

# Predicting and Ranking Formulations in a Sparse Categorical DOE for Potency

Francisca Galindo Garre

Senior Principal Statistician

Manufacturing and Applied Statistics

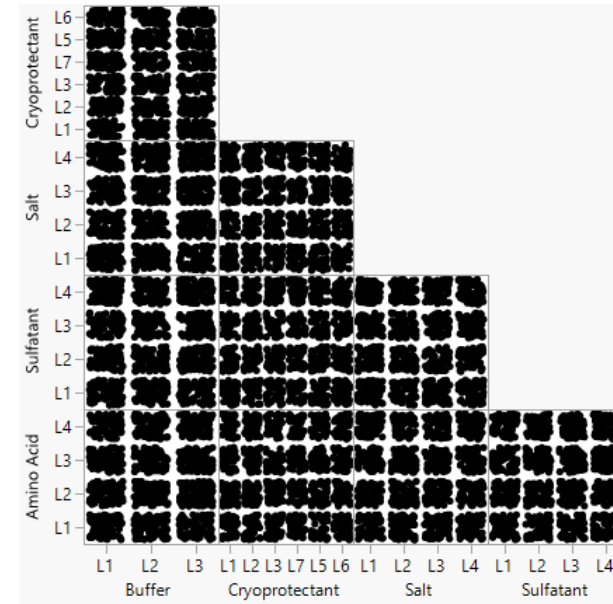
# How can we find promising formulations in a sparsely sampled categorical space?

The objective of this study is to identify stabilizing excipients for Lenti virus vector during relevant stress conditions.

A formulation study was design to study potency:

|                |                                     |
|----------------|-------------------------------------|
| Buffer         | hard to change factor with 3 levels |
| Cryoprotectant | 6 levels                            |
| Salt           | 4 levels                            |
| Surfactant     | 4 levels                            |
| Amino acids    | 4 levels                            |

Full factorial Design



> 1000 possible combinations

# A 65 runs D-optimal split-plot DOE screened five factor

65

RUNS

D-optimal design for a model with main effects as necessary and second order interactions as if-possible results in a 65 runs design.

19

19 runs had observed potency below the level of quantification of the assay.

?

DECISION GAP

Loss of design information from missing runs to calculate probability of success and select the set of formulation candidates



# Enumerate, predict, and rank the discrete design space

## CENSORED MIXED-EFFECTS MODEL

A mixed effects censored (Tobit) regression was fitted, treating no detects at the assay limit of quantification as left censored.

$$y_i^* \sim \mathcal{N}(\mathbf{x}_i^\top \boldsymbol{\beta} + b_{WP(i)}, \sigma^2)$$

$$b_{WP} \sim \mathcal{N}(0, \sigma_{WP}^2)$$

$$y_i = \begin{cases} L_i, & y_i^* \leq L_i \\ y_i^*, & y_i^* > L_i \end{cases}$$

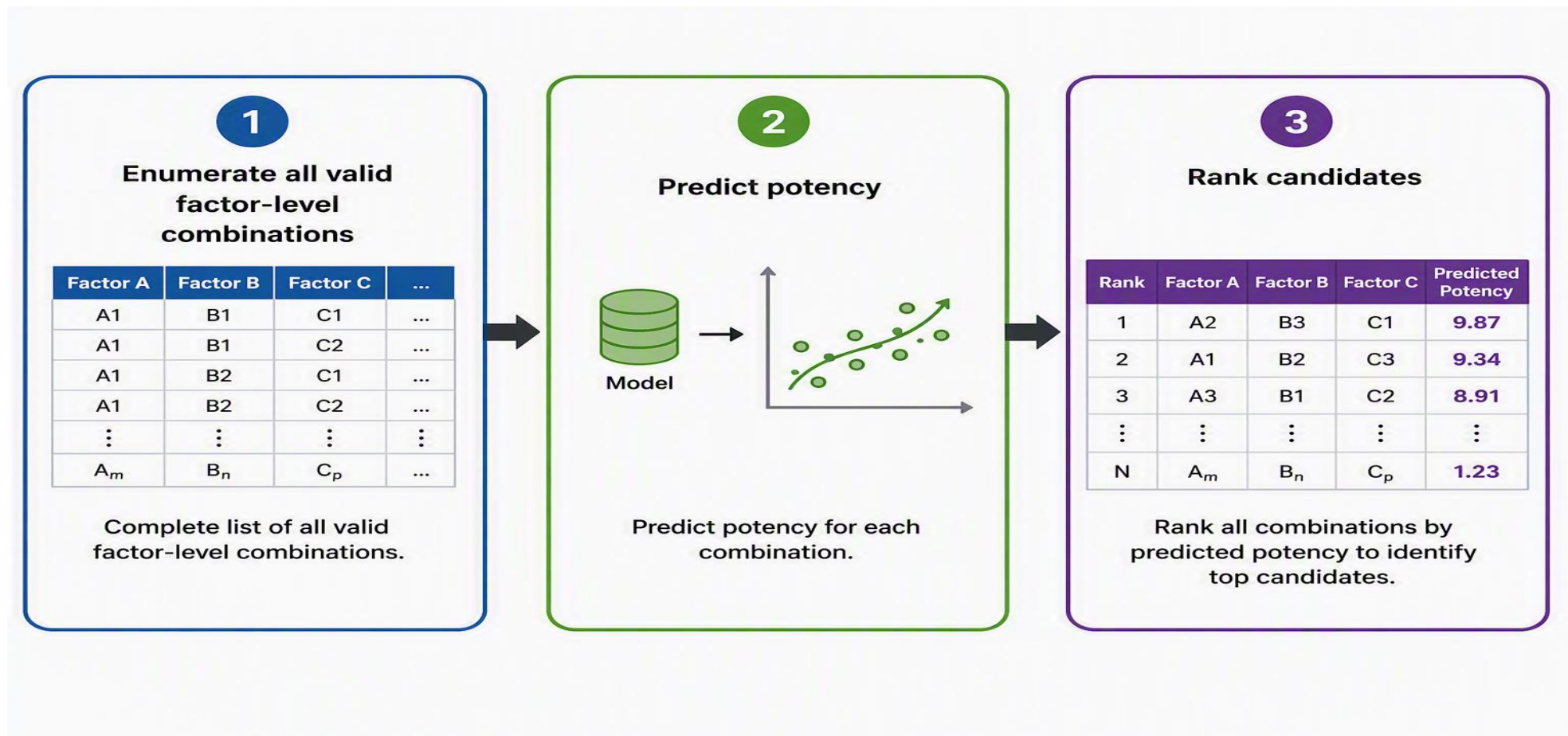
The model was fitted using the Bayesian approach with the R package brms.

