



Enhancing First-Time-in-Human (FTiH) Trials through Integrated Pharmaceutical Statistics: the Critical Role of Chemistry, Manufacturing, and Controls (CMC), Non-Clinical, and Translational Phases

Dan Lin

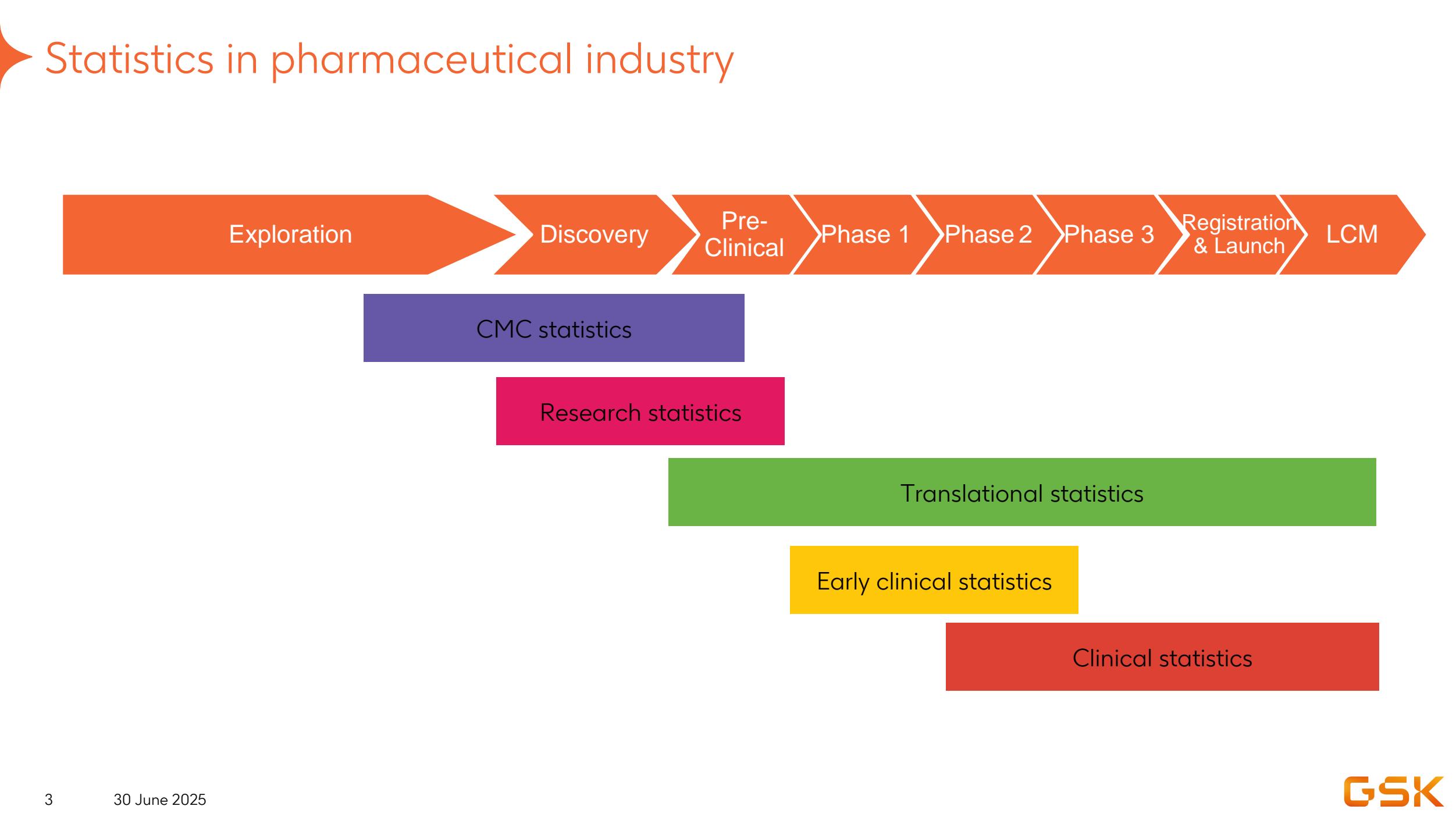
I am an employee of the GSK group of companies.
This work was sponsored by GlaxoSmithKline Biologicals SA.

Outline

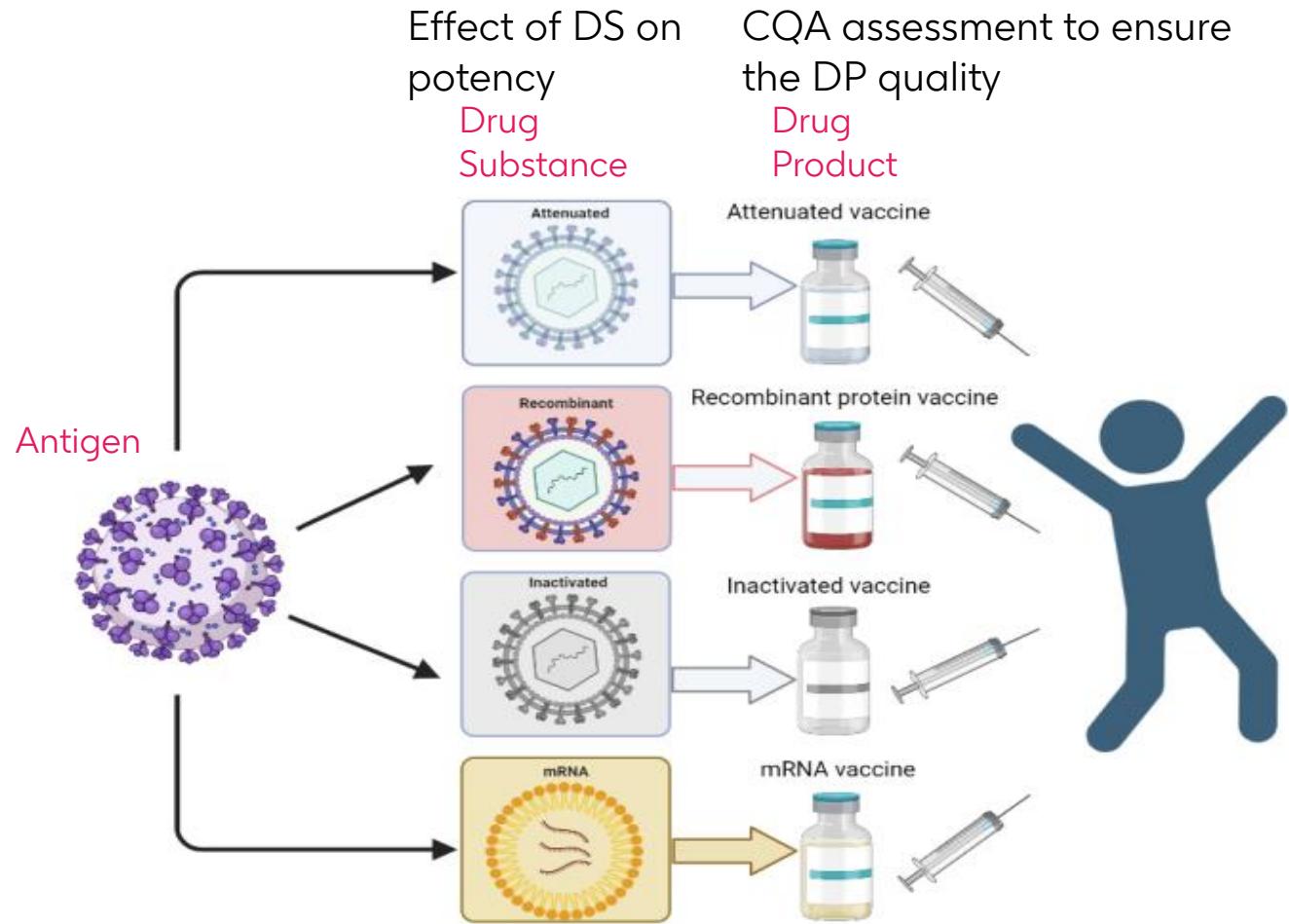
Role of non-clinical statistician

- CMC statistics
- Pre-clinical statistics
- Translational statistics

1. DoE in process development and optimisation
2. Assay qualification & validation
3. Non-clinical study types
4. Translational model from CMC, animal to human
5. Translational topic: dimensionality reduction and single cell data visualisation



Vaccine potency and immunization effect



Matthieu Van Tilbeurgh et al., Predictive Markers of Immunogenicity and Efficacy for Human Vaccines

Type of vaccines (examples):

- Live attenuated: MMR (measles, mumps, rubella), Rota virus, Dengue, Polio, TB
- Recombinant: Hep B, HPV, COVID (Novavax)
- Inactivated: Polio, HepA, pertussis
- mRNA: COVID 19, RSV ped (under dev), Influenza (under dev)

Effect of immunization on human immune system:

- Humoral response (potency)
- Cellular response (functional)
- Neutralization of virus (functional)
- Bacterial adhesion (functional)

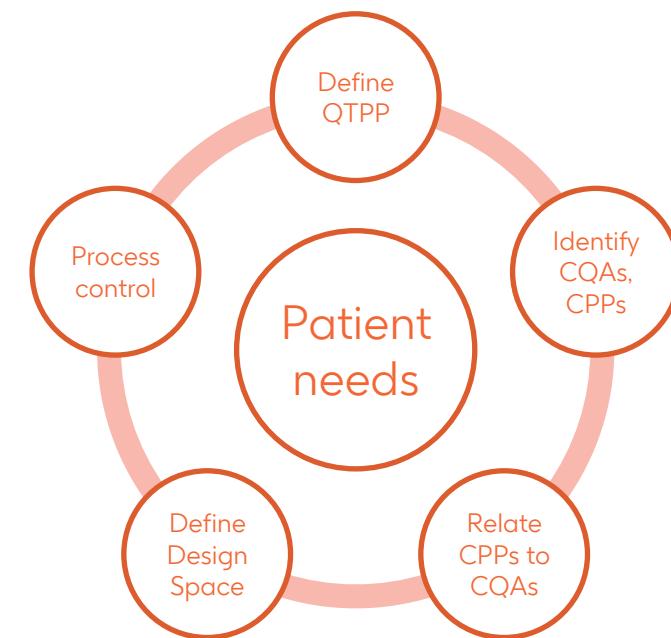
Objective of CMC development

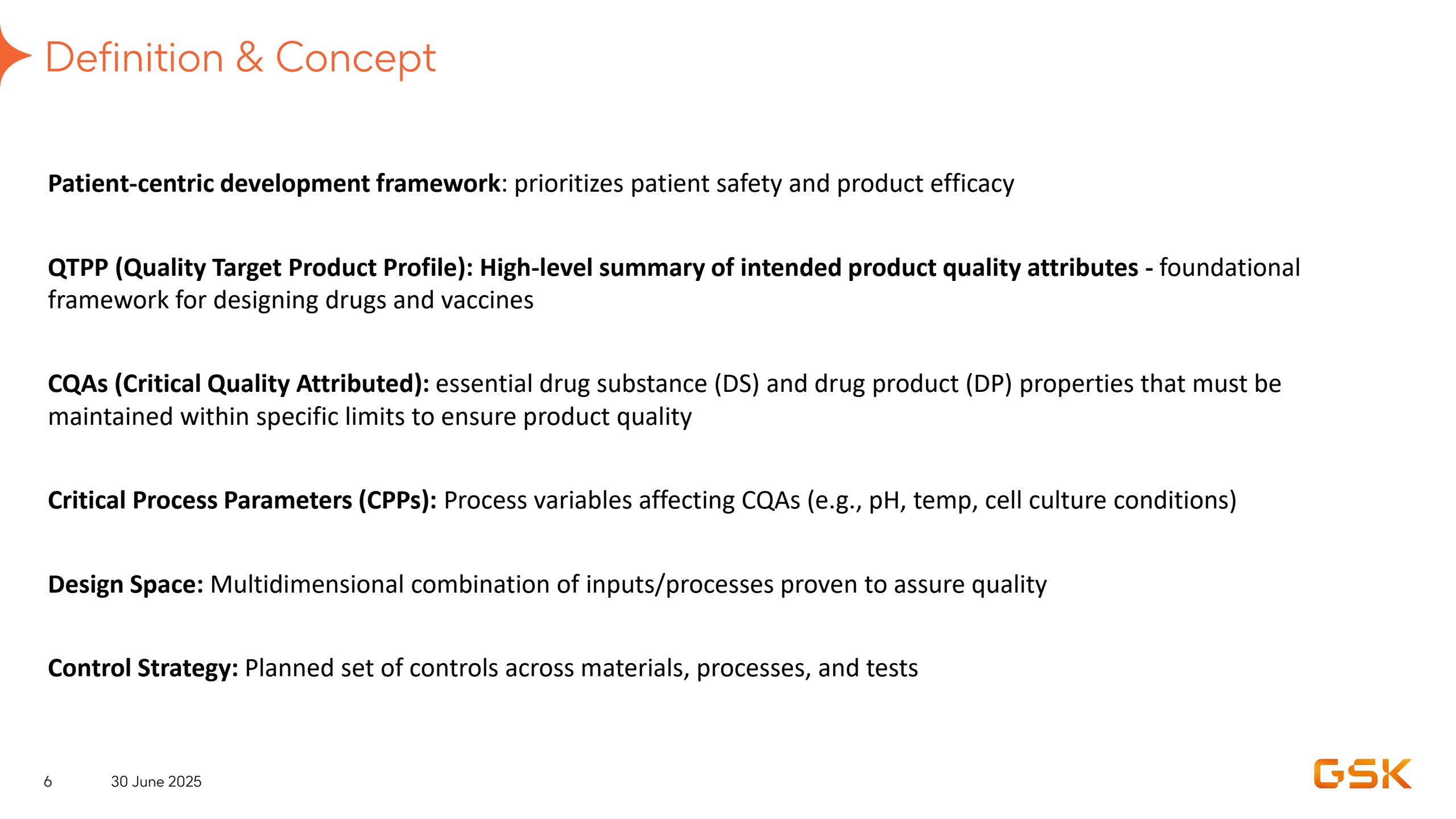
Develop robust processes & analytical methods ensuring product quality in a QbD framework

Quality by Design: ICH Q8 (R2) applies to both pharmaceutical and biological products

- Quality Target Product Profile (QTPP) definition
- Critical Quality Attributes (CQAs) **identification**
- Ensure quality over its **shelf life**
- CQAs – **design space** to ensure quality
- Control Strategy – **specification limit**

Patient Centric Approach





Definition & Concept

Patient-centric development framework: prioritizes patient safety and product efficacy

QTPP (Quality Target Product Profile): High-level summary of intended product quality attributes - foundational framework for designing drugs and vaccines

CQAs (Critical Quality Attributed): essential drug substance (DS) and drug product (DP) properties that must be maintained within specific limits to ensure product quality

