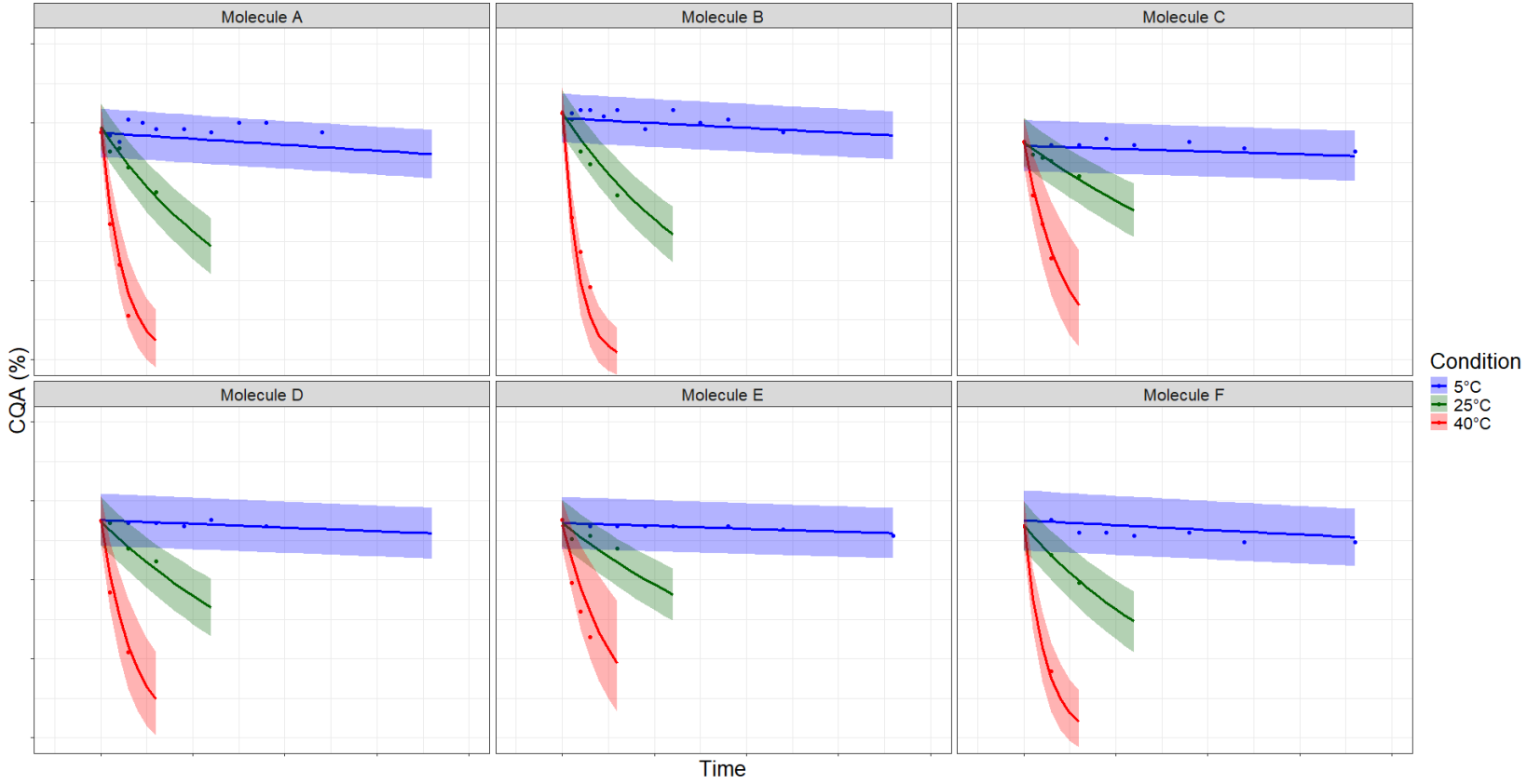
Case Study: Stability

- Determining real time product stability for biologics is a common problem
- In many cases stability between similar molecules is comparable
- Data available from 6 IgG1 molecules at 3 temperatures
- A non-linear hierarchical Bayesian regression model was used
  - Kinetic model

- Using degradation rates between molecules and temperatures, expectation for missing combinations can be calculated

Stability Data

- Predictions show good agreement across molecules over all conditions
- Despite varying rates at accelerated conditions



Is Our Prior Knowledge Relevant?

- In general, inconsistency causes uncertainty

- Good practice to justify priors scientifically
  - These molecules are highly homologous
  - The reaction rate is positive and increases with temperature

- Even better to justify with data
  - The accelerated stability of this molecule is consistent with the others

Summary

- Bayesian statistics inherently lend themselves to use of prior knowledge
  - Not all they are used for
- There are many sources and methods to incorporate prior knowledge
  - Methods are primarily statistical
  - Sources and justification are primarily scientific
- Using prior knowledge quantitatively allows better decision making

- Consistent priors and data decrease uncertainty
- Inconsistent priors and data can, and should, increase uncertainty

Conclusion

- Use of Bayesian methods for non-clinical applications is increasing rapidly
  - Already well established for clinical studies

- Clear case for Bayesian methods as an additional set of tools in CMC

- Not as much established practice as for Frequentist statistics (yet)
  - Several groups working to address this
  - Active area for Regulatory engagement

AstraZeneca PLC, 1 Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge, CB2 0AA, UK
+44(0)203 749 5000
www.astrazeneca.com