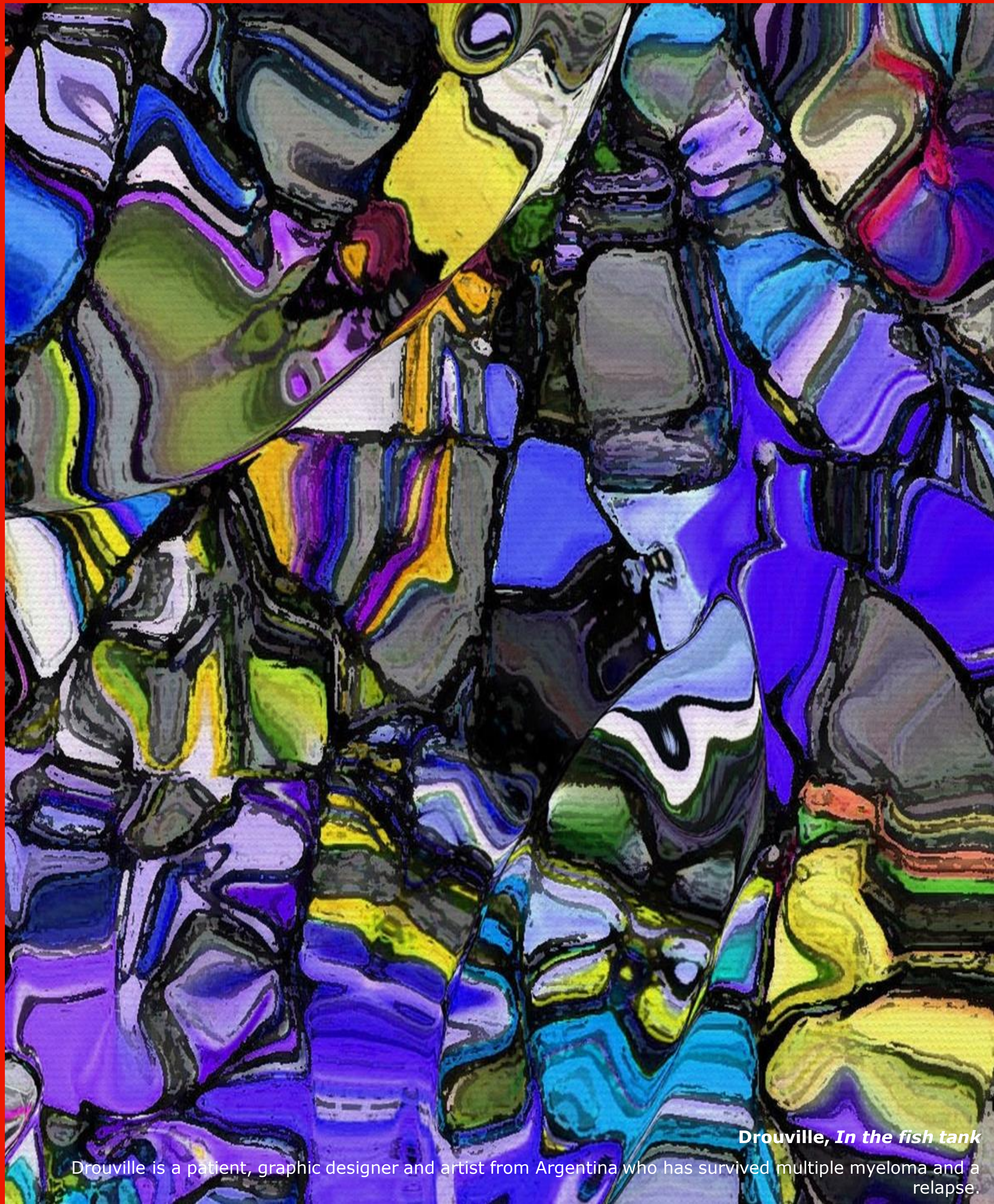
Leveraging Bayesian Techniques in DOE Model Prediction and Simulation to Enhance Decision-Making in the Context of Large-Molecule Process Characterization in the Pharmaceutical Setting

Yang Cao & Jyh-Ming Shoung

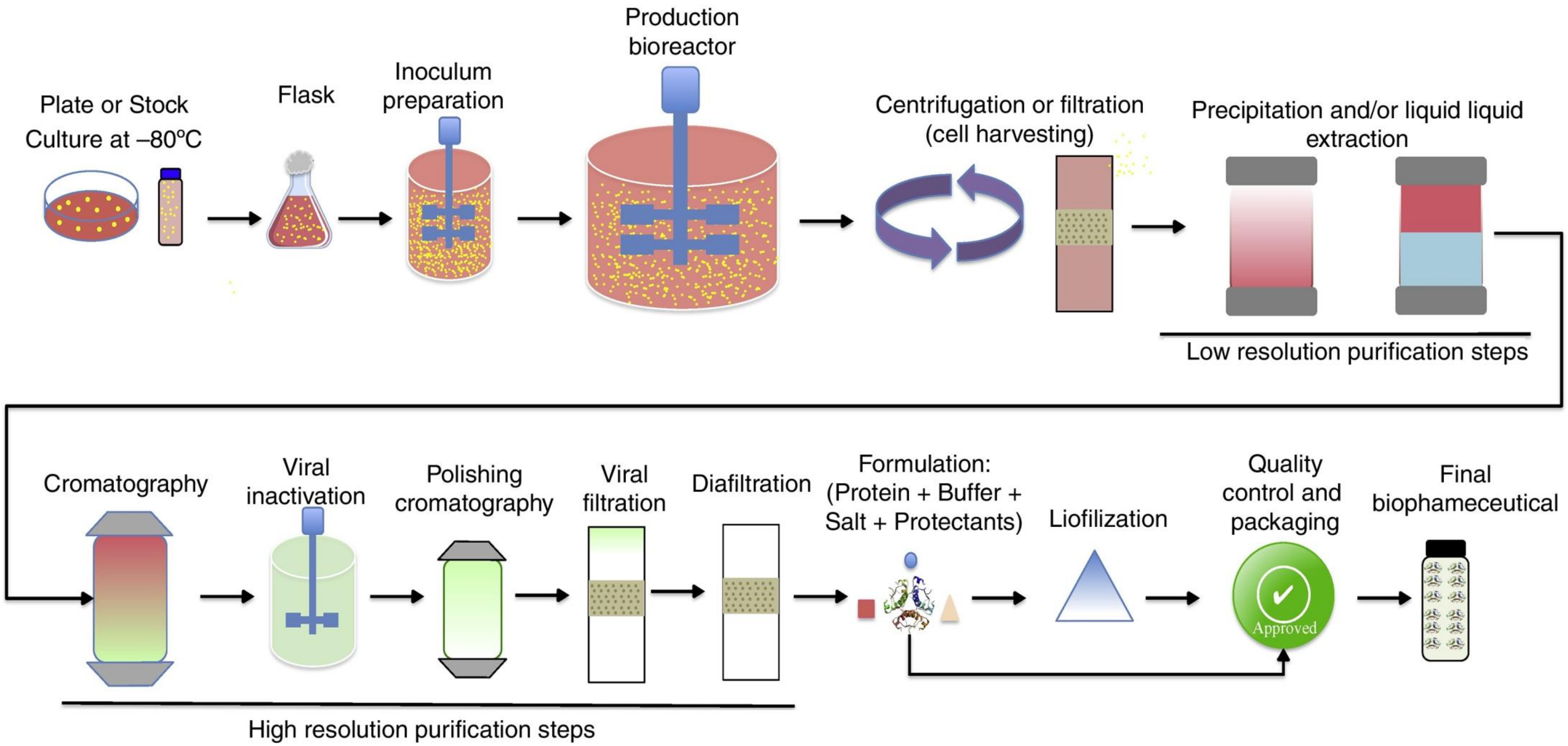
Non-Clinical Statistics Conference
September 25th, 2024
Wiesbaden, Germany

Johnson&Johnson
Innovative Medicine



Drouville, *In the fish tank*
Drouville is a patient, graphic designer and artist from Argentina who has survived multiple myeloma and a relapse.

process characterization



Criticality—which parameters critically impact the quality of product?

Proven acceptable range (PAR)--what ranges of the parameters are acceptable?