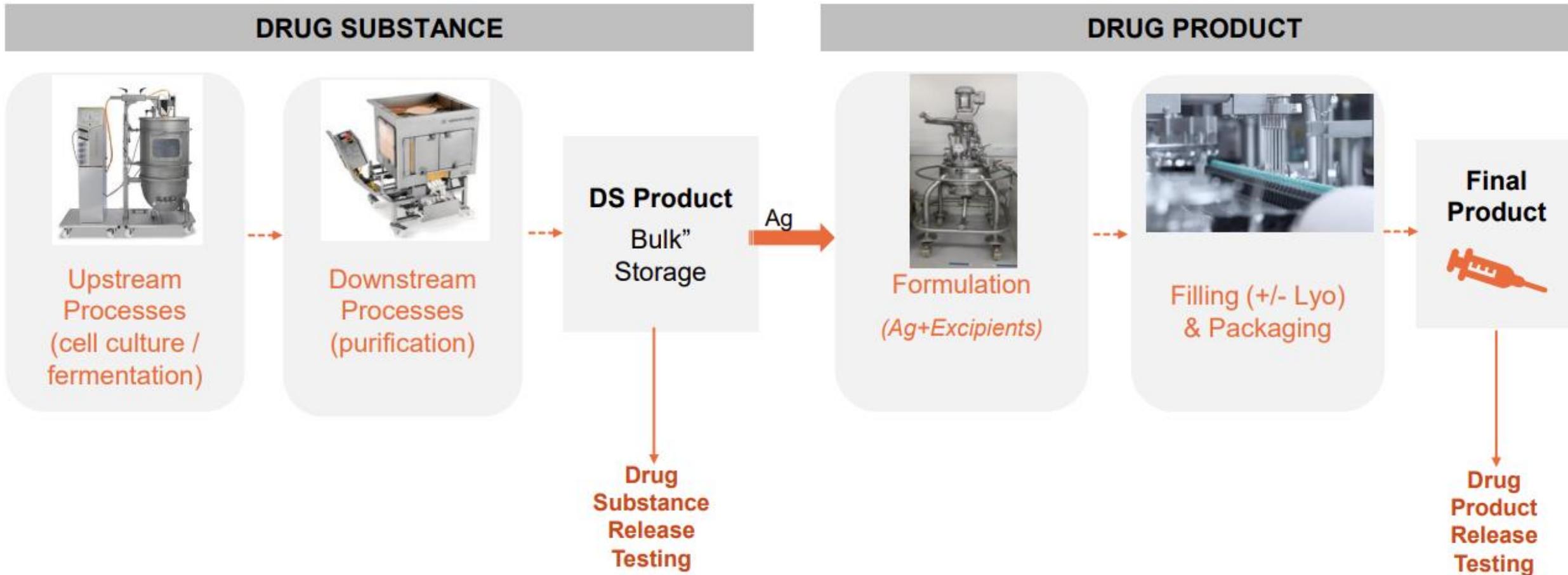
Critical Process Parameters (CPPs): Process variables affecting CQAs (e.g., pH, temp, cell culture conditions)

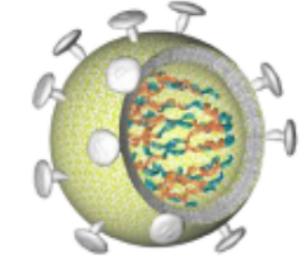
Design Space: Multidimensional combination of inputs/processes proven to assure quality

Control Strategy: Planned set of controls across materials, processes, and tests

Vaccine development process

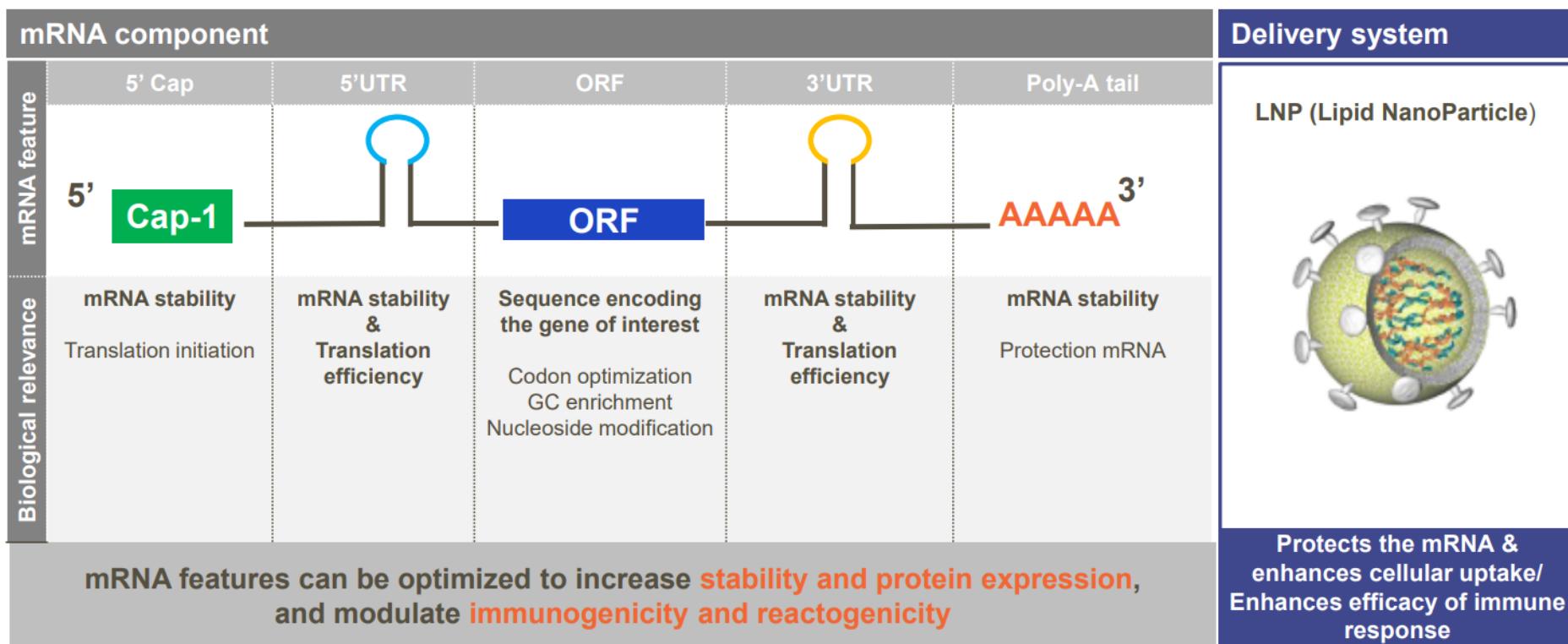
From antigen to final product in Syringe / vials



mRNA component						Delivery system	Core Principle
mRNA feature	5' Cap	5'UTR	ORF	3'UTR	Poly-A tail	LNP (Lipid NanoParticle)	LNP delivers the mRNA encoding immunogen(s) into the host cell cytoplasm where it is translated into protein(s) that will be then recognized by the immune system
Biological relevance	mRNA stability Translation initiation	mRNA stability & Translation efficiency	Sequence encoding the gene of interest Codon optimization GC enrichment Nucleoside modification	mRNA stability & Translation efficiency	mRNA stability Protection mRNA		Protects the mRNA & enhances cellular uptake/ Enhances efficacy of immune response
mRNA features can be optimized to increase stability and protein expression, and modulate immunogenicity and reactogenicity							

Non-replicative mRNA-based vaccines

components: mRNA (ORF, CAP, UTRs, Poly-A tail) +LNP (different lipids)



Nance and Meier. ACS Cent Sci. 2021;7:748–756; Linares-Fernández et al. Trends in Molecular Medicine. 2021;26(3):311-323; Jackson npj Vaccines 2020

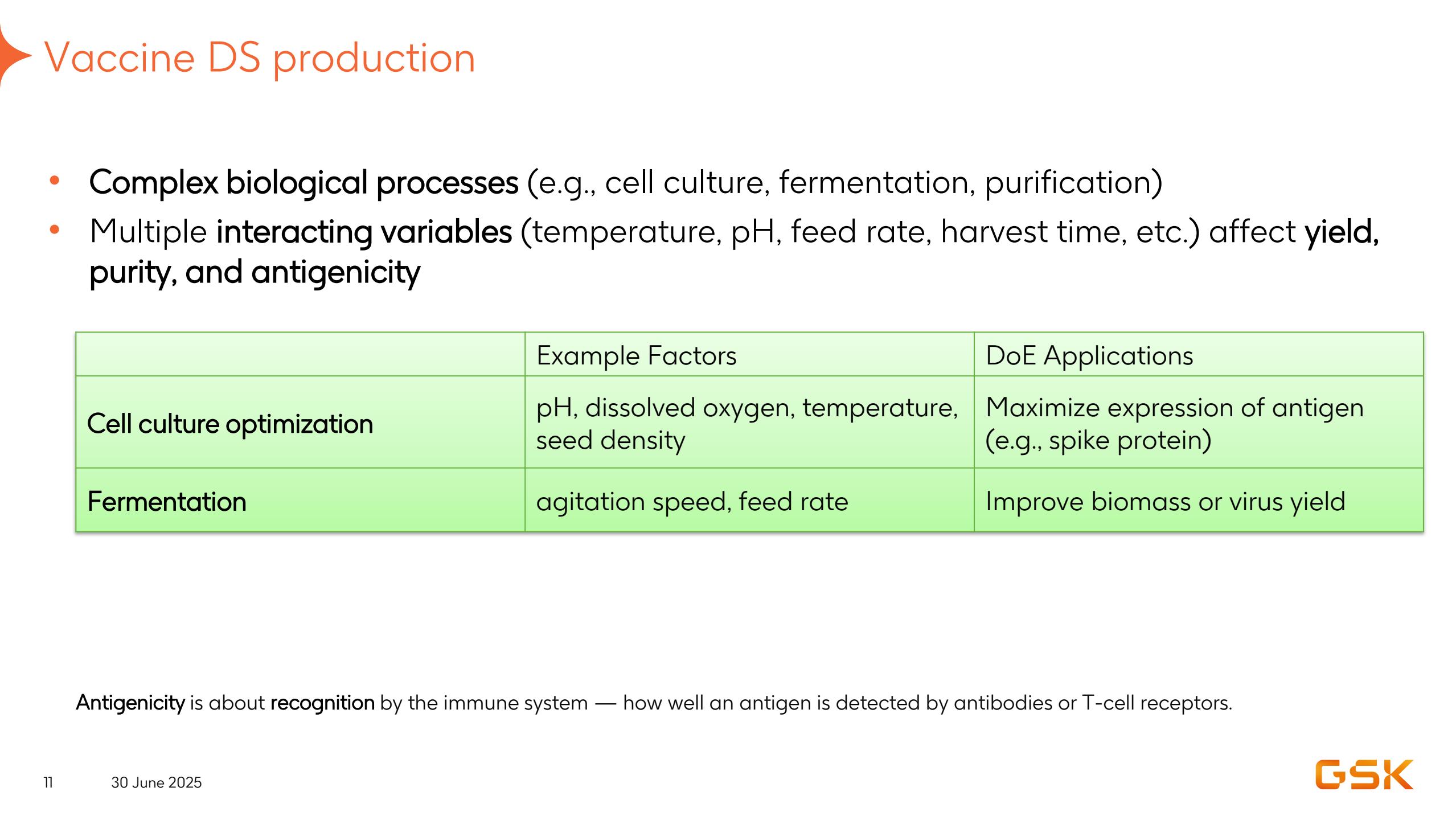
CAP, DNA, deoxyribonucleic acid; LNP, lipid nanoparticle; mRNA, messenger ribonucleic acid;
UTR, untranslated region

mRNA=Messenger RiboNucleic Acid
LNP = Lipid NanoParticle

ORF=Open Reading Frame
UTRs=UnTranslated Regions

Design of Experiment





Vaccine DP – formulation and stability

- Formulation influences **stability, potency, and delivery** (e.g., adjuvants, preservatives, pH)
- **Robust vaccine formulation** under QbD framework

	Example Factors	DoE Applications
Lyophilization	Freezing rate, excipients, drying temp	Optimize cake appearance and reconstitution
Stability testing	Temp, humidity, container type	Assess degradation pathways and shelf life



Glass vials

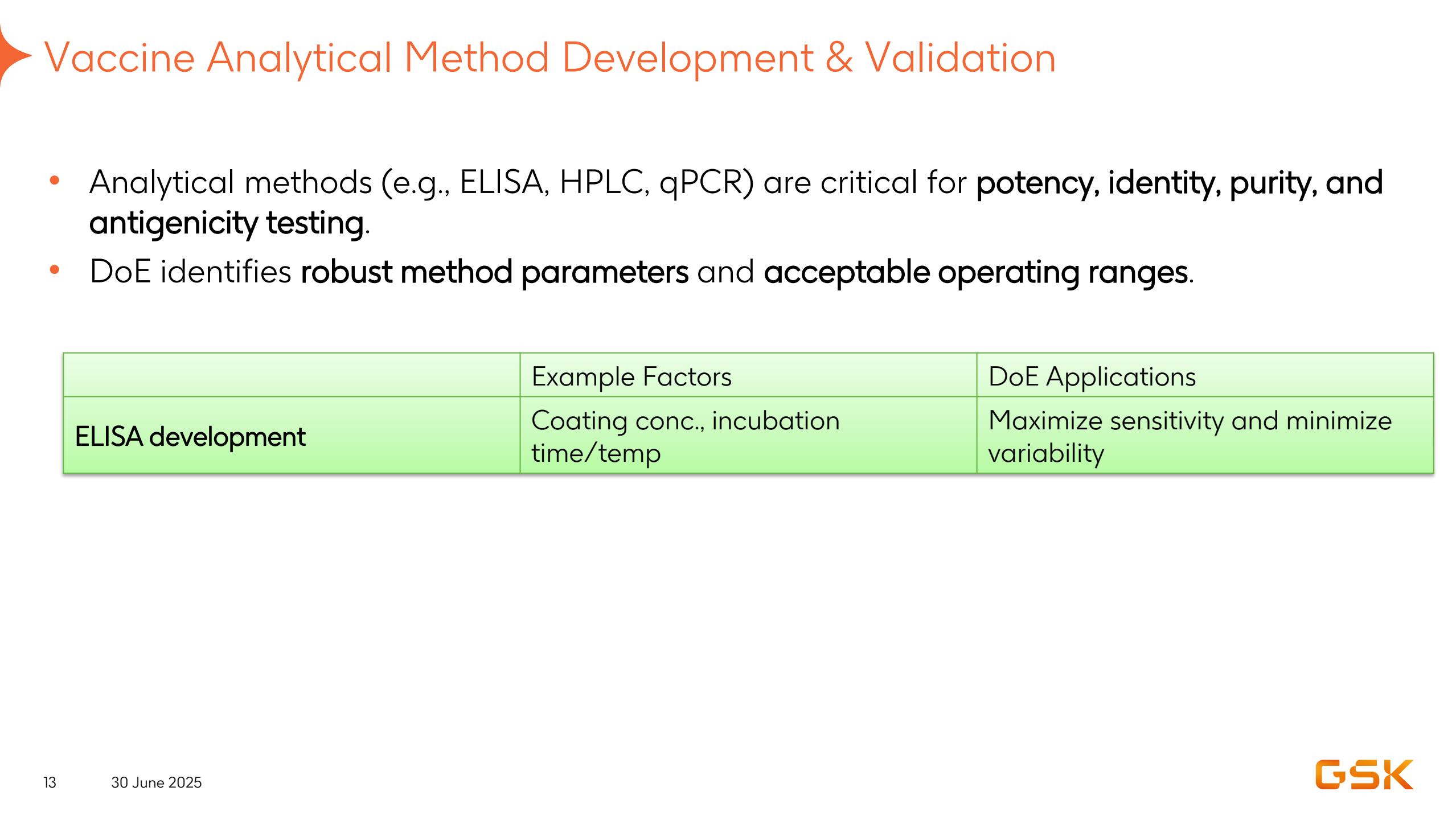


Pre-filled syringe (PFS)

12

30 June 2025





Vaccine Analytical Method Development & Validation

- Analytical methods (e.g., ELISA, HPLC, qPCR) are critical for **potency, identity, purity, and antigenicity testing**.
- DoE identifies **robust method parameters and acceptable operating ranges**.