**Fixed effects:** Cryoprotectant, Salt, Buffer, Surfactant, Amino Acids, and selected two-way interactions.

**Random effect:** WP (Whole Plots) intercept.

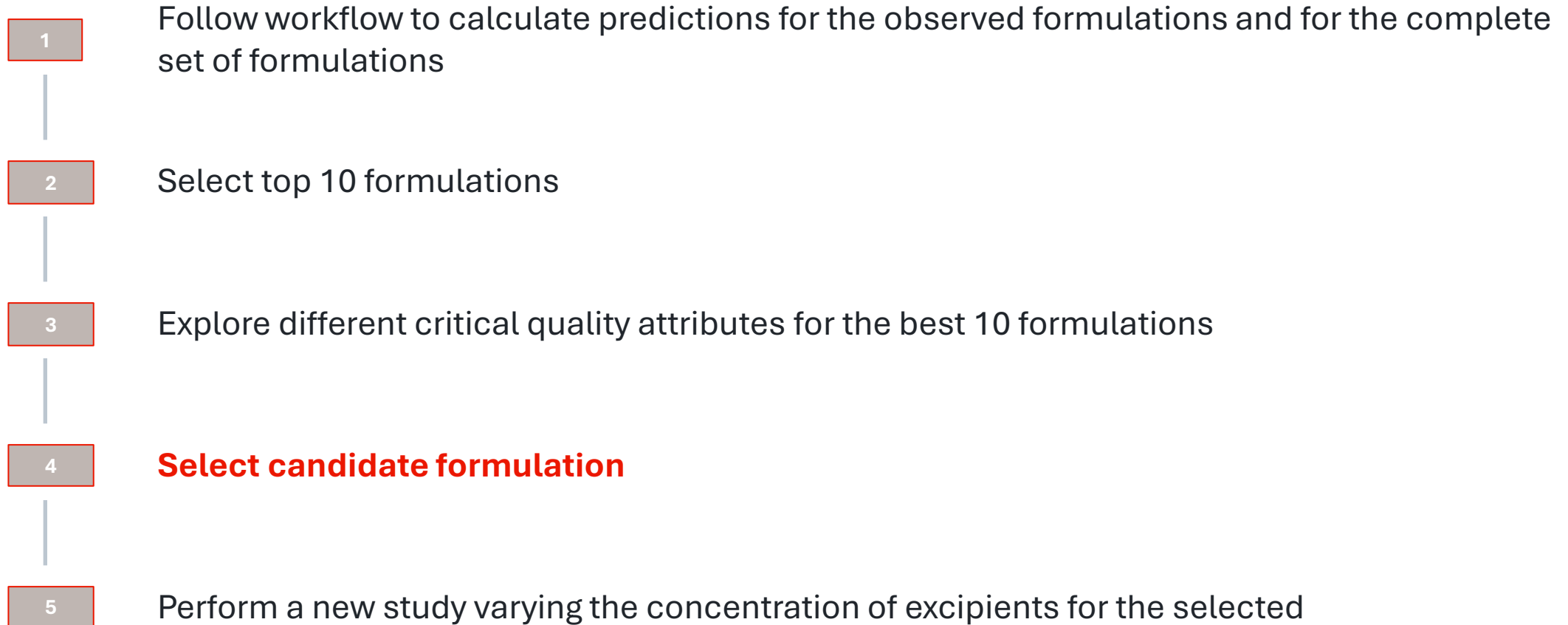
**Censoring:** Measurements below the detection limit  $L_i$  are treated as left-censored.

# Enumerate, predict, and rank the discrete design space



Ranking based on probability of having potency above LLOD, predicted potency or prediction uncertainty.

# Sparse categorical DOE can guide formulation decisions



# Thank you

Francisca Galindo Garre,  
Senior Principal statistician,  
Manufacturing and Applied  
Statistics (MAS).  
[fgalind1@its.jnj.com](mailto:fgalind1@its.jnj.com)

**Johnson&Johnson**  
Innovative Medicine