Image source: https://www.sciencedirect.com/science/article/pii/S1517838216310413

Proposed statistical workflow for criticality and PAR assessment

DoE studies

Statistical model

Y=a+b*X

Criticality

Effect-to-noise ratio

CPP/Non-CPP

PAR

Monte Carlo simulation

Failure rate of simulated attribute meets criteria

Yes

PAR confirmed

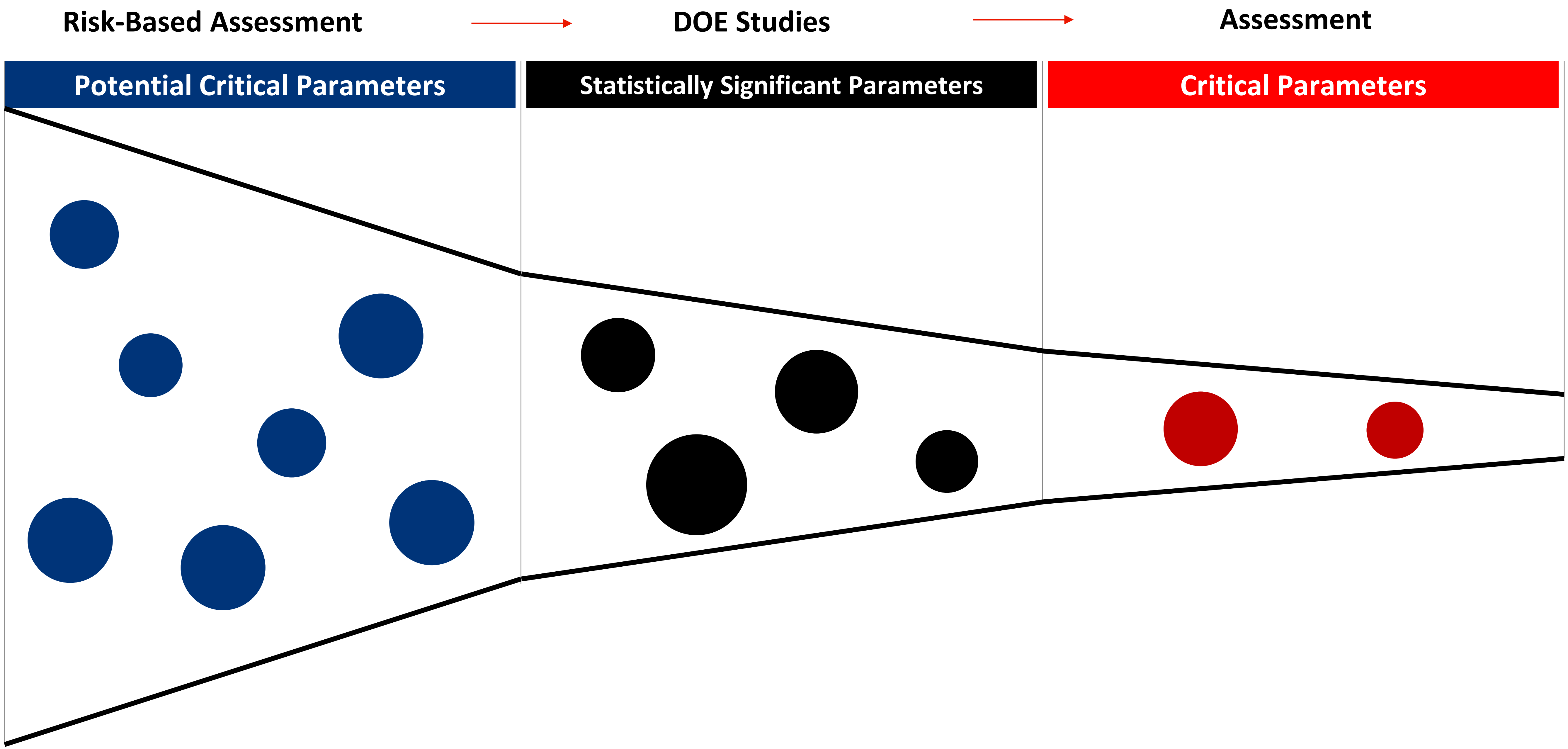
No

Adjust PAR until criteria is met

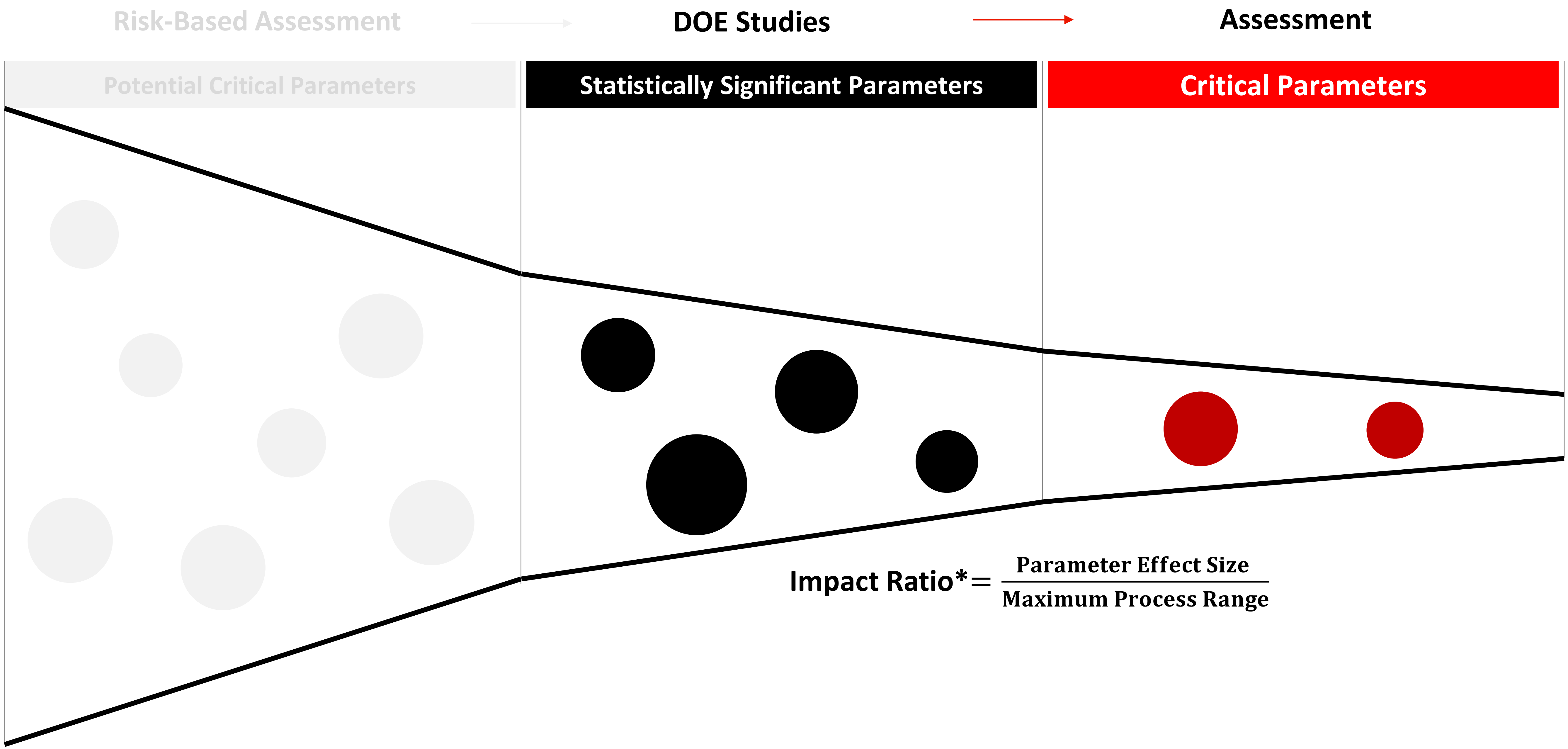
CPP: Critical Process Parameter

PAR: Proven Acceptable Range

Which parameters are critical?