	Example Factors	DoE Applications
ELISA development	Coating conc., incubation time/temp	Maximize sensitivity and minimize variability



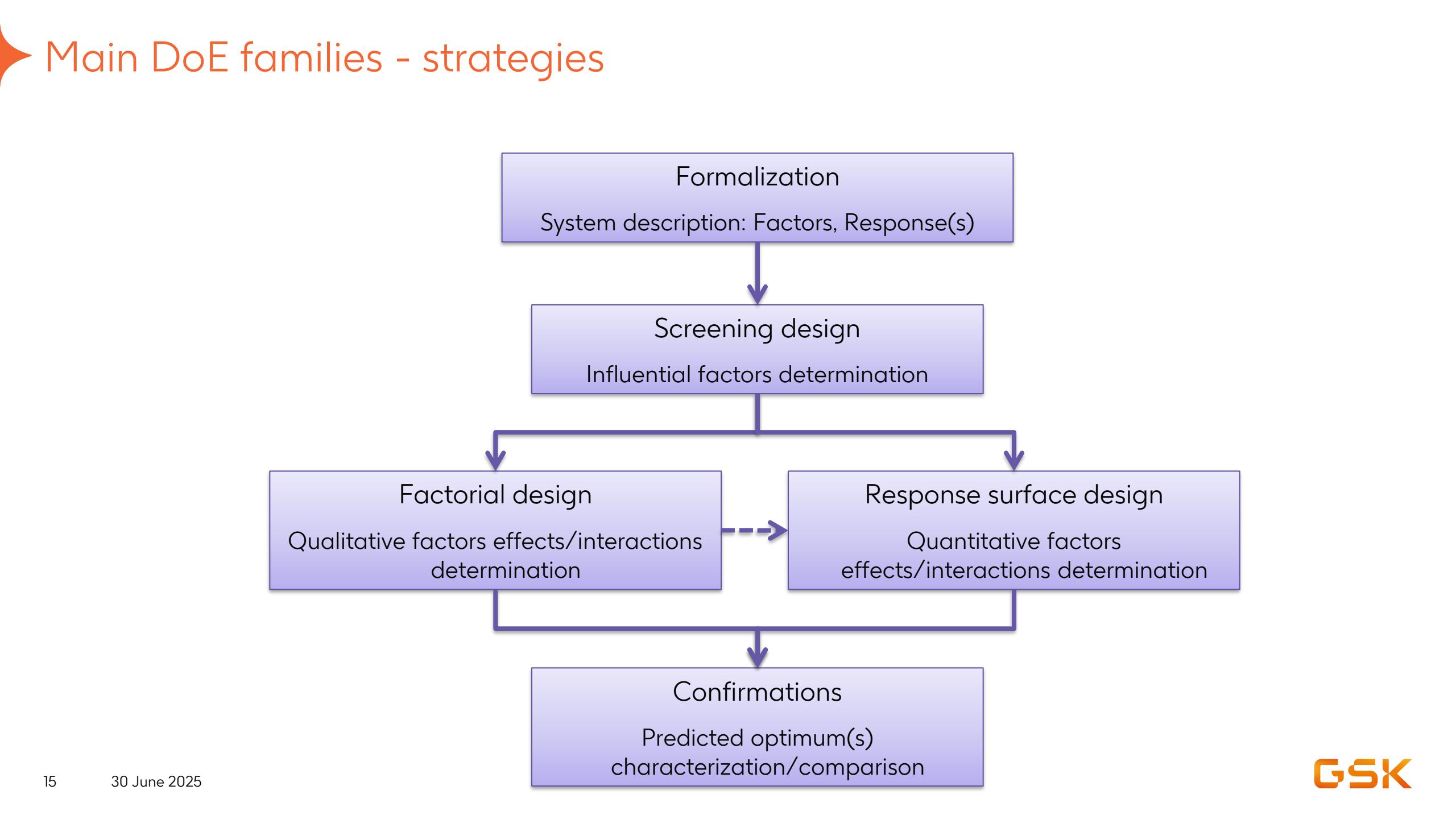
Quotes...

“An approximate answer to the right problem is worth a good deal more than an exact answer to an approximate problem.”

John Tukey

“To call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may be able to say what the experiment died of.”

Ronald Aylmer Fisher



DoE definition

Define → Measure → Analyze → Improve → Control

Screening of multiple factors (M)

Which factors have the **biggest impact** on the outputs?

Process understanding (A)

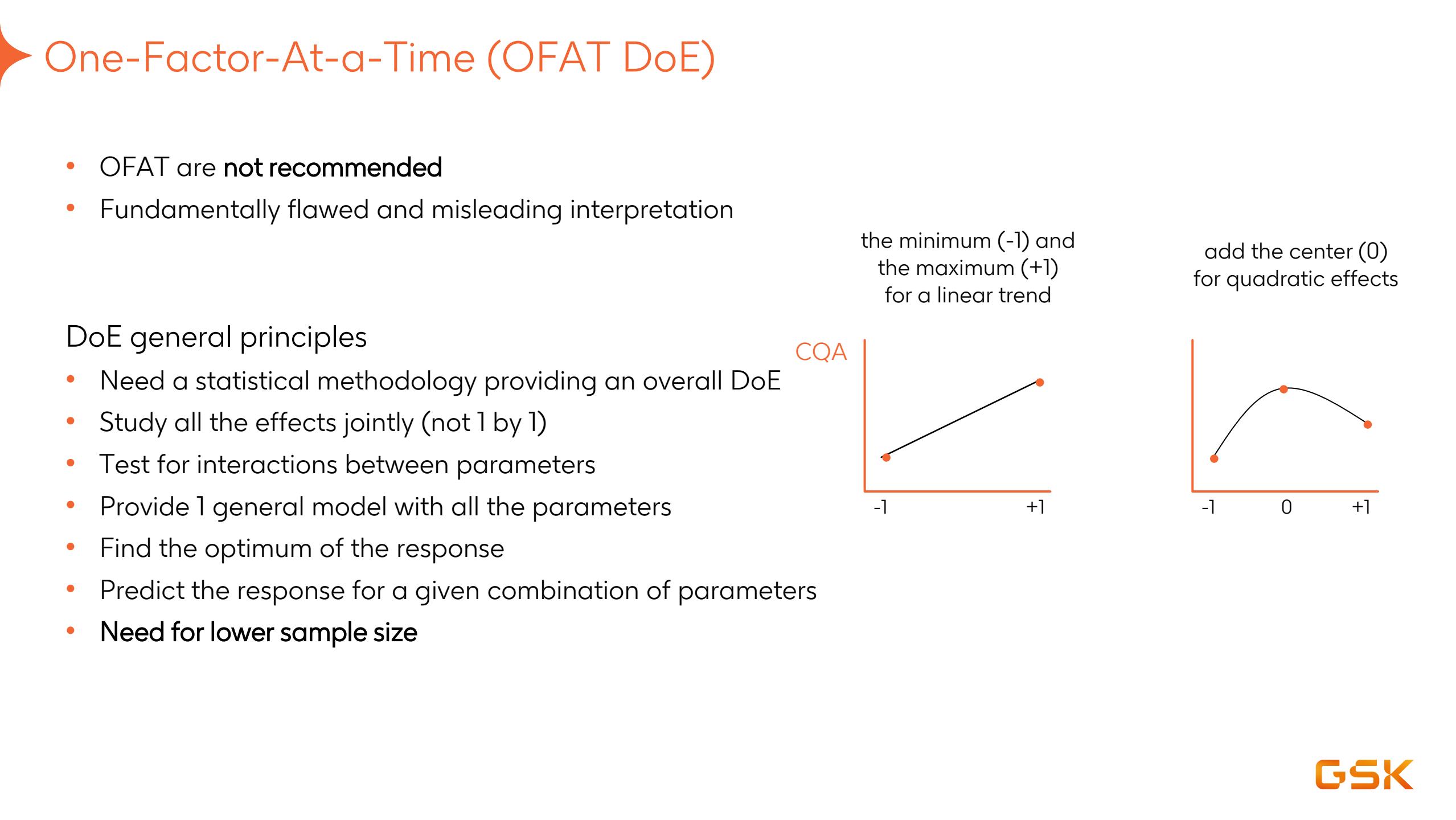
How do the input **factors interact and impact** the response?

Process optimization (I)

What are the **factor best settings** that optimize the response?

Process robustness (C)

How do **small variations of the input factors** impact the response?



DoE types

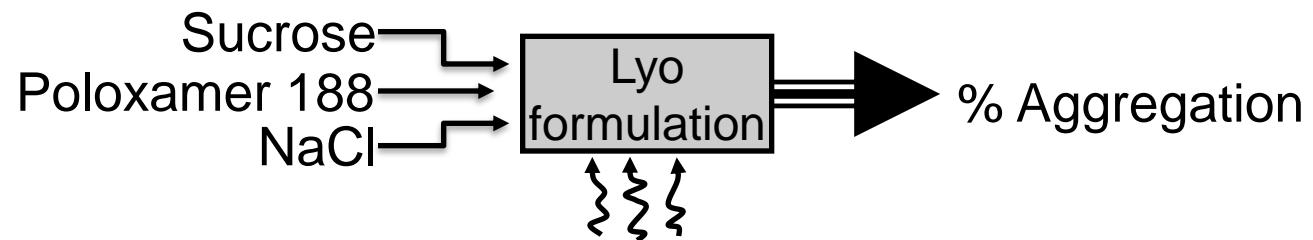
Covered here

- Process Screening:
 - Definitive Screening Design
 - Full factorial / fractional factorial
- Process Optimizations: Response surface modeling
 - Central Composite Design (CCD)
 - Custom Optimal Design
- Screening and optimization simultaneously:
 - OMARS (Orthogonal Minimally Aliased Response Surface) design

Factorial design application

Formulation optimization

Objective: Test several excipients to reduce protein aggregation after lyophilisation



Tested levels:

- Sucrose: 3 - 10%
- PX188: 0 - 0.2%
- NaCl: 0 - 80mM

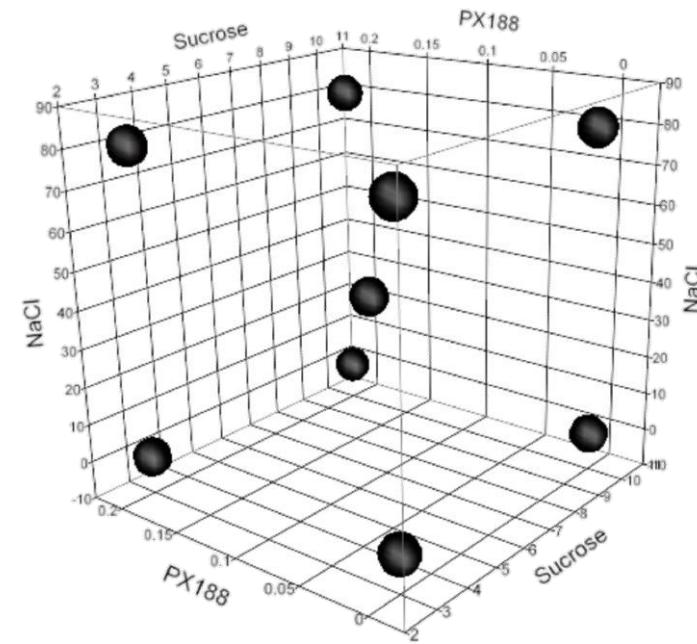
What is their impact on protein aggregation, and how do they interact ?
=> **Factorial design**

Factorial design application

8 combinations of **qualitative** factors (2 levels) + 3x central point = 11 experiments

	Factors		
	Sucrose	PX188	NaCl
Experiments	1	3	0
	2	10	0
	3	3	0.2
	4	10	0.2
	5	3	0
	6	10	0
	7	3	0.2
	8	10	0.2
	9	6.5	0.1

X3 (central conditions)



At the end of each experiment, final read-out collected (% aggregation) and analyzed
Analysis of variance: significance of the **main and interaction effects**

Response surface design

When question is: how does the response “behave” into the **range** associated to studied factors?

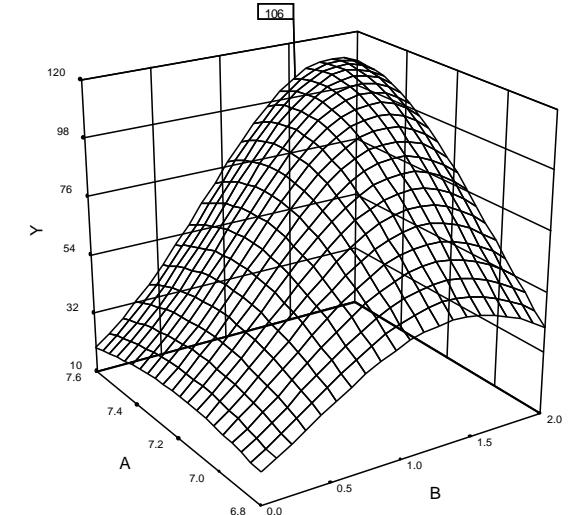
Factors: quantitative

Pros :

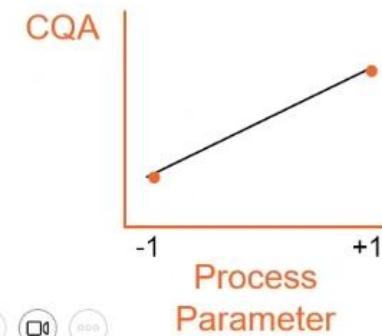
- Allow to predict responses in the whole domain
- The link between response(s) and quantitative factors can be non-linear

Cons :

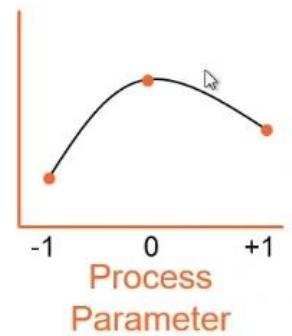
- This link must stay simple (to be modeled)
- Results are more complex to handle



the minimum (-1) and
the maximum (+1)
for a linear trend



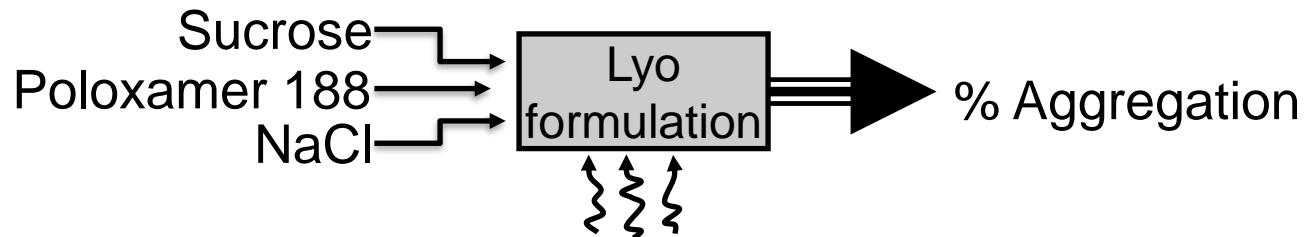
add the center (0)
for quadratic effects



Response surface design application

Formulation optimization

Objective: **optimize the dose** of several excipients to reduce protein aggregation after lyophilisation



Tested domains:

- Sucrose: 3 to 10%
- PX188: 0 to 0.2%
- NaCl: 0 to 80mM

Because we want to optimize the doses:
=> **response surface design**.