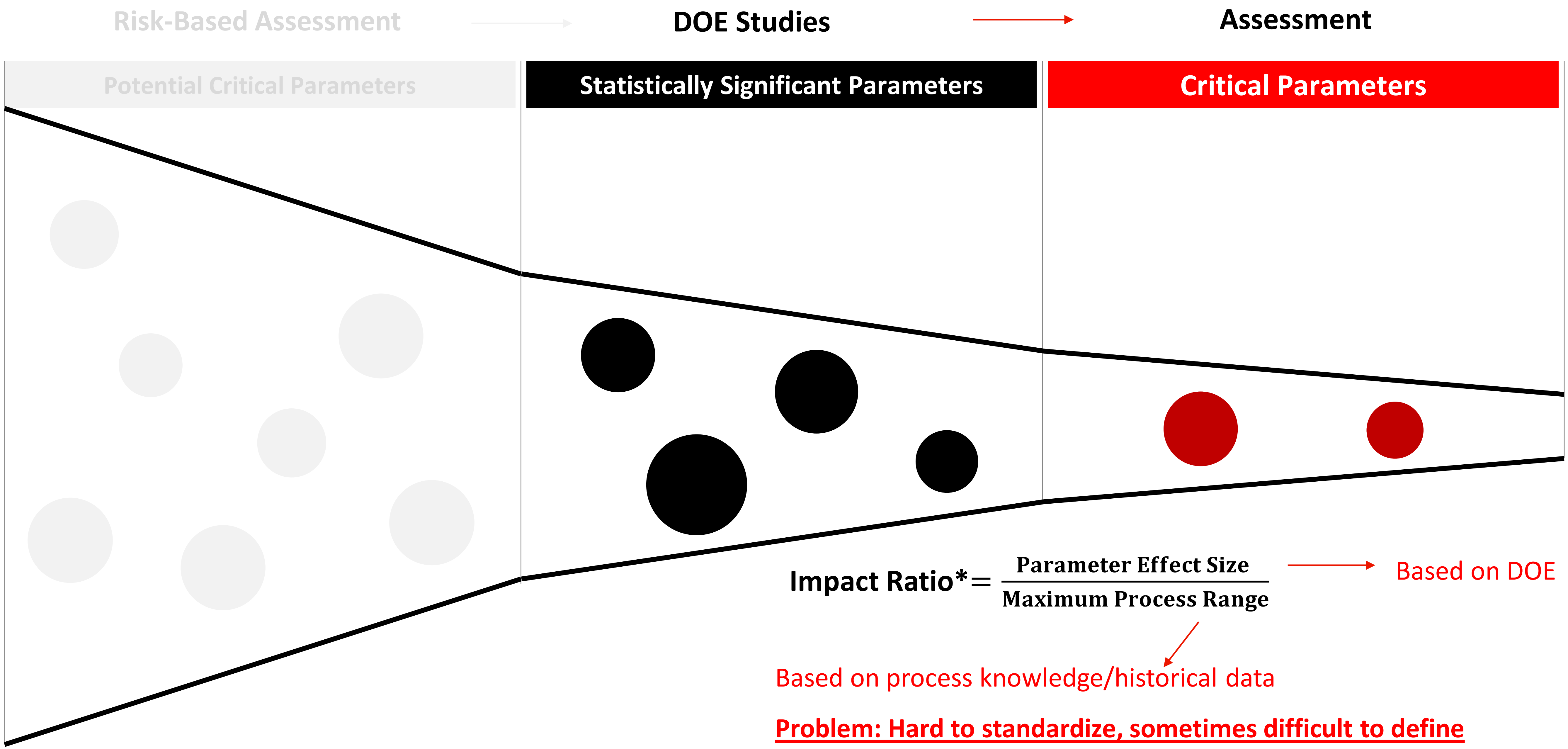
Which parameters are critical?



$$\text{Impact Ratio}^* = \frac{\textbf{Parameter Effect Size}}{\textbf{Maximum Process Range}}$$

*Hakemeyer C, McKnight N, St John R, Meier S, Trexler-Schmidt M, Kelley X2, Zettl F, Puskeiler R, Kleinjans X1, Lim F, Wurth C. Process characterization and Design Space definition. Biologicals. 2016 Sep;44(5):306-18. doi: 10.1016/j.X2iologicals.2016.06.004. Epub 2016 Jul 25. PMID: 27464992.

Which parameters are critical?



$$\text{Impact Ratio*} = \frac{\text{Parameter Effect Size}}{\text{Maximum Process Range}}$$

Parameter Effect Size → Based on DOE

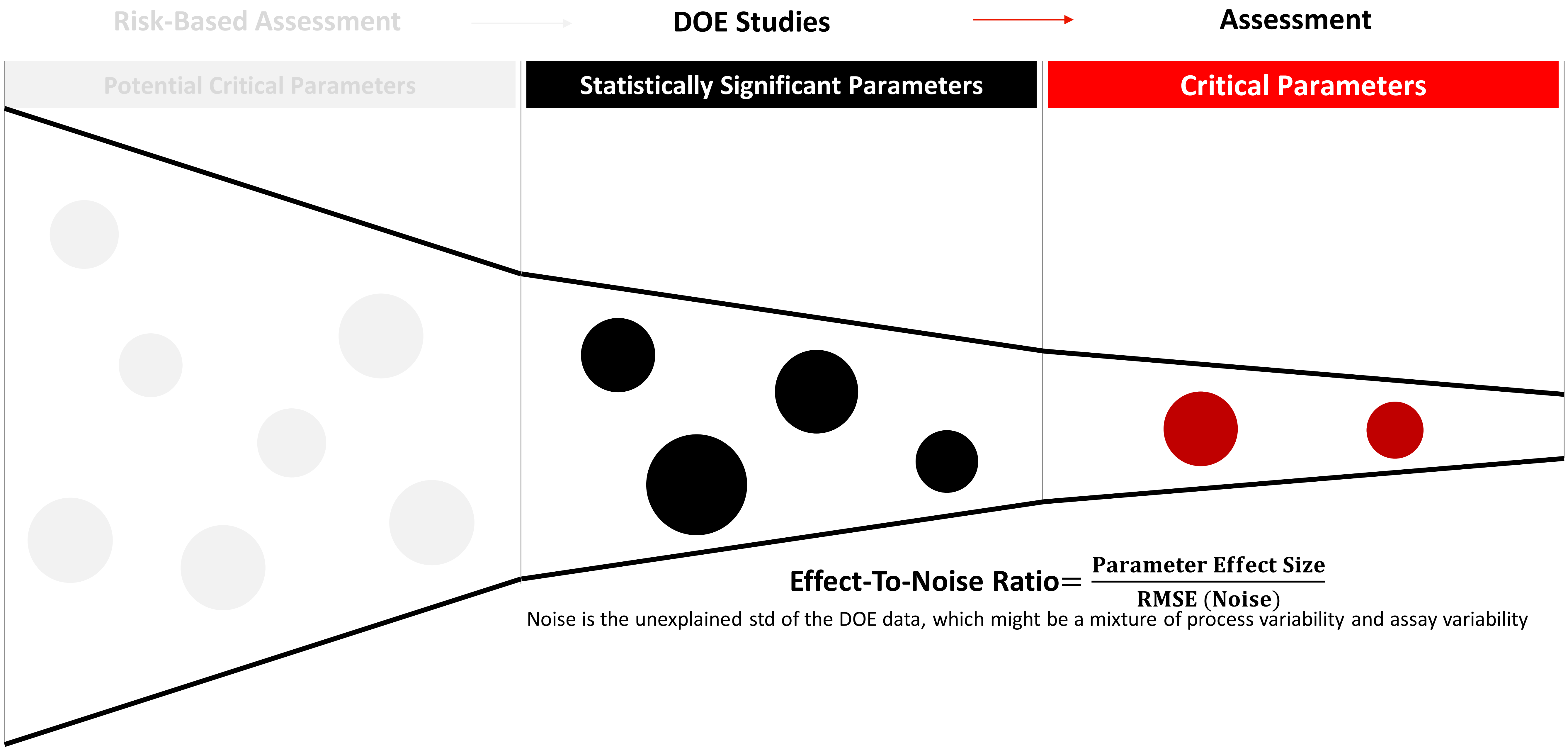
Maximum Process Range → Based on process knowledge/historical data

Problem: Hard to standardize, sometimes difficult to define

*Hakemeyer C, McKnight N, St John R, Meier S, Trexler-Schmidt M, Kelley X2, Zettl F, Puskeiler R, Kleinjans X1, Lim F, Wurth C. Process characterization and Design Space definition. Biologicals. 2016 Sep;44(5):306-18. doi: 10.1016/j.X2iologicals.2016.06.004. Epub 2016 Jul 25. PMID: 27464992.

We are proposing an effect-to-noise ratio, calculated based on DOE model

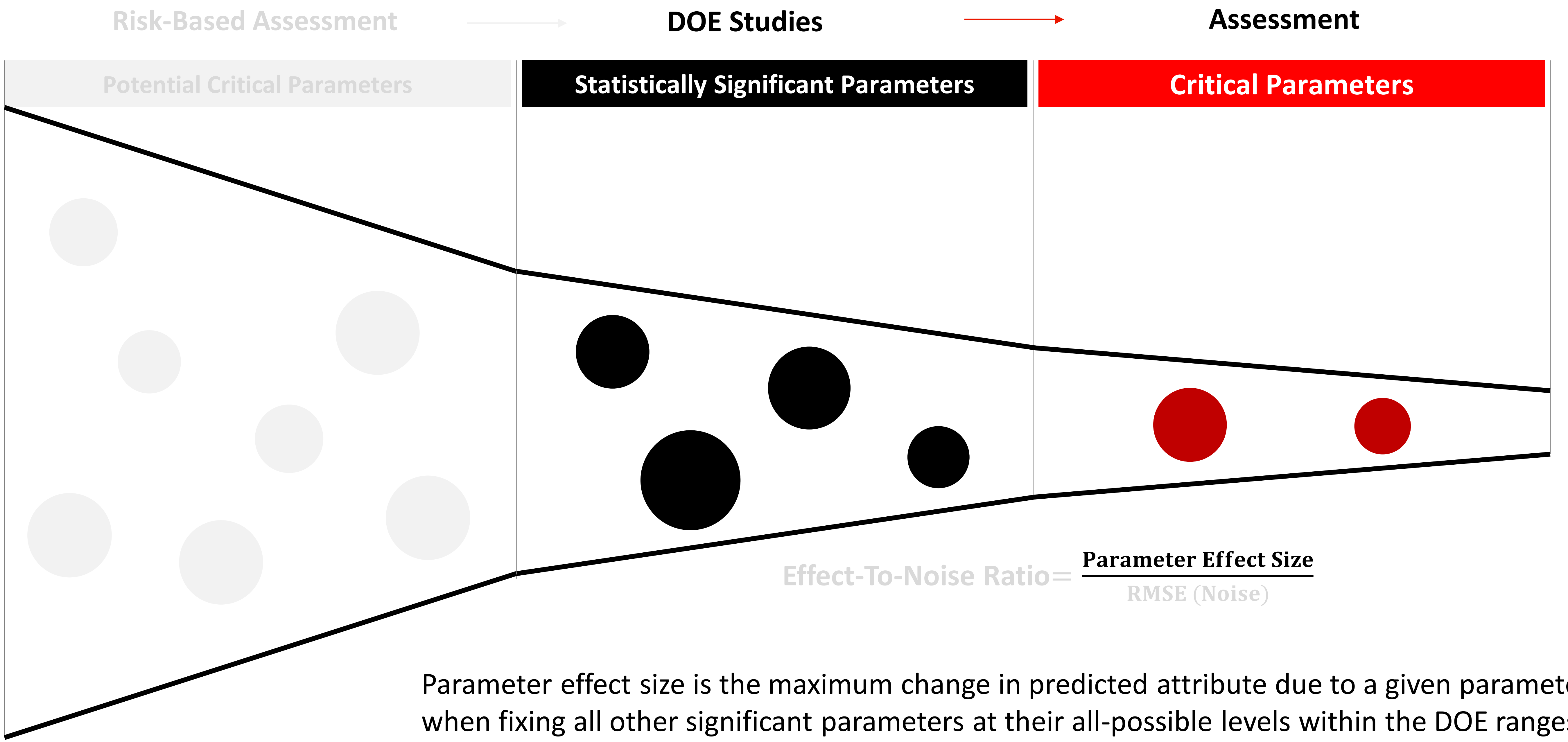
It allows for consistent and fair comparison, even when process knowledge is limited



$$\text{Effect-To-Noise Ratio} = \frac{\textbf{Parameter Effect Size}}{\textbf{RMSE (Noise)}}$$

Noise is the unexplained std of the DOE data, which might be a mixture of process variability and assay variability

We are proposing an effect-to-noise ratio, calculated based on DOE model

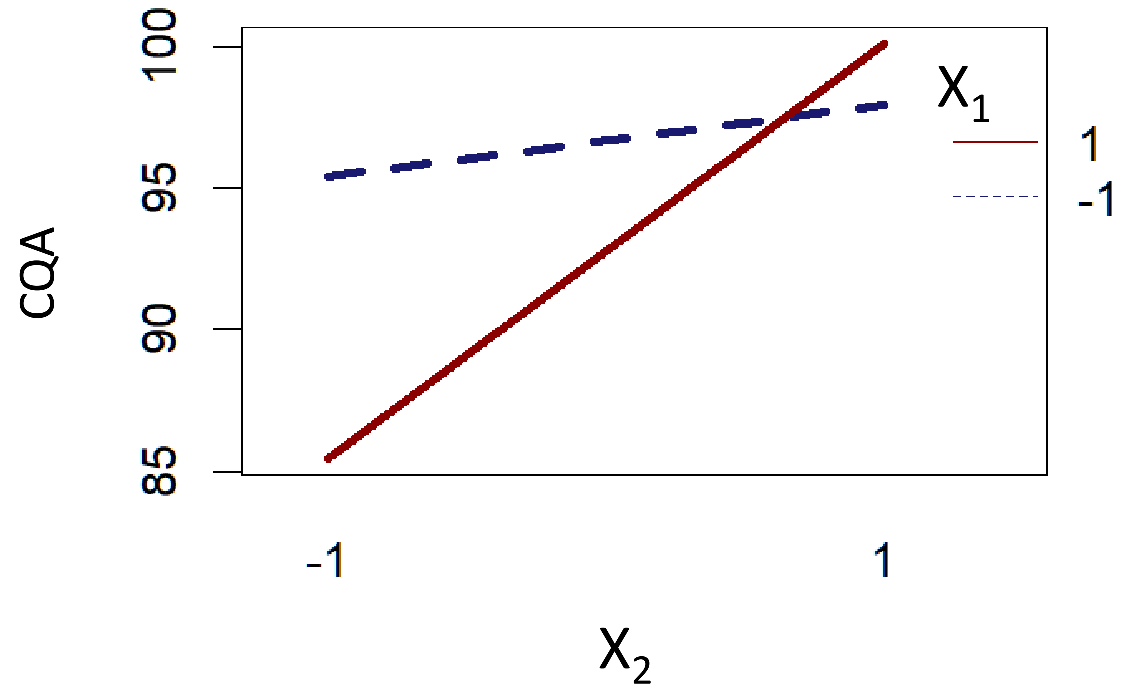


$$\text{Effect-To-Noise Ratio} = \frac{\textbf{Parameter Effect Size}}{\text{RMSE (Noise)}}$$

Parameter effect size is the maximum change in predicted attribute due to a given parameter when fixing all other significant parameters at their all-possible levels within the DOE ranges

Parameter effect size (X_2): main effects +interactions

DOE Model: CQA=94.7 -2* X_1 +4.3* X_2 -3* X_1 X_2

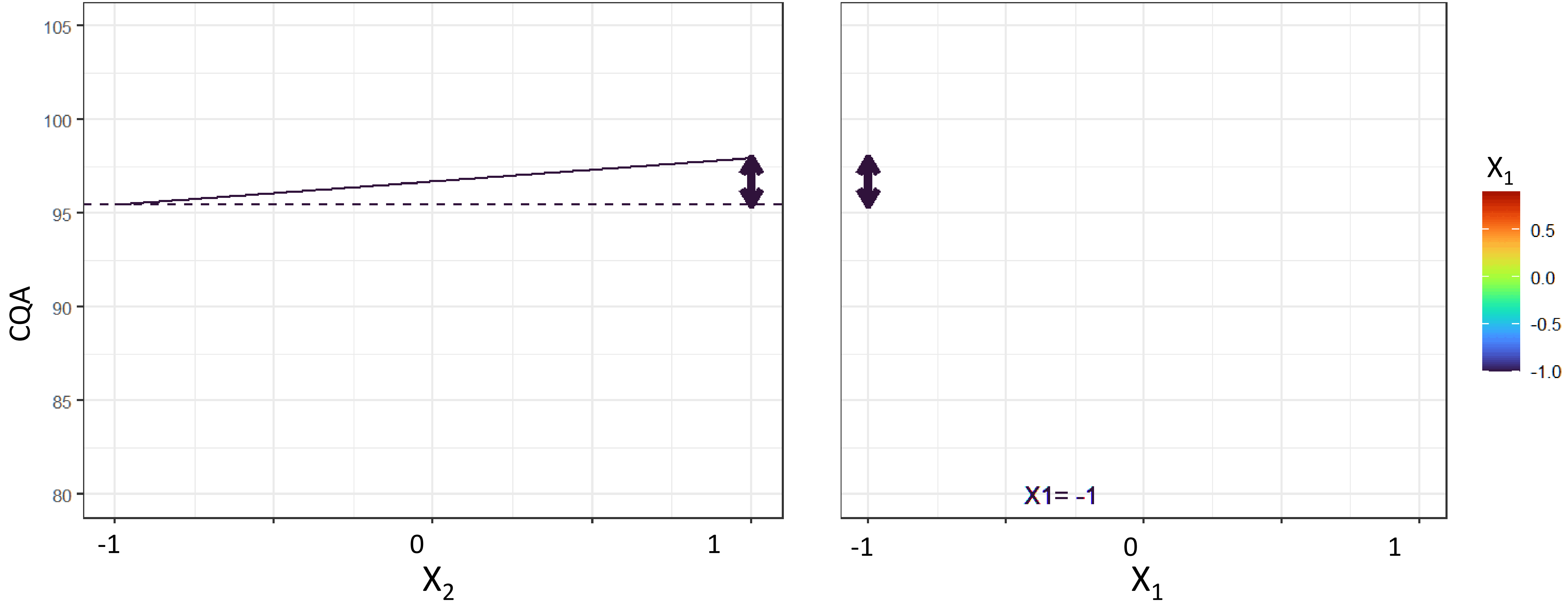


What is the parameter effect size for X_2 ?

Parameter effect size: main effects +interactions

Grid-Search Illustration*--can be universally applied, accounting for applicable interactions and curvatures

DOE Model: $CQA=94.7 -2^* X_1 +4.3^* X_2 -3^* X_1 X_2$



Parameter effect size of X_2 = the max length of the vertical arrows=14.65

It quantifies the magnitude change in CQA due to X_2 when fixing other significant parameter at a level that results in the greatest impact

J&J Innovative Medicine

*F. Li *et al*., "Removing Subjectivity from the Assessment of Critical Process Parameters and Their Impact," *Pharmaceutical Technology* 42 (1) 2018.