Note: this design may also be used to generate knowledge around already determined formulation (robustness)

Response surface design application

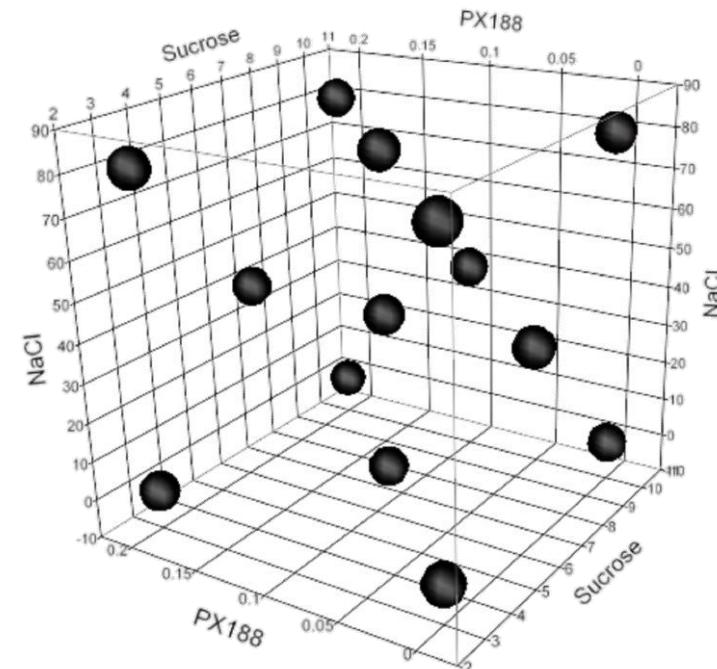
Central composite design

15 combinations of **quantitative** factors, 17 experiments

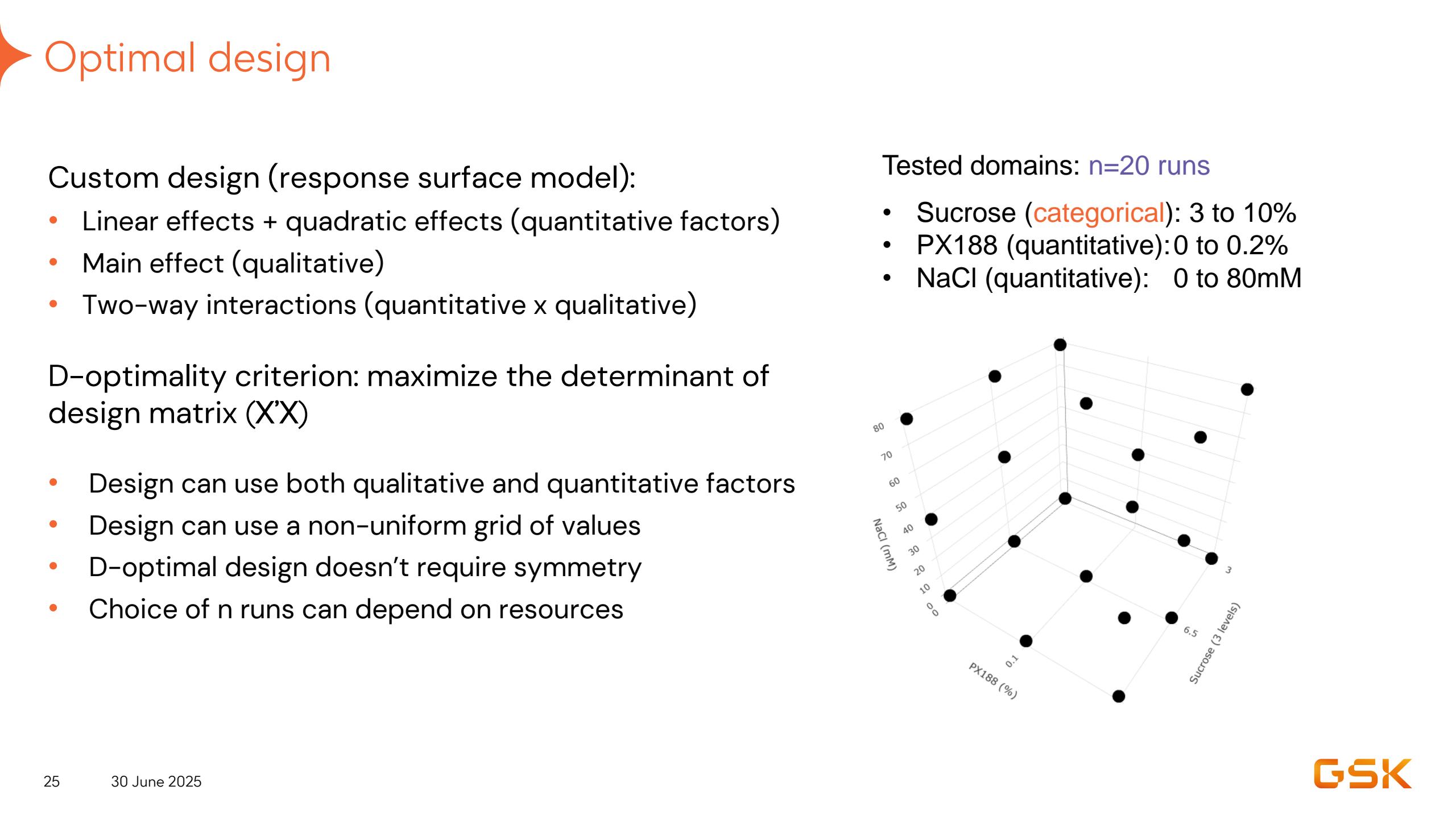
Experiments	Factors		
	Sucrose	PX188	NaCl
1	3	0	0
2	10	0	0
3	3	0.2	0
4	10	0.2	0
5	3	0	80
6	10	0	80
7	3	0.2	80
8	10	0.2	80
9	6.5	0.1	40
10	10	0.1	40
11	6.5	0	40
12	6.5	0.2	40
13	6.5	0.1	0
14	6.5	0.1	80
15	3	0.1	40

X3 (central conditions)

6 Axial points:
x2 per factor

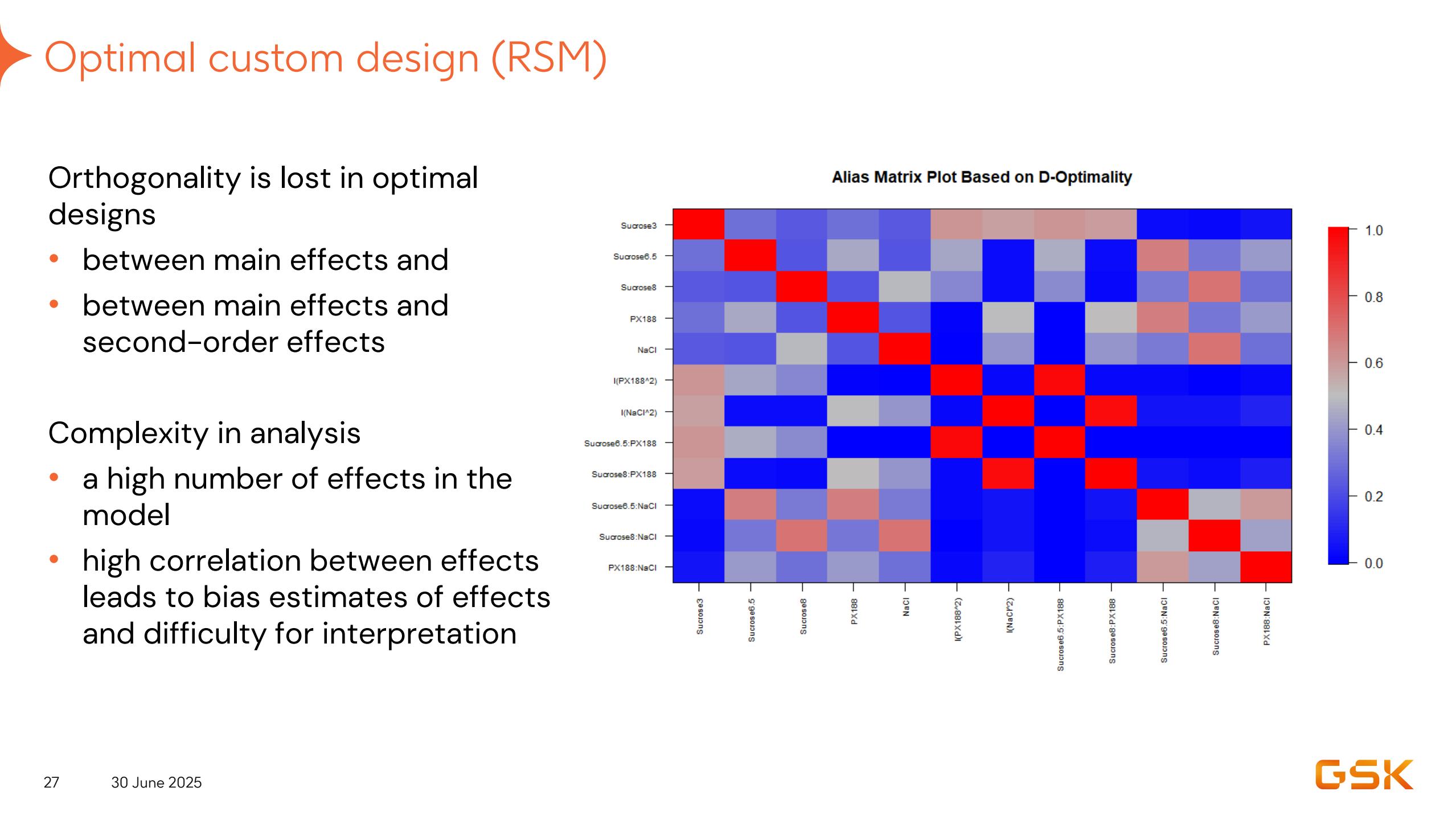


At the end of each experiment, final read-out collected (% aggregation) and analyzed
Analysis: include main and interaction effects as well as **quadratic effects**.



Choice of criterion for optimal design

Criterion	Description	Focus / Goal	Key Measure
D-Optimality (D-opt)	Maximizes determinant of design matrix ($X'X$)	Maximize the overall information content of the design	Determinant of design matrix ($X'X$)
A-Optimality (A-opt)	Minimizes trace of inverse design matrix ($(X'X)^{-1}$)	Minimize the average variance of parameter estimates	Trace of $(X'X)^{-1}$
E-Optimality (E-opt)	Maximizes smallest eigenvalue of $X'X$	Minimize the maximum variance among parameters	Smallest eigenvalue of $X'X$
I-Optimality (I-opt)	Minimizes average prediction variance over design space	Improve overall prediction precision	Average prediction variance
G-Optimality (G-opt)	Minimizes maximum prediction variance over design space	Control worst-case prediction variance	Maximum prediction variance
P-Optimality (P-opt)	Based on weighted sum or product of parameter variances	Flexible, tailored to specific parameters	Weighted variance function
V-Optimality (V-opt)	Minimizes average prediction variance at specified points	Prediction precision at chosen design points	Average variance of predictions at points
c-Optimality (c-opt)	Minimizes variance of a specific linear combination of parameters	Precision for specific contrasts or parameters	Variance of $c'\beta c^{-1}\beta$ (linear contrast)

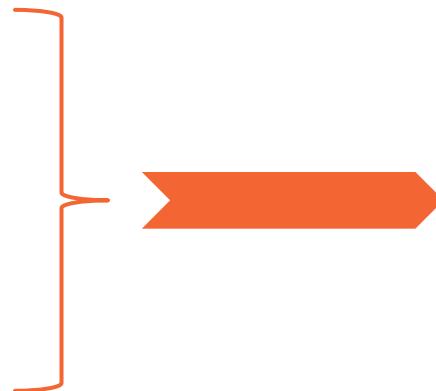


Screening design application

VERO Cells culture improvement

Objective: Maximization of the cells yield, by medium modification

Growth factors (groups)	Low level	High level
A : Insulin (mg/L)	5	15
B : EGF (μg/L)	5	15
C : Hormones I (μg/L)	0	16
	0	200
D : Hormones II	0	25
	0	33
E : Peptones I (mg/L)	100	200
	50	100
F : Peptones II (g/L)	0	1
	0	1
G : Cholesterol (mg/L)	0	0,2
H: Lipid (μg/mL)	1	10

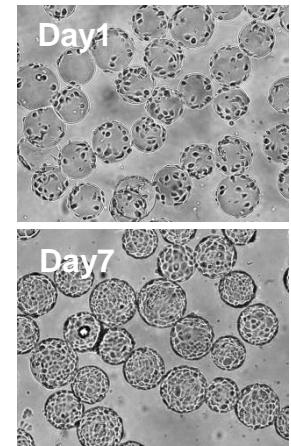


Cell culture



VERO Cells density

Among 8 factors, which ones impact the response significantly?
=> Screening design



Screening design application

Definitive Screening Design:

Cons: DSDs can be extended to incorporate **two-level categorical** factors, but the resulting designs are no longer orthogonal

Allow the estimation of

- Main effects
- **Quadratic effects**
- Two-factor interactions

Introduced by Jones & Nachtsheim
(Technometrics, 2011)

Fold-over runs in red boxes

	X1	X2	X3	X4	X5	X6	X7	X8
1	0	1	1	1	1	1	1	1
2	0	-1	-1	-1	-1	-1	-1	-1
3	1	0	1	1	-1	1	-1	-1
4	-1	0	-1	-1	1	-1	1	1
5	1	-1	0	1	1	-1	1	-1
6	-1	1	0	-1	-1	1	-1	1
7	1	-1	-1	0	1	1	-1	1
8	-1	1	1	0	-1	-1	1	-1
9	1	1	-1	-1	0	1	1	-1
10	-1	-1	1	1	0	-1	-1	1
11	1	-1	1	-1	-1	0	1	1
All rows	17							
Selected	0							
Excluded	0							
Hidden	0							
Labelled	0							

Screening design application

Definitive Screening Design:

Each run will always have a center point for a factor

	X1	X2	X3	X4	X5	X6	X7	X8
1	0	1	1	1	1	1	1	1
2	0	1	-1	-1	-1	-1	-1	-1
3	1	0	1	1	-1	1	-1	-1
4	-1	0	-1	-1	1	-1	1	1
5	1	1	0	1	1	-1	1	-1
6	-1	1	0	-1	-1	1	-1	1
7	1	-1	-1	-1	0	1	1	-1
8	-1	1	1	0	0	-1	-1	-1
9	1	1	-1	-1	0	1	1	-1
10	-1	-1	1	1	1	0	-1	1
11	1	-1	1	-1	-1	1	0	1
12	-1	1	-1	1	1	1	0	-1
13	1	1	-1	1	-1	-1	-1	0
14	-1	-1	1	-1	1	1	0	-1
15	1	1	1	-1	1	-1	-1	0
16	-1	-1	-1	1	-1	1	1	0
17	0	0	0	0	0	0	0	0

Goos and Núñez Ares. (webinar) OMARS Designs: Bridging the Gap between Definitive Screening Designs and Standard Response Surface Designs
<https://falltechnicalconference.org/wp-content/uploads/2017/11/webinar-FTC-Peter-Goos-OMARS-designs.pdf>

Screening design application

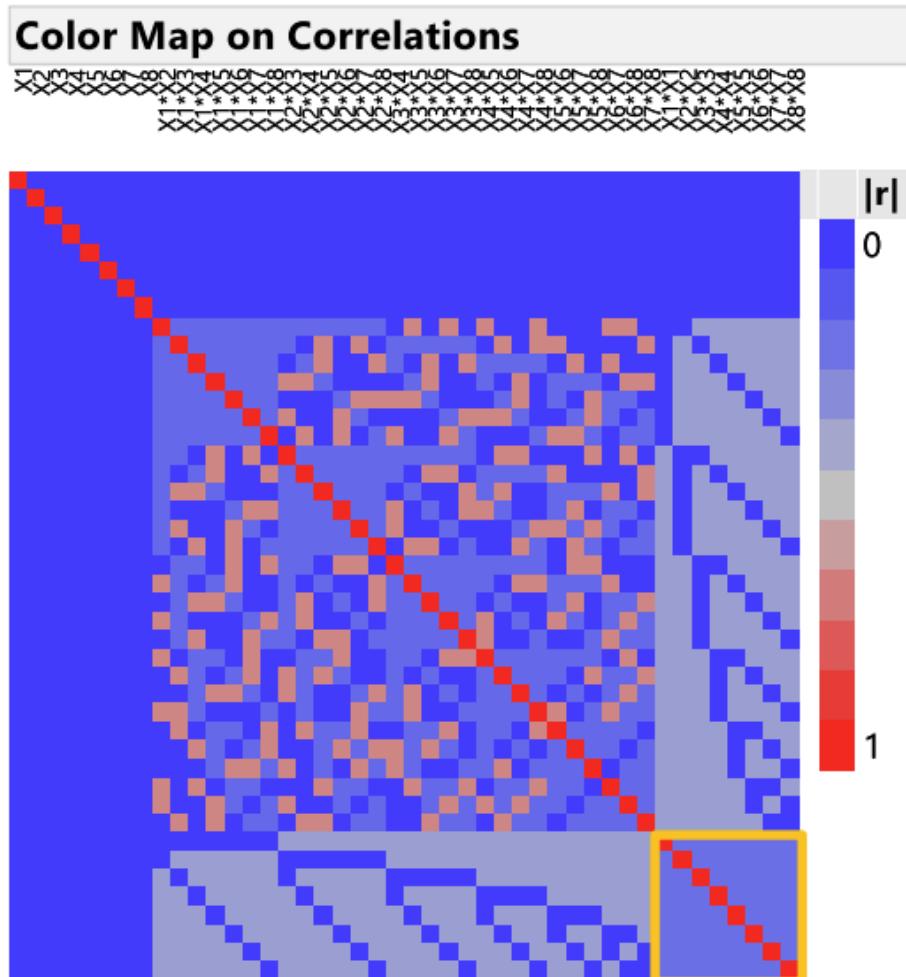
Definitive Screening Design:

One run of center points for all factors

	X1	X2	X3	X4	X5	X6	X7	X8
1	0	1	1	1	1	1	1	1
2	0	-1	-1	-1	-1	-1	-1	-1
3	1	0	1	1	-1	1	-1	-1
4	-1	0	-1	-1	1	-1	1	1
5	1	-1	0	1	1	-1	1	-1
6	-1	1	0	-1	-1	1	-1	1
7	1	-1	-1	0	1	1	-1	1
8	-1	1	1	0	-1	-1	1	-1
9	1	1	-1	-1	0	1	1	-1
10	-1	-1	1	1	0	-1	-1	1
11	1	-1	1	-1	-1	0	1	1
All rows	17	12	-1	1	-1	1	0	-1
Selected	0	13	1	1	-1	1	-1	0
Excluded	0	14	-1	-1	1	-1	1	0
Hidden	0	15	1	1	1	-1	-1	-1
Labelled	0	16	1	1	1	1	1	0
	17	0	0	0	0	0	0	0

Screening design application

Definitive Screening Design:



Orthogonality properties:

- Main effects are orthogonal to each other
- Main effects are orthogonal to two-way interactions and to quadratic effects
- Therefore, they are called minimally aliased designs
- Two-factor interactions are often strongly aliased with each other
- Two-factor interactions are sometimes strongly aliased with quadratic effects as well

Analysis can be **painful and leave ambiguity** if more than a few factors matter.

OMARS design

How many OMARS designs exist?

- **Orthogonal**

main effects estimated independently from each other

- **Minimally Aliased**

main effects estimated independently from all second-order effects

- **Response Surface Designs**

allow the estimation of a partial or complete second-order effects model

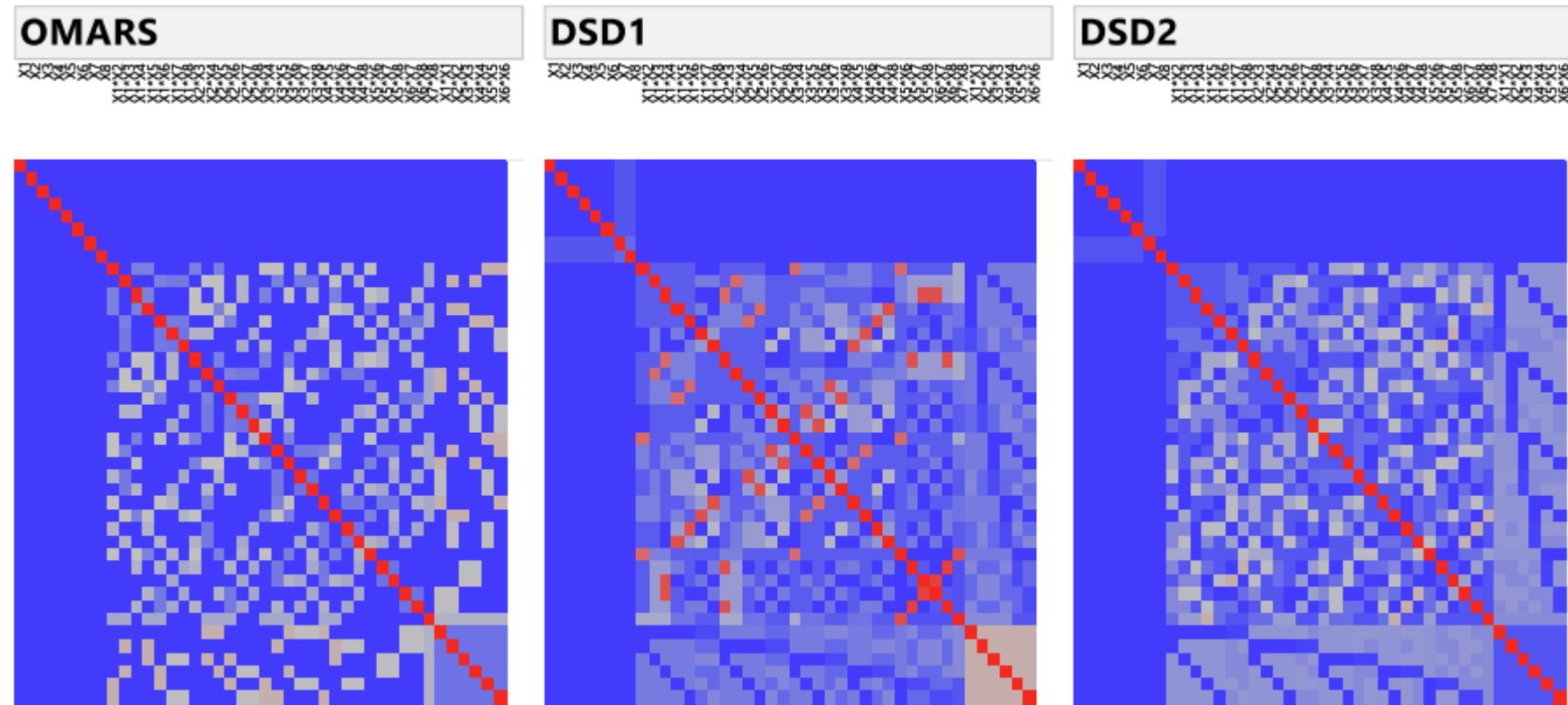
Can optimize multiple criteria (optimality, power, precision of estimate, etc) simultaneously

# Factors	3	4	5	6	7
# Runs	8-14	8-24	12-44	12-50	14-70
# Designs	5	41	5399	1406	1082

Integer linear programming techniques

Source: Núñez Ares, Goos 2020, 2023 & Núñez Ares, Schoen, Goos 2023

OMARS design vs. definitive screening designs for 6 quantitative and 2 categorical factors



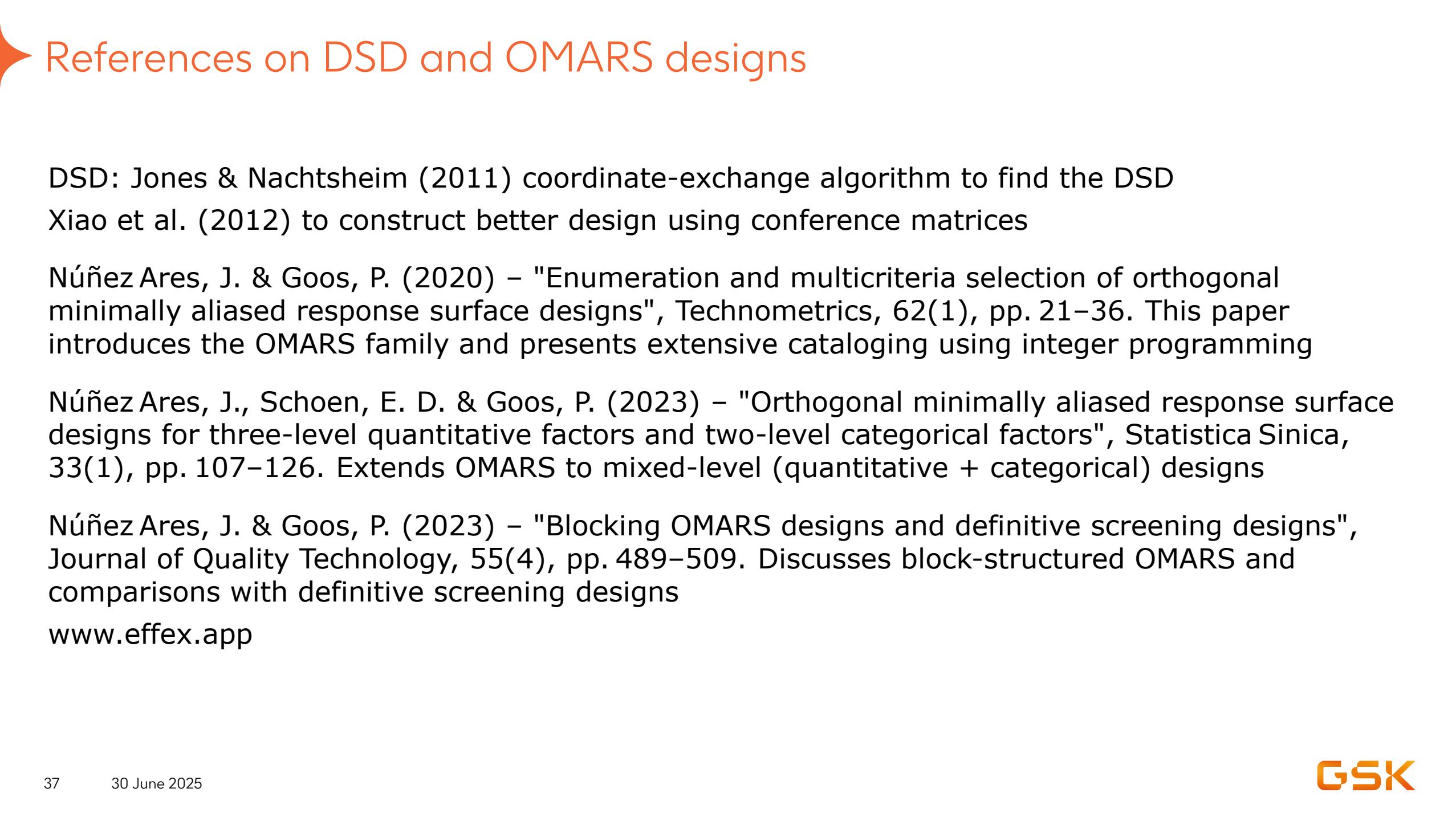
KU LEUVEN

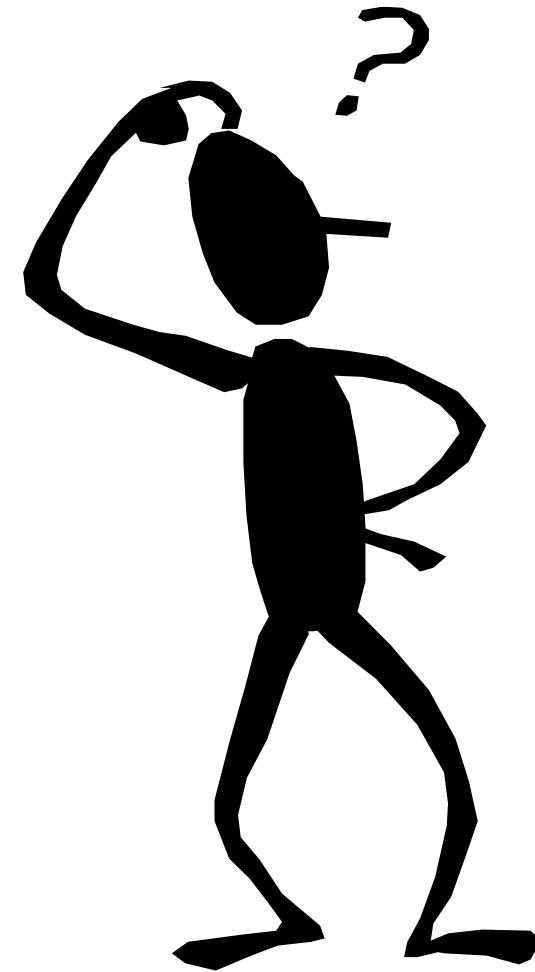
Goos and Núñez Ares. (webinar) OMARS Designs: Bridging the Gap between Definitive Screening Designs and Standard Response Surface Designs. <https://falltechnicalconference.org/wp-content/uploads/2017/11/webinar-FTC-Peter-Goos-OMARS-designs.pdf>



Summary

Objective	Design	Types of factors	Orthogonality	Model terms
Screening	Factorial	only qualitative factors	+ orthogonal of all effects	main and two-way interaction effects
	Fractional Factorial		+ orthogonal depending on resolution	
Optimisation	Central Composite	only quantitative factors	+ orthogonal depending on axial distance	main, two-way interaction and quadratic effects
	Optimal	both qualitative and quantitative factors	- high alias of effects	
Screening	Definitive Screening	both quantitative + categorical (2 levels) factors	+ main orthogonal of two-way and quadratic effects - aliasing and confounding among interaction and between interaction and quadratic effects	main, two-way interaction, and quadratic effects (quantitative factors)
Screening + Optimisation	OMARS	both categorical and quantitative factors		