Effect-to-noise ratio

$$CQA = 94.7 - 2 * X_1 + 4.3 * X_2 - 3 * X_1 X_2$$

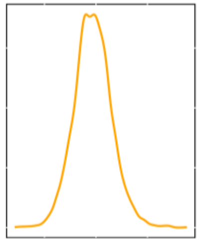
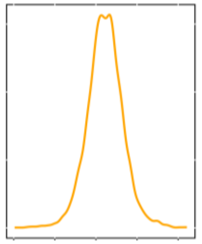
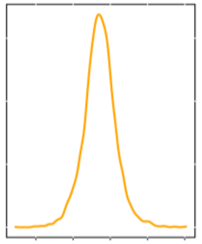
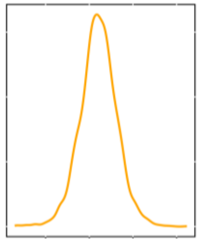
$$\textbf{Effect-to-Noise Ratio } (X_2) = \frac{\textbf{Parameter Effect Size}(X_2)}{\textbf{Noise (RMSE)}} = \frac{\textbf{14.65}}{\textbf{2.55}} = \textbf{5.75}$$

- RMSE is directly derived from the DOE model

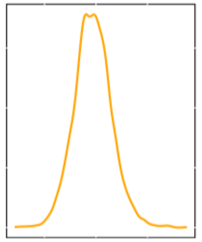
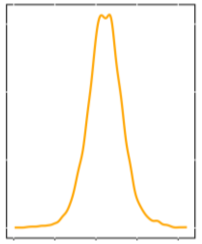
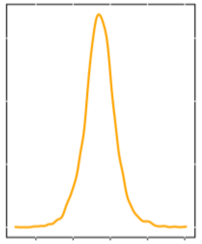
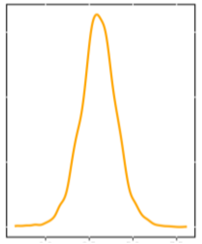
The max magnitude change in CQA due to X_2 is 5.75 times the noise

Now, can we do better and account for model uncertainty?

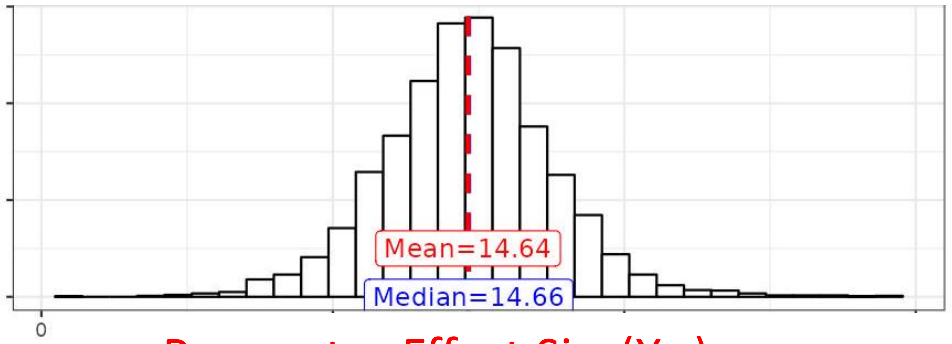
Yes ! Instead of estimates, we can get distributions, thanks to Bayesian

CQA=  +  * X_1 +  * X_2 +  * $X_1 X_2$

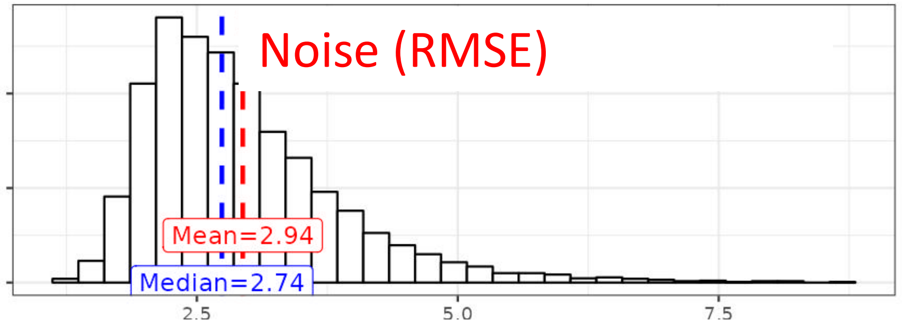
Now, we have a distribution of effect-to-noise ratio from Bayesian

CQA=  +  * X_1 +  * X_2 +  * X_1 X_2

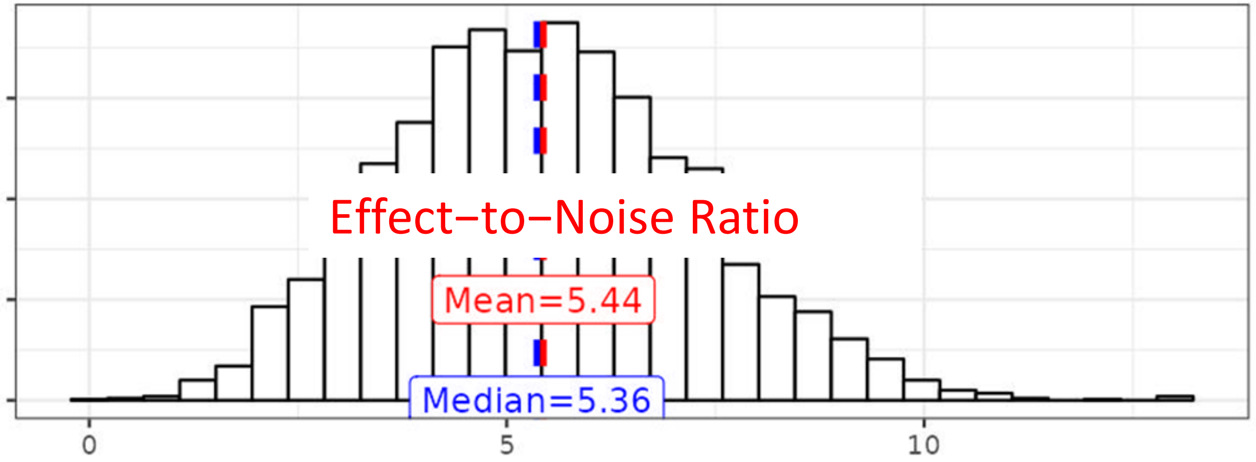
Effect-to-Noise Ratio (X_2)=



Parameter Effect Size(X_2)

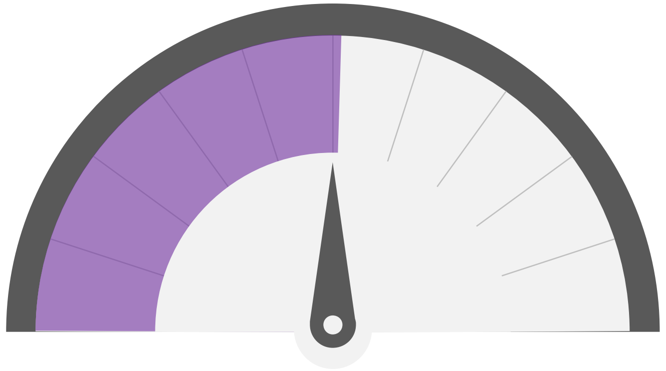


=



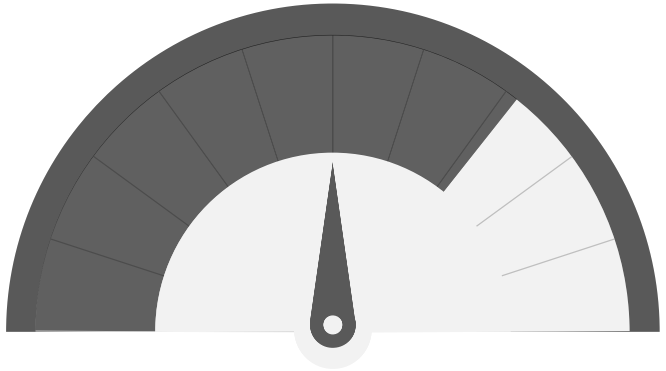
What is the certainty that the effect is real rather than noise?