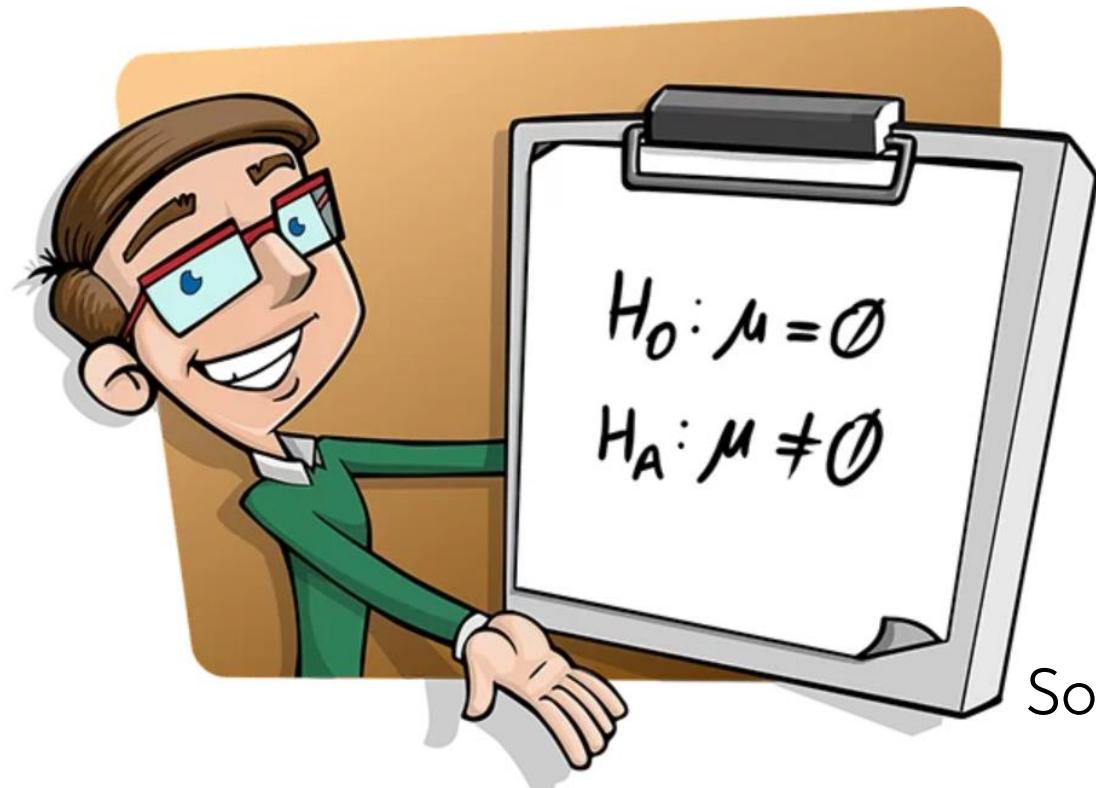


Assay qualification & validation

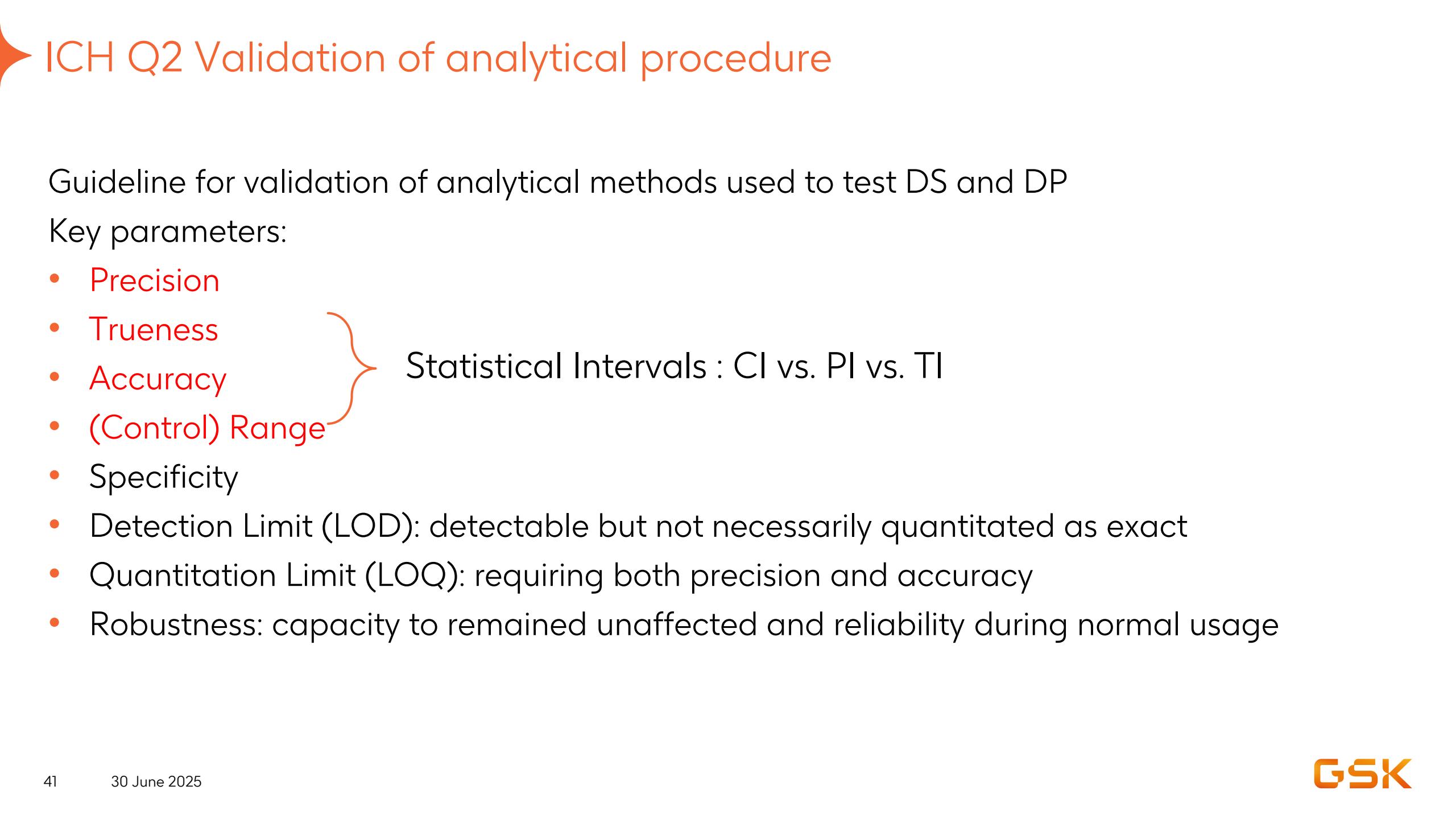


“If I only had one day left to live, I’d spend it in a statistics class - that way the day would seem a lot longer”

-Anonymous



Source: <https://magnimetrics.com/>
GSK

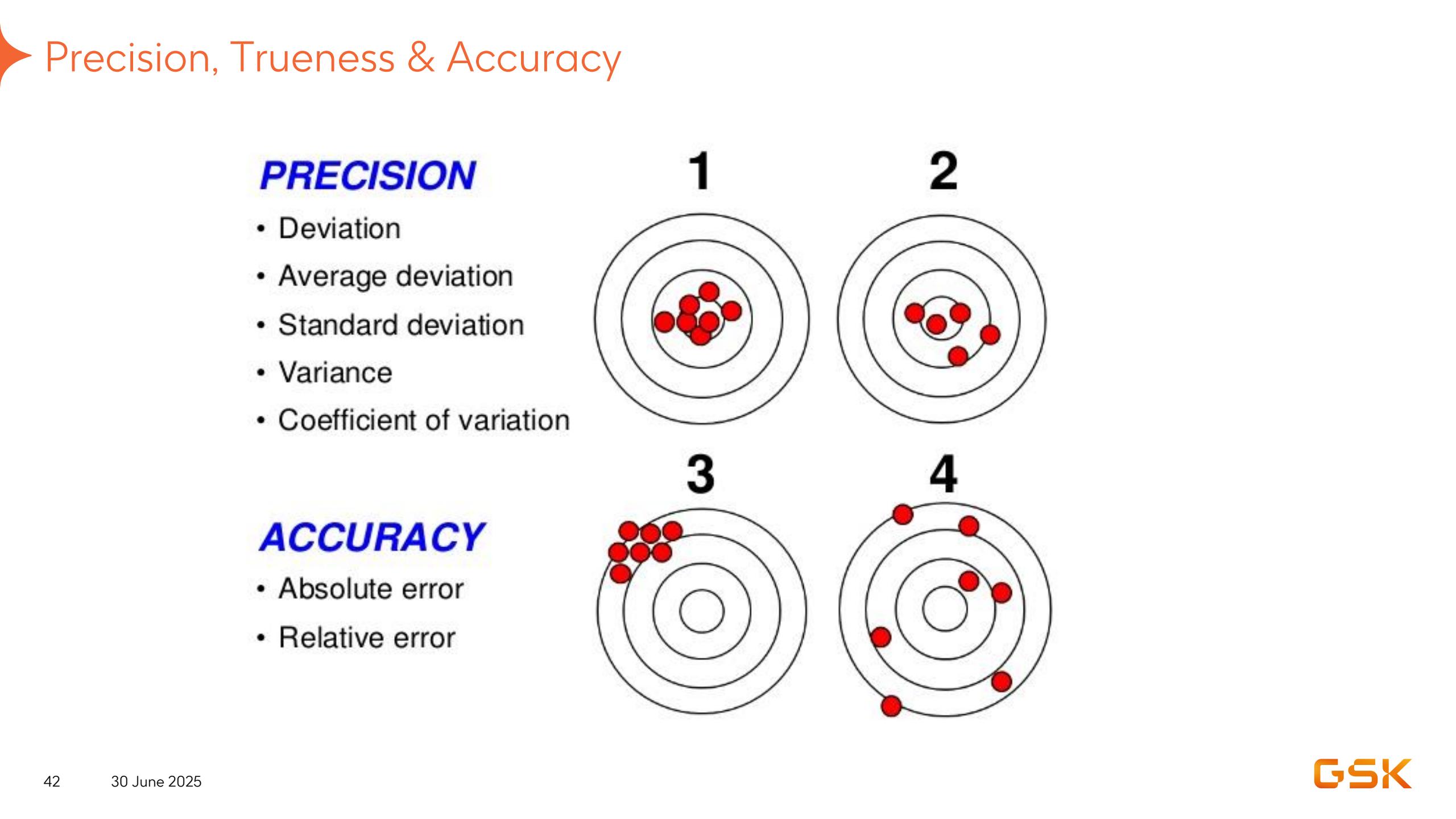


ICH Q2 Validation of analytical procedure

Guideline for validation of analytical methods used to test DS and DP

Key parameters:

- Precision
- Trueness
- Accuracy
- (Control) Range
- Specificity
- Detection Limit (LOD): detectable but not necessarily quantitated as exact
- Quantitation Limit (LOQ): requiring both precision and accuracy
- Robustness: capacity to remained unaffected and reliability during normal usage

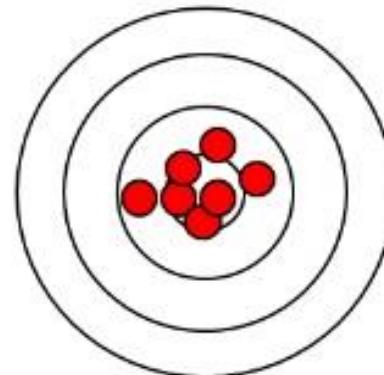


Precision, Trueness & Accuracy

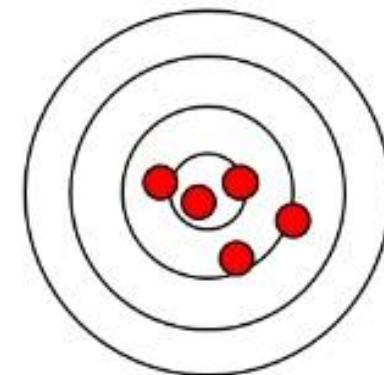
PRECISION

- Deviation
- Average deviation
- Standard deviation
- Variance
- Coefficient of variation

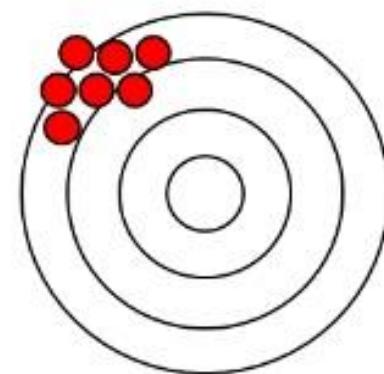
1



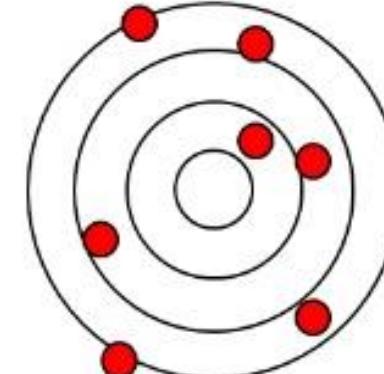
2



3



4



ACCURACY

- Absolute error
- Relative error

Consistency of the result of the test under specified conditions whether test is a status (qualitative) or a metrics (quantitative)

ISO

- Closeness of individual measurements to each other (random error).

FDA

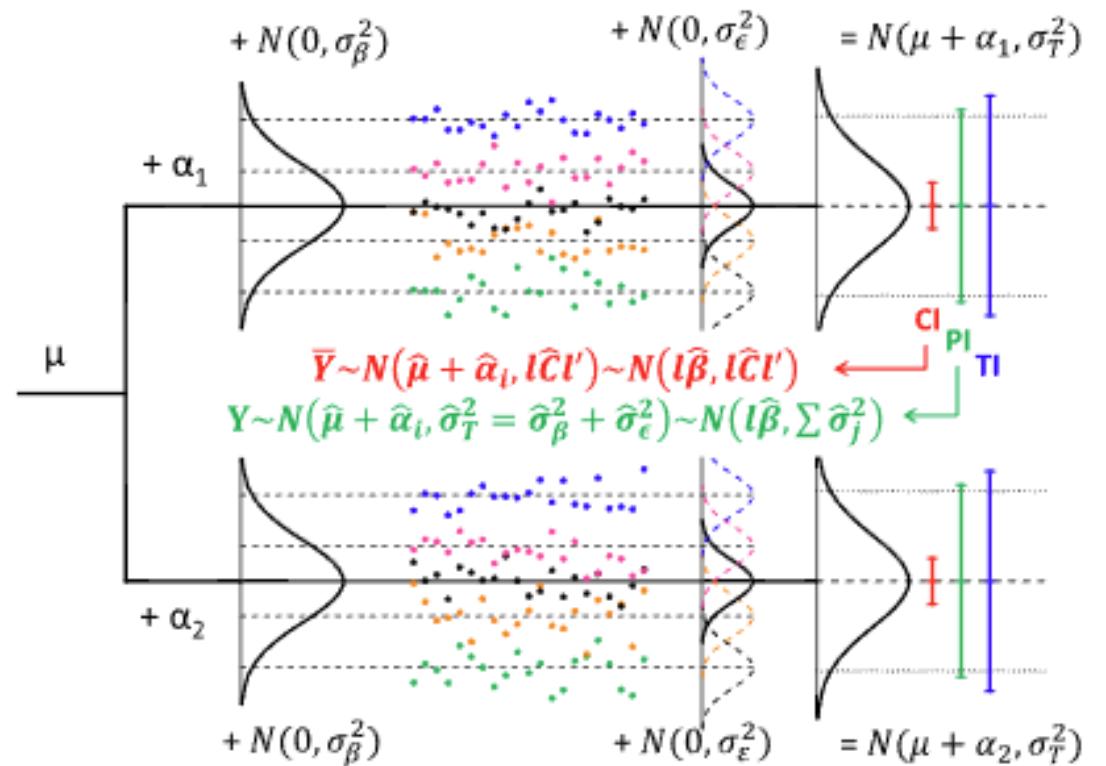
- Closeness of agreement among individual measurements when the procedure is repeatedly applied to multiple aliquots of the same homogeneous sample.

ISO: international organization for standardization

Precision

Variance decomposition using random effects model:

- Given factors in DoE for precision: cross, nested or more complex design
- Precision of response measured is quantified by
 - SD of the random factors involved
 - CV% ($= \text{SD}/\mu$) vs. GCV% ($= \sqrt{\exp(\sigma^2)-1}$) for log-transformed response



Ability to give a true classification or true measure of the substance of interest

ISO

- Closeness of the mean measurement to the true/reference value (systematic error).

FDA

- **Accuracy:** closeness of mean test results to the actual (true) value of the analyte



Accuracy

A broader term encompassing both **trueness + precision**

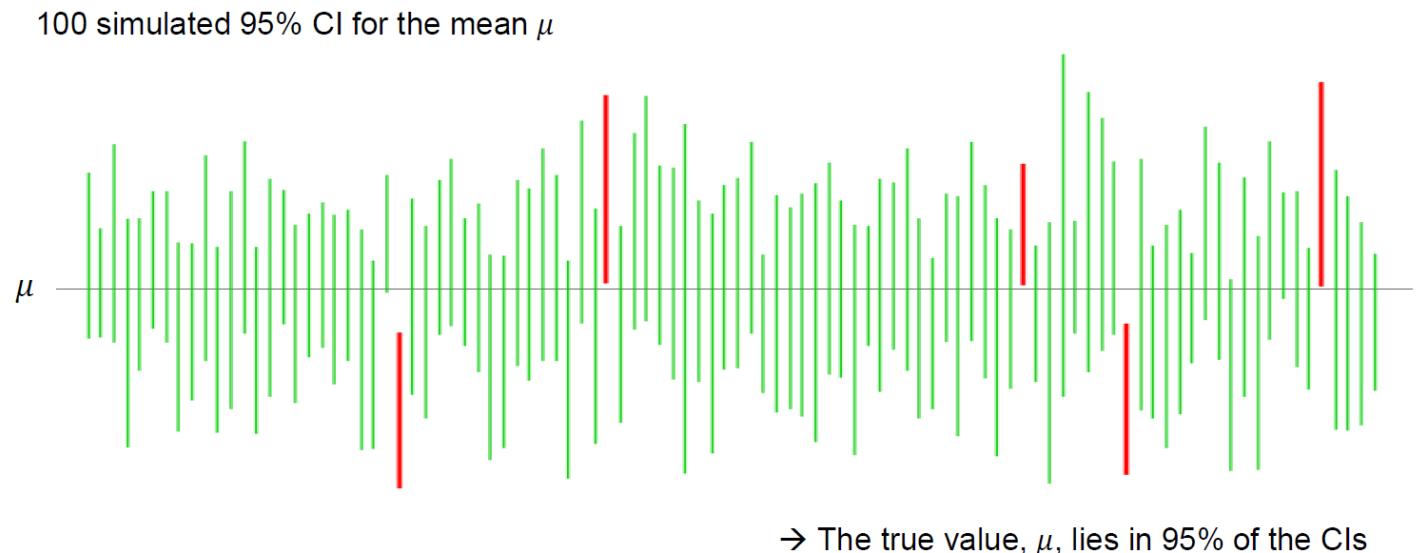
	Type of Error	Quantification Metrics
Trueness	Systematic Error	CI
Precision	Random Error	CV%, SD
Accuracy	Total Error	PI or TI

Confidence Intervals/Limits

Confidence limits are the maximum and minimum values bracketing the statistic of interest (usually the arithmetic or geometric mean) based on the distribution of the data (usually the normal or lognormal distribution) at a certain confidence level (usually 95%)

- ✓ Small # samples = Wide CI
- ✓ Large # samples = Tighter CI

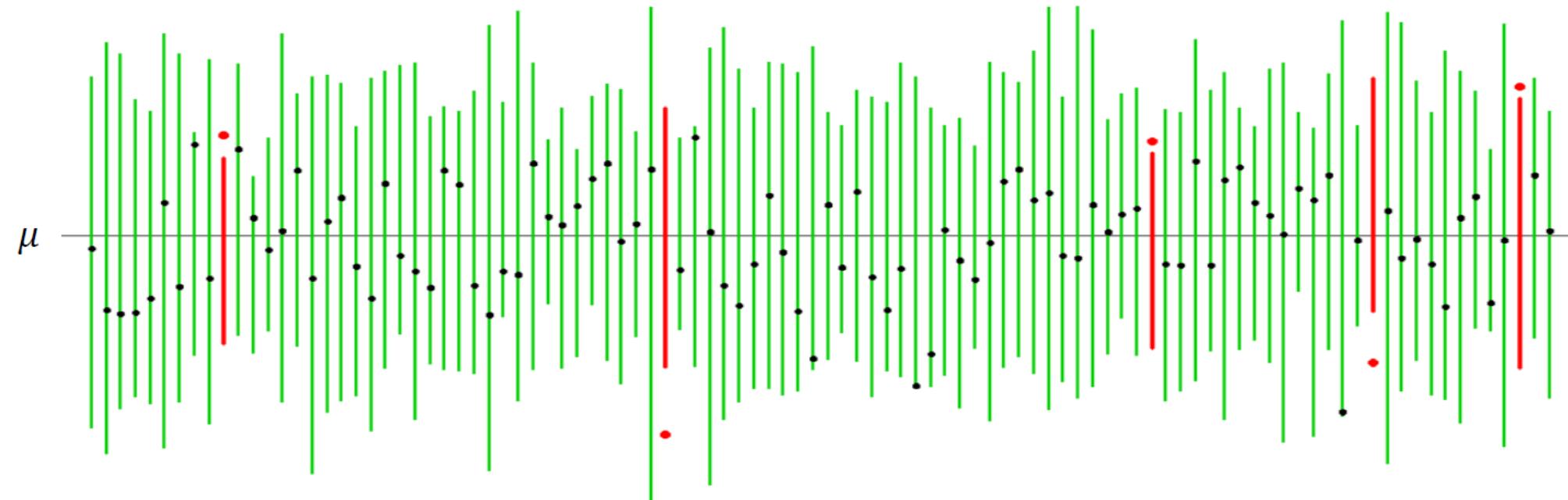
$$\bar{X} \pm t_{1-\alpha/2, n-1} S / \sqrt{n},$$



Note: in Bayesian statistics, credible intervals are commonly used

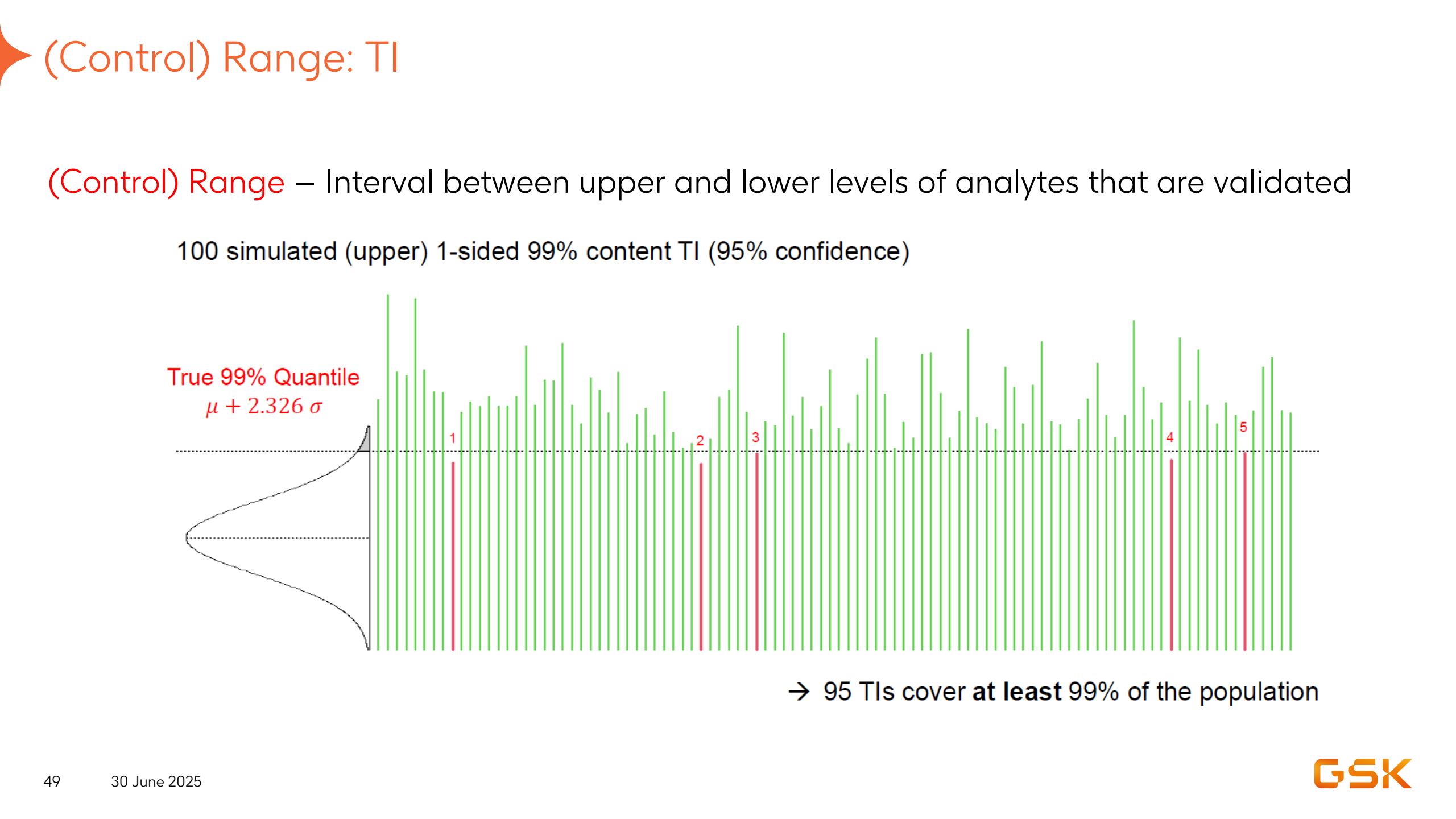
Accuracy: PI

100 simulated 95% PI for a future observation



→ The « future » observation lies into 95% PIs

Note: in Bayesian statistics, PI can be obtained by credible intervals from the posterior distribution





Prediction and Tolerance

Prediction Intervals: β expectation

Prediction limits are the maximum and minimum values covering $\beta\%$ (e.g. 95%) of the population samples on average.

Tolerance Intervals: β expectation - γ content

Tolerance limits are the maximum and minimum values covering $\beta\%$ (e.g. 95%) of the population samples at a certain $1-\gamma\%$ confidence level (e.g. 80%).

95% PI \approx 95% TI with 50% confidence level

CI vs. PI vs. TI in i.i.d. setting

Independent &
identically distributed

CI

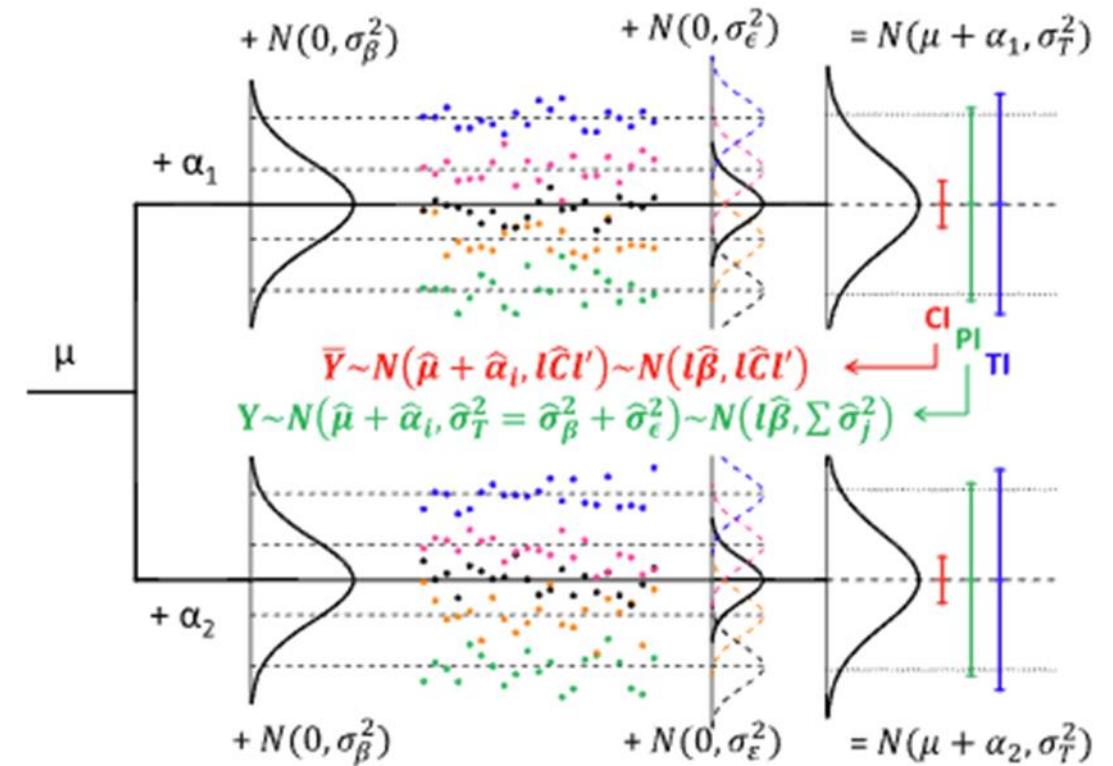
$$\bar{X} \pm t_{1-\alpha/2, n-1} S / \sqrt{n},$$

PI

$$\bar{X} \pm t_{1-\psi/2, n-1} S \sqrt{1 + 1/n}.$$

TI

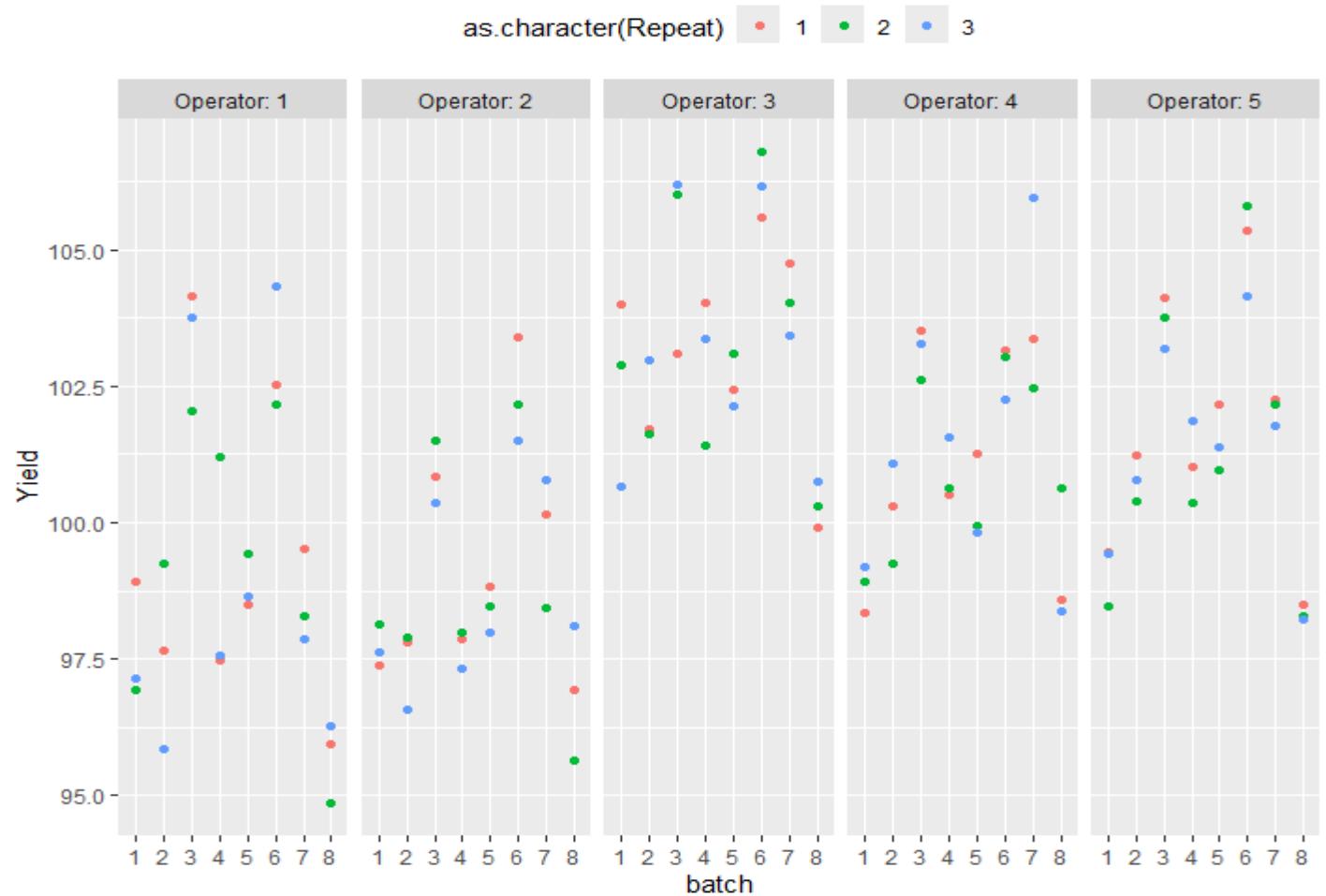
$$\bar{X} \pm z_{1-\psi/2} S \sqrt{1 + 1/n} \sqrt{\frac{n-1}{\chi^2_{\alpha, n-1}}},$$