Low



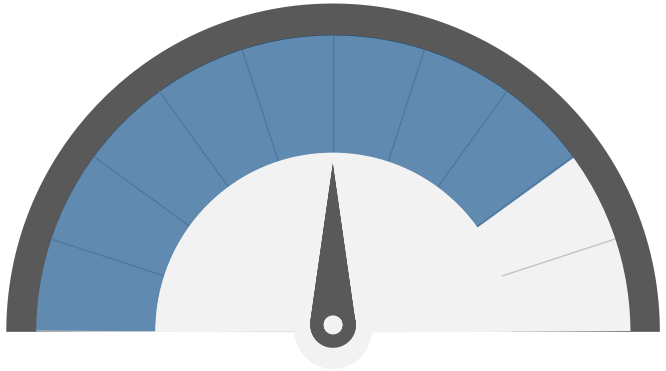
$\mathrm{Prob}\ (effect > noise) < 50\%$

Medium



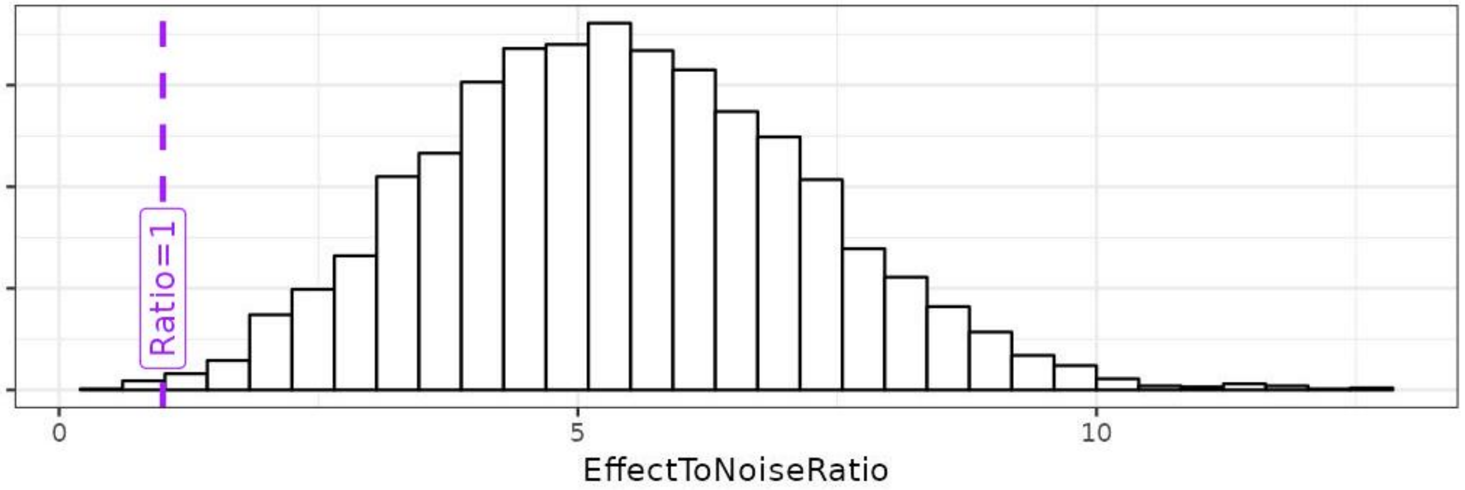
$50\% \leq \mathrm{Prob}(effect > noise) < 80\%$

High



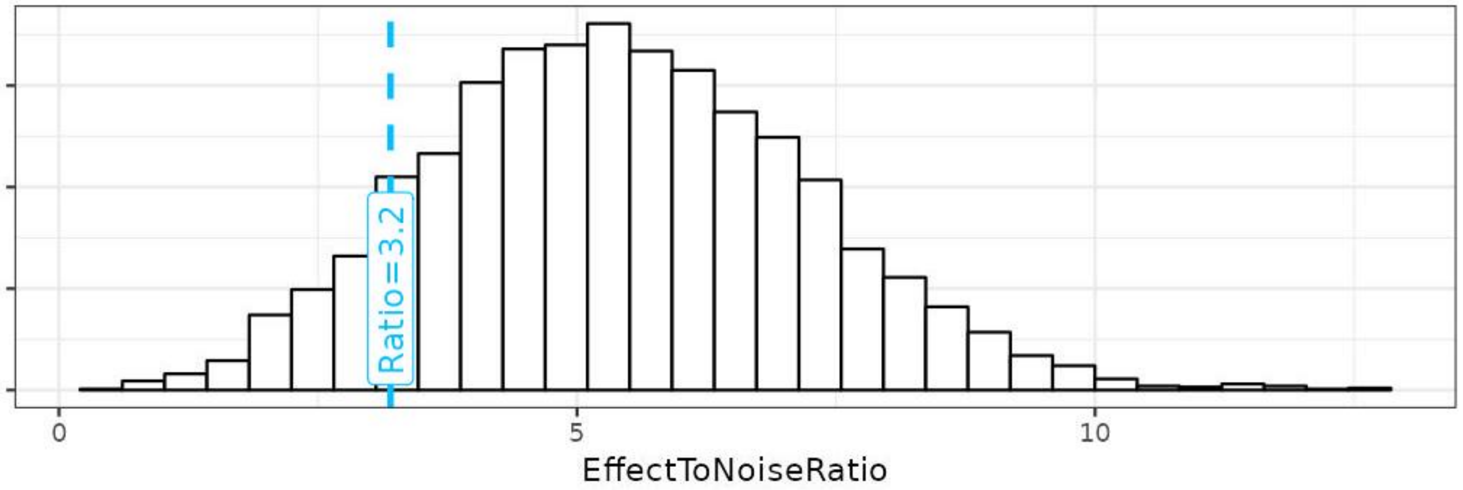
$80\% \leq \mathrm{Prob}\ (effect > noise)$

We can leverage the distribution of effect-to-noise ratio to assess certainty



Estimate based on Frequentist=5.7								
Estimates based on Bayesian								
Lower Bound	**99.8%**	**90%**	**80%**	**75%**	**50%**	**25%**	**20%**	**10%**
Ratio	1.0	3.2	3.9	4.2	5.3	6.6	6.9	7.8
99.8% certainty that the effect is greater than noise								

We can leverage the distribution of effect-to-noise ratio to assess criticality



Estimate based on Frequentist=5.7								
Estimates based on Bayesian								
Lower Bound	**99.8%**	**90%**	**80%**	**75%**	**50%**	**25%**	**20%**	**10%**
Ratio	1.0	3.2	3.9	4.2	5.3	6.6	6.9	7.8
90% chance that the effect is at least 3.2 times the noise								

X_2 might be considered a CPP since there is a high chance that its impact is practically significant relative to the noise

Proposed statistical workflow for criticality and PAR assessment

DoE studies

↓

Statistical model

Y=a+b*X

↓

PAR

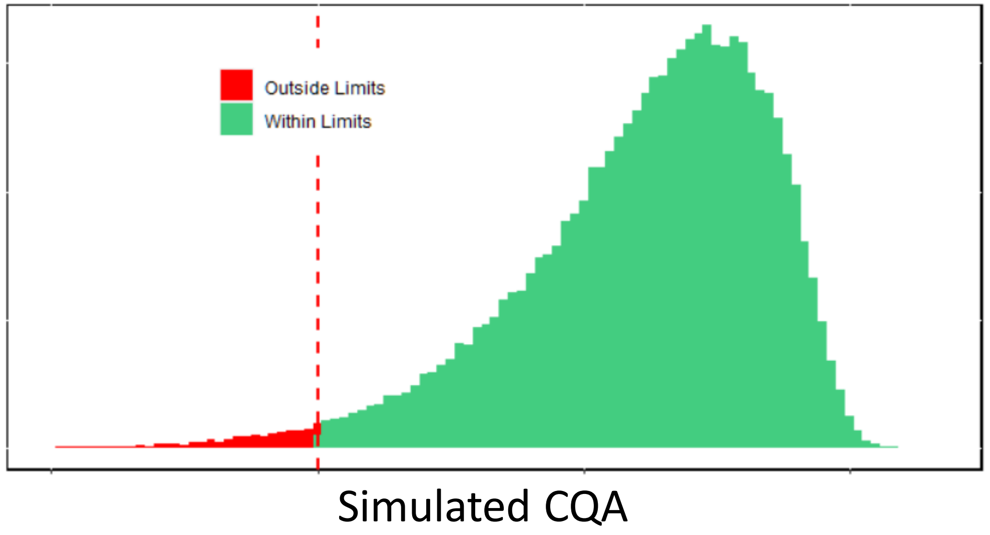
Monte Carlo simulation

↓

Failure rate of simulated attribute meets criteria

Yes → **PAR confirmed**

No → **Adjust PAR until criteria is met**



CPP: Critical Process Parameter

PAR: Proven Acceptable Range

Proven acceptable range (PAR) definition

PAR defined in ICH Q8 (R2) : “a characterized range of a process parameter for which operation within this range, while keeping other parameters constant, will result in producing a material meeting relevant quality criteria”.

Proven acceptable range (PAR) definition

at all possible extreme conditions

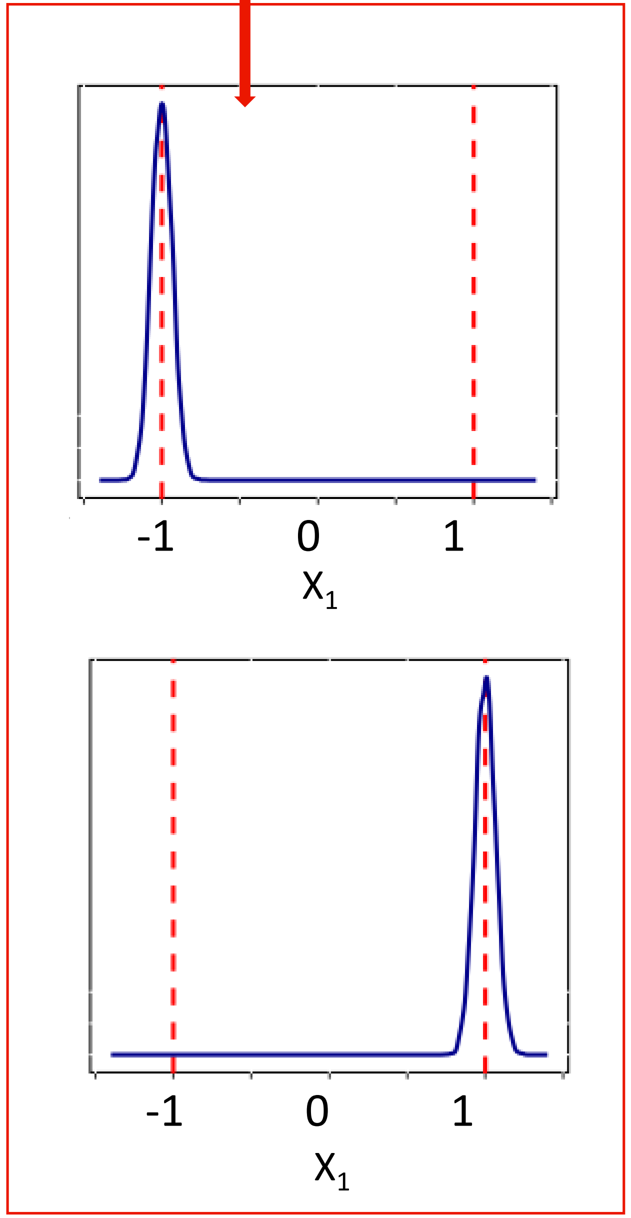
PAR defined in ICH Q8 (R2) : “a characterized range of a process parameter for which operation within this range, while keeping other parameters constant, will result in producing a material meeting relevant quality criteria”.

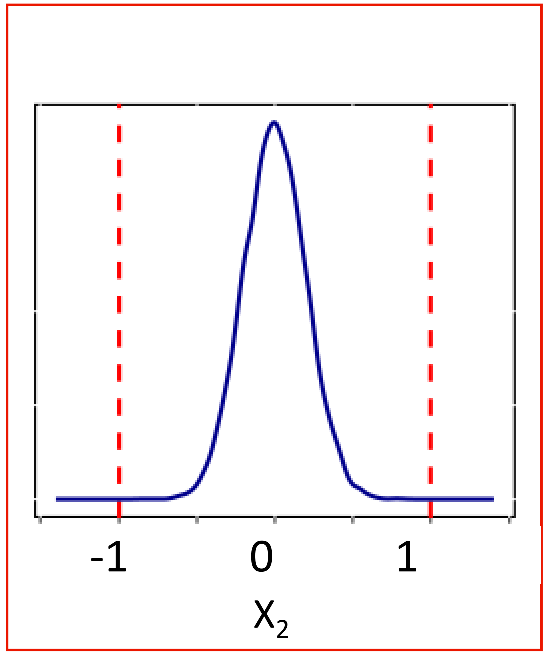
at target or normal operating range

Proven acceptable range (PAR) for X_1 in Monte Carlo simulation

Run simulations at the extreme case conditions for X_1, while keeping X_2 at target

$CQA = 94.7 - 2 * X_1 + 4.3 * X_2 - 3 * X_1 X_2$



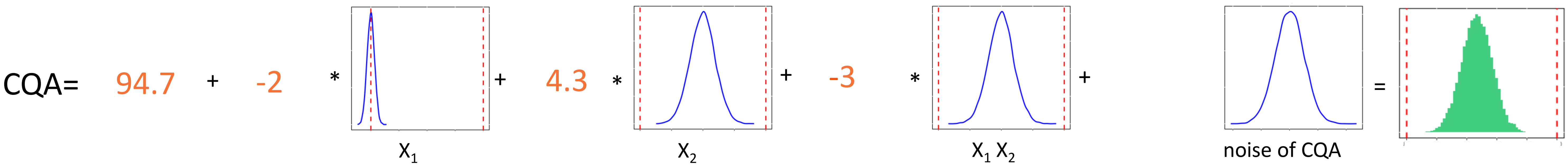


Proven acceptable range (PAR) for X_1 in Monte Carlo simulation

Run simulations at the extreme case conditions for X_1, while keeping X_2 at target

$CQA=94.7 -2* X_1 +4.3* X_2 -3* X_1 X_2$

Extreme Condition 1(X_1 =-1)



We applied fixed coefficients, which didn’t account for model uncertainty ~~~

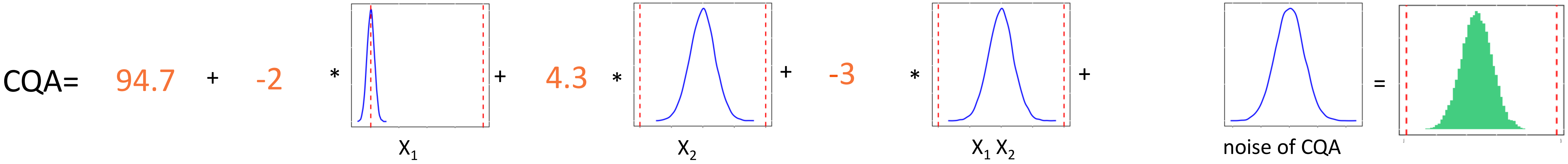
Can we improve?

Yes ! We can simulate Y using a distribution of model coefficients, thanks to Bayesian

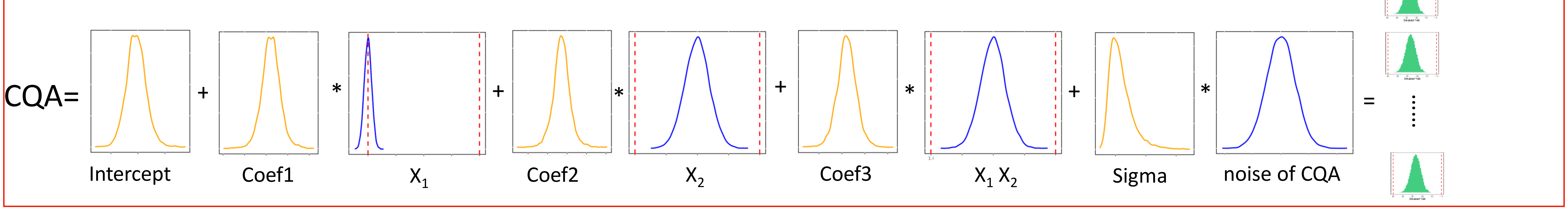
Run simulations at the extreme case conditions for X_1, while keeping X_2 at target

$CQA=94.7 -2* X_1 +4.3* X_2 -3* X_1 X_2$

Extreme Condition 1(X_1 =-1)--**Frequentist**

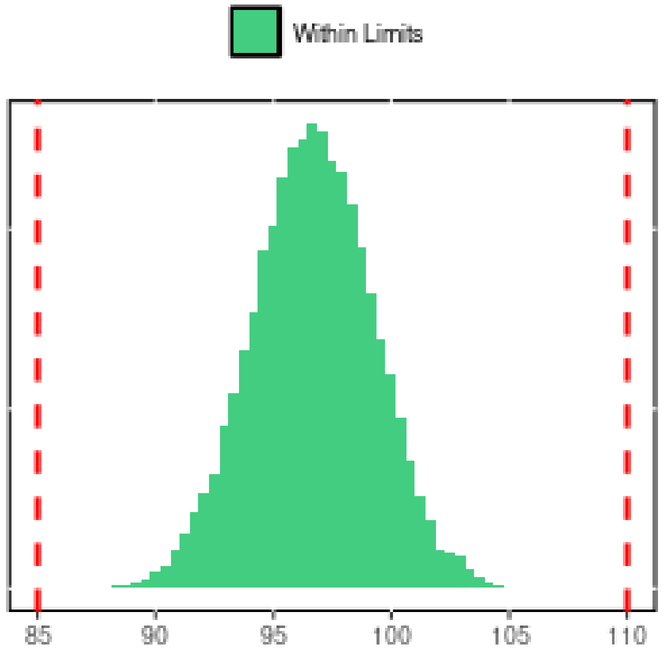


Extreme Condition 1(X_1 =-1)--**Bayesian**



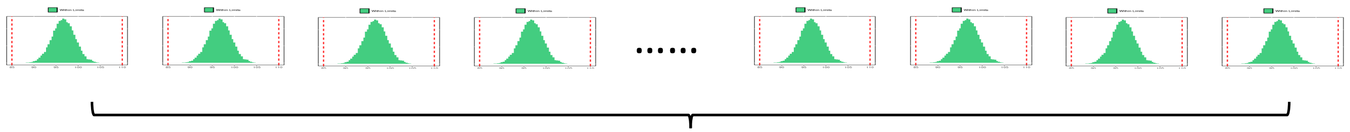
A single estimate of failure %(Frequentist) v.s. a distribution of failure % (Bayesian)

Frequentist

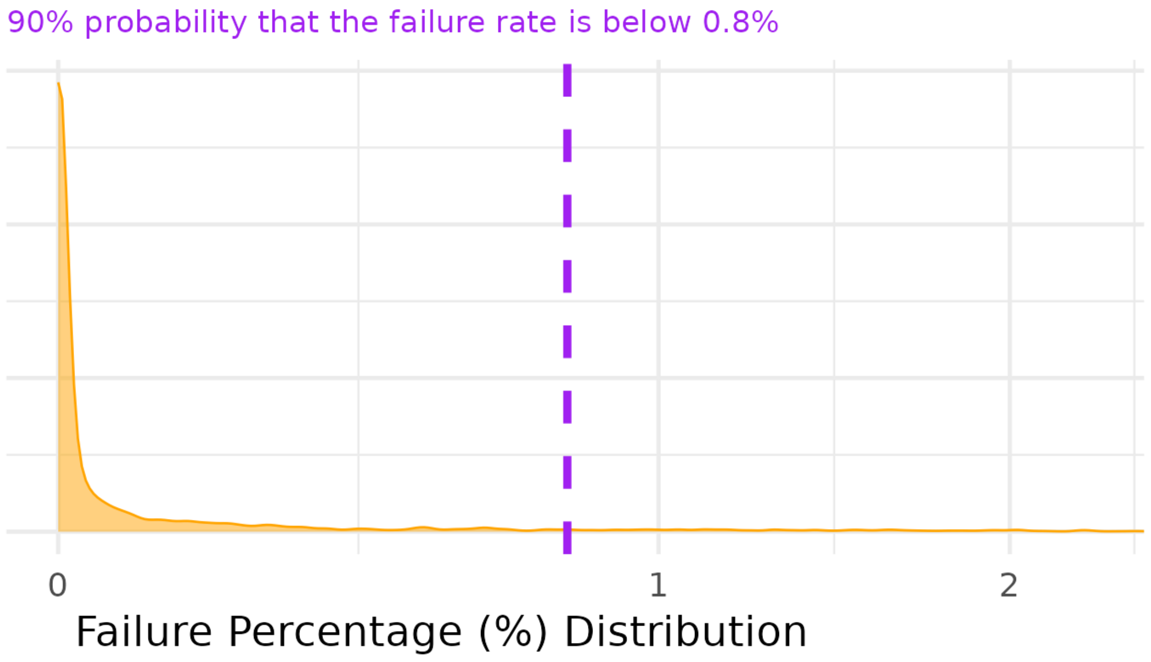


Failure Rate= 0%

Bayesian



Failure Rate =



Upper Bound	50%	80%	90%	95%	99%
Failure (%)	0.0	0.2	0.8	2.3	9.2

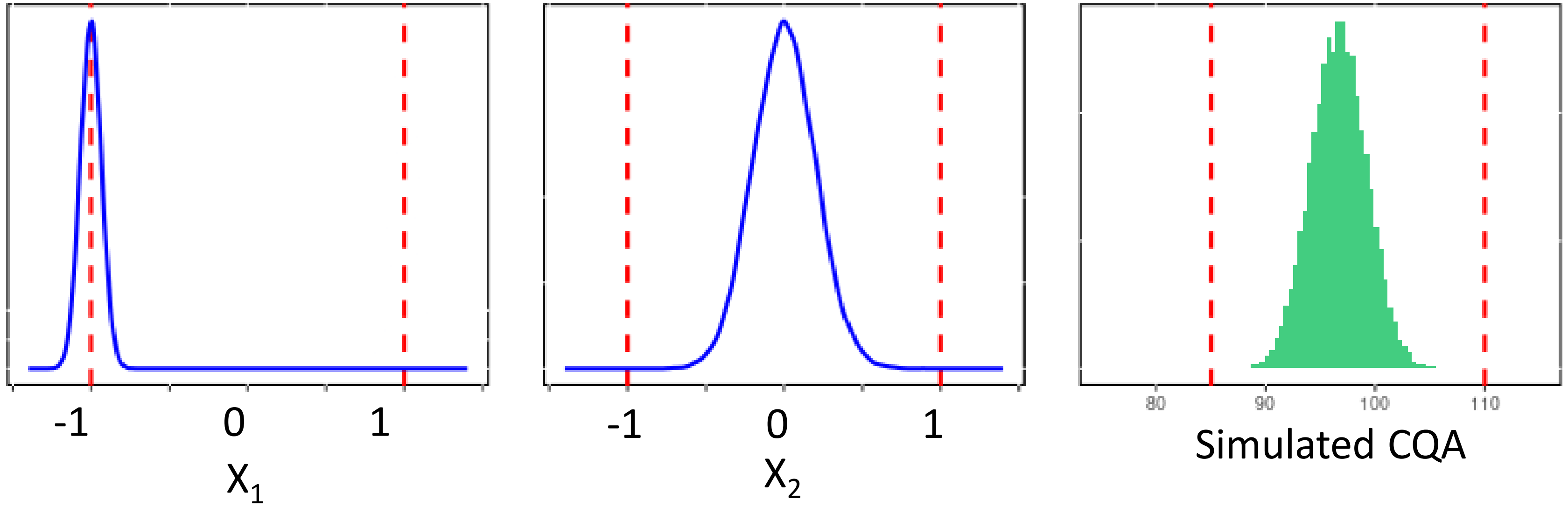
Failure Rate=% simulated CQA fall outside the specifications/Acceptable limits

Bayesian approach additionally accounts for model uncertainty

Frequentist Simulation at X1=-1.00
Failure=0%

Within Limits

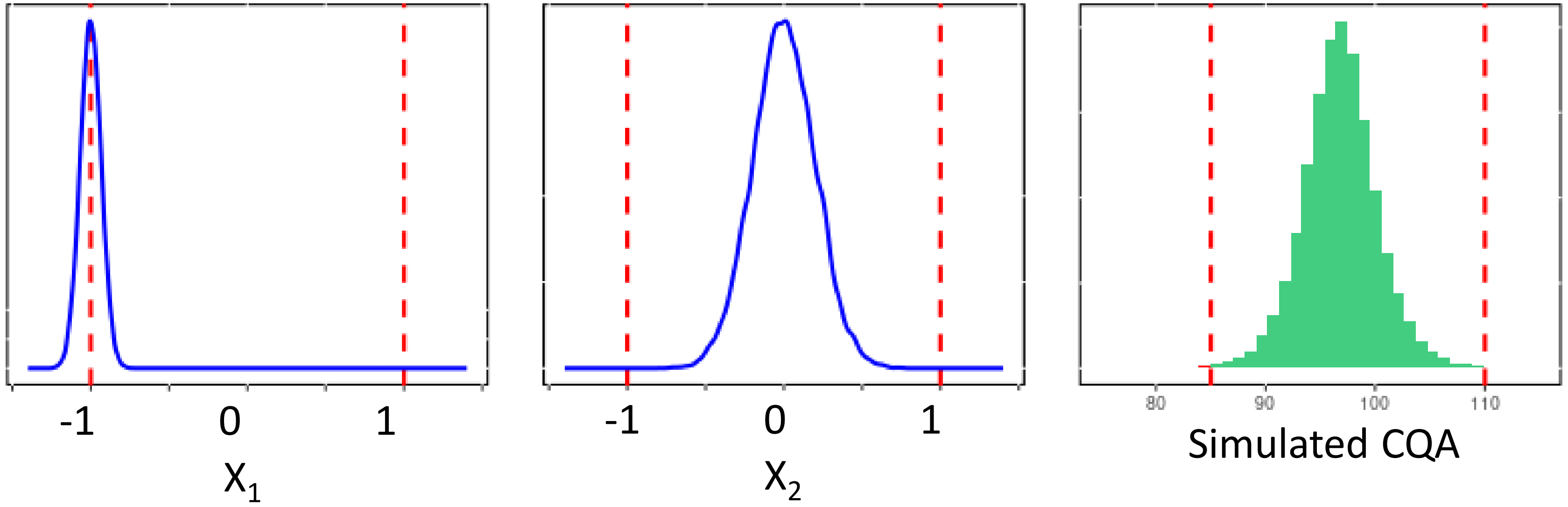
Account for noises of (parameters +CQA)



Bayesian Simulation at X1= -1.00
90% probability that Failure is below 0.8%

Outside Limits
Within Limits

Account for noises of (parameters +CQA)+ model uncertainty



J&J Innovative Medicine

PAR of X_1 might need to be narrowed to ensure a failure rate <5% with 90% probability

What we proposed

Bayesian enhanced statistical workflow to facilitate the decision-making in the context of process characterization

Future work

Improve Bayesian distribution with scientific knowledge (e.g., informative prior)

- **Continue the discussion to finalize the metrics**
- **Run more proof-of-concept examples**

sanofi

William Eberley

Yi Li

Canghai Lu

Yukun Ren

Jiaya Zhang

Johnson&Johnson

Bill Pikounis

Hong Shen