CI vs. PI vs. TI in i.i.d. setting vs. mixed model		
Independent & identically distributed		Mixed Model
CI	$\bar{X} \pm t_{1-\alpha/2,n-1} S / \sqrt{n},$	$\hat{\mu} \pm t_{1-\alpha/2,k} \sqrt{\text{var}(\hat{\mu})}. \quad df = k$
PI	$\bar{X} \pm t_{1-\psi/2,n-1} S \sqrt{1 + 1/n}.$	$\hat{\mu} \pm t_{1-\psi/2,r} \sqrt{\text{var}(\hat{\mu}) + \hat{\sigma}_T^2}, \quad df = r$
TI	$\bar{X} \pm z_{1-\psi/2} S \sqrt{1 + 1/n} \sqrt{\frac{n-1}{\chi_{\alpha,n-1}^2}},$	$\hat{Y}_{l\beta,n+1} \pm z_{1-\psi/2} \sqrt{l\hat{C}_{11}l' + \hat{\sigma}_T^2} \sqrt{1 + \frac{1}{\hat{\sigma}_T^2} \sqrt{\sum_{j=1}^q H_j^2 k_j^2 \text{EMS}_j^2}},$
		Modification of df in mixed model according to variance decomposition
Francq, Lin, Hoyer 2019		
52	30 June 2025	

Example: protein yield data (simulated)

- Response: protein yield
- Measured by fluorescence-based protein assay
- 8 Batches of vaccine formulation
- 5 Operators
- Each operator measures each batch 3 times
- Mean Yield = 100





Precision

Repeatability: Precision under **the same conditions** over a short time

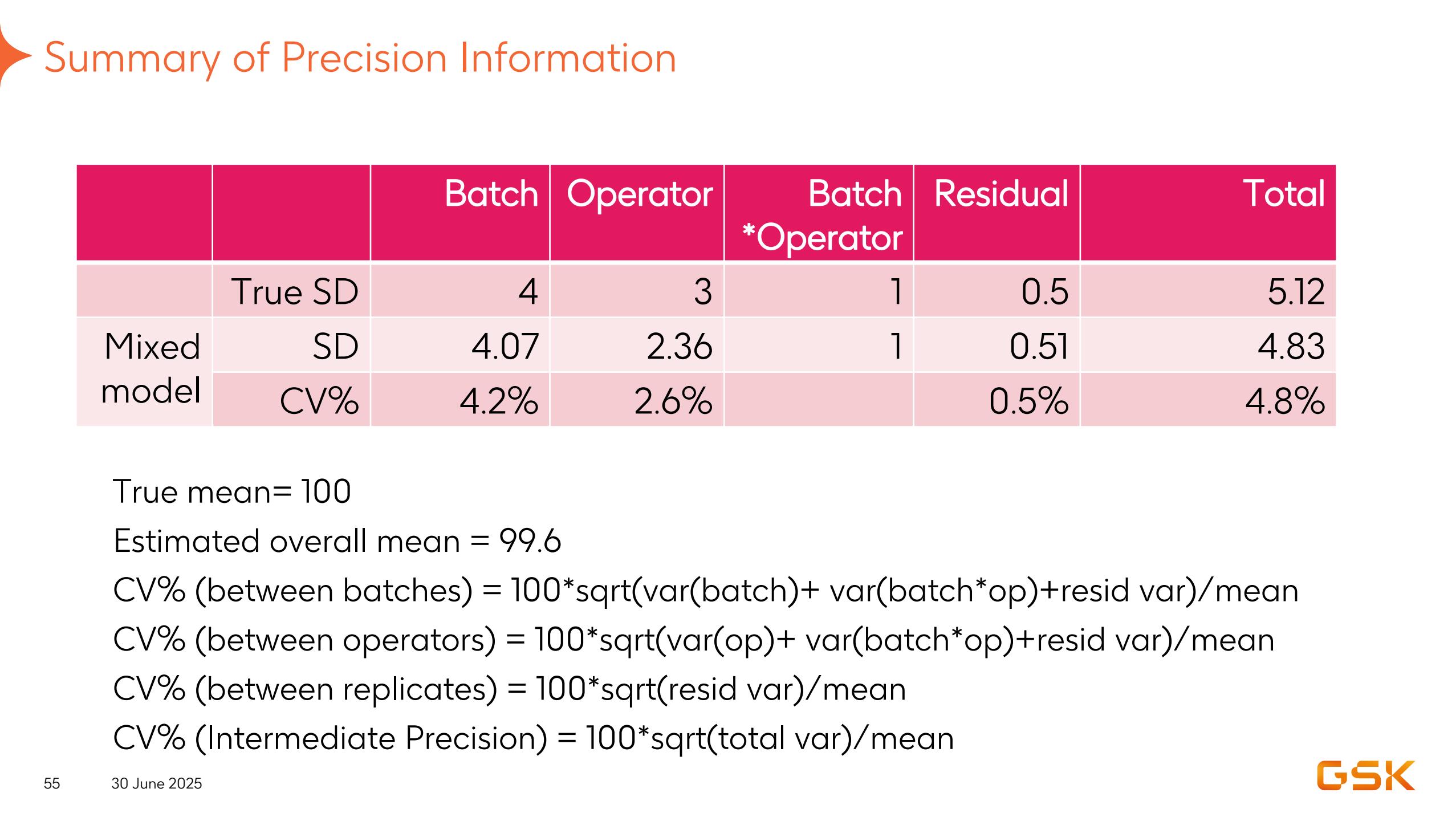
- Same analyst
- Same instrument
- Same lab
- Same day

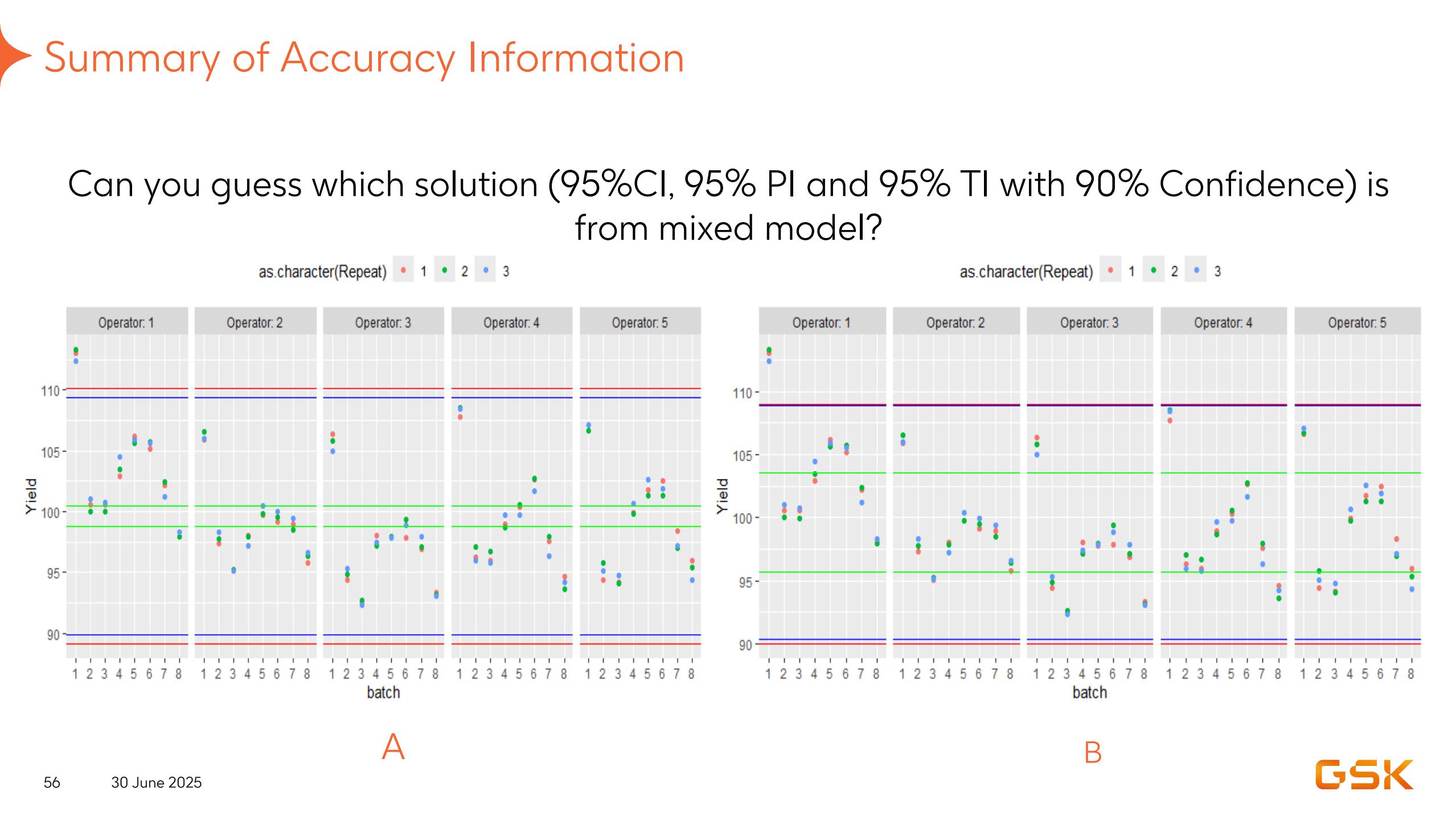
Intermediate Precision (IP): Precision under **different conditions within the same laboratory**

- Different days
- Different analysts
- Different equipments

Reproducibility: Precision under **different laboratories or study sites**

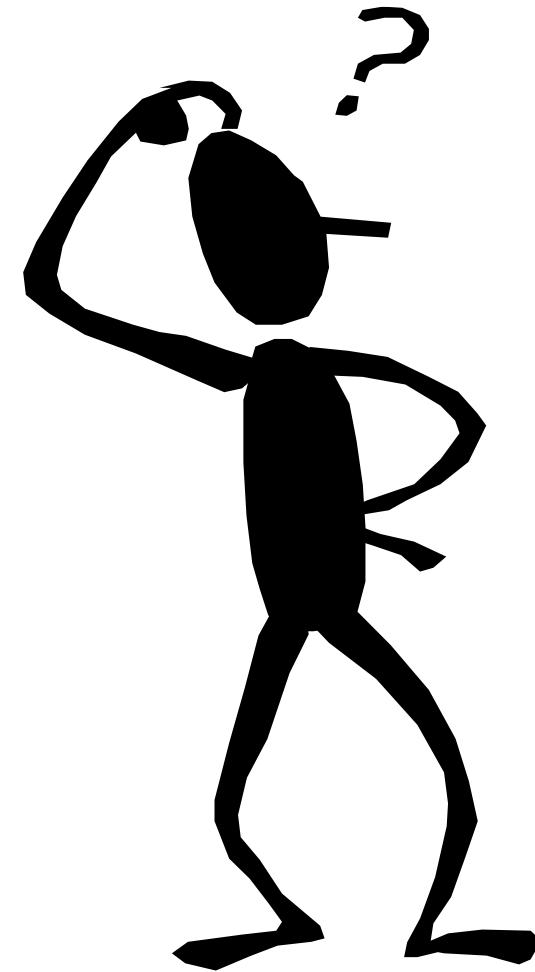
- Different labs
- Different operators
- Possibly different instruments or reagent lots

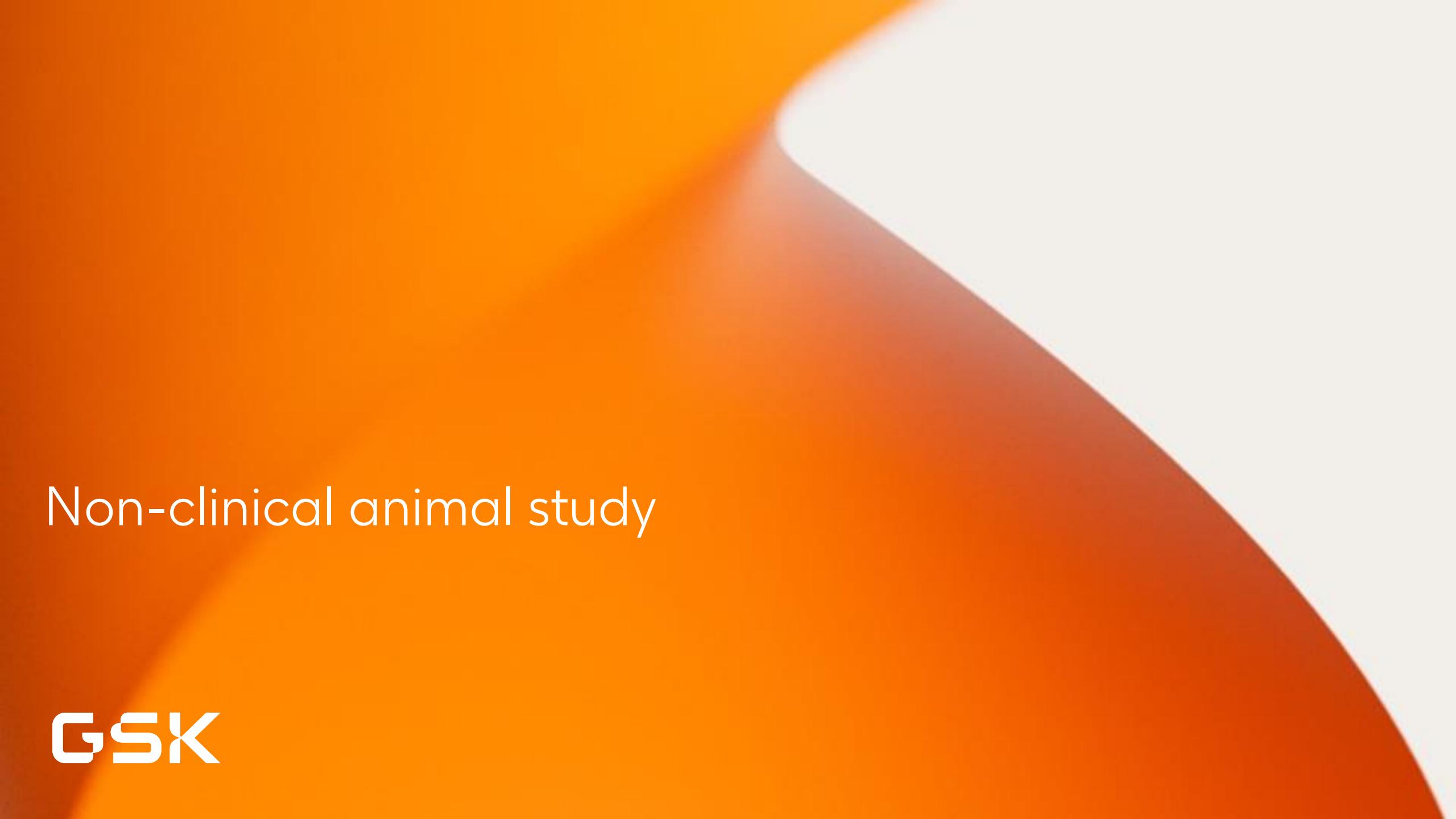




- Francq, Lin, Hoyer. Confidence, Prediction and Tolerance in Linear Mixed Models. *Statistics in Medicine* (2019) ***
- Francq, Lin, Hoyer: Confidence and Prediction in Linear Mixed Models: Do Not Concatenate the Random Effects. Application in an Assay Qualification Study. *Statistics in Biopharmaceutical research* (2020)
- Menssen & Schaarschmidt Prediction intervals for all of M future observations based on linear random effects models, *Statistica Neerlandica* (2022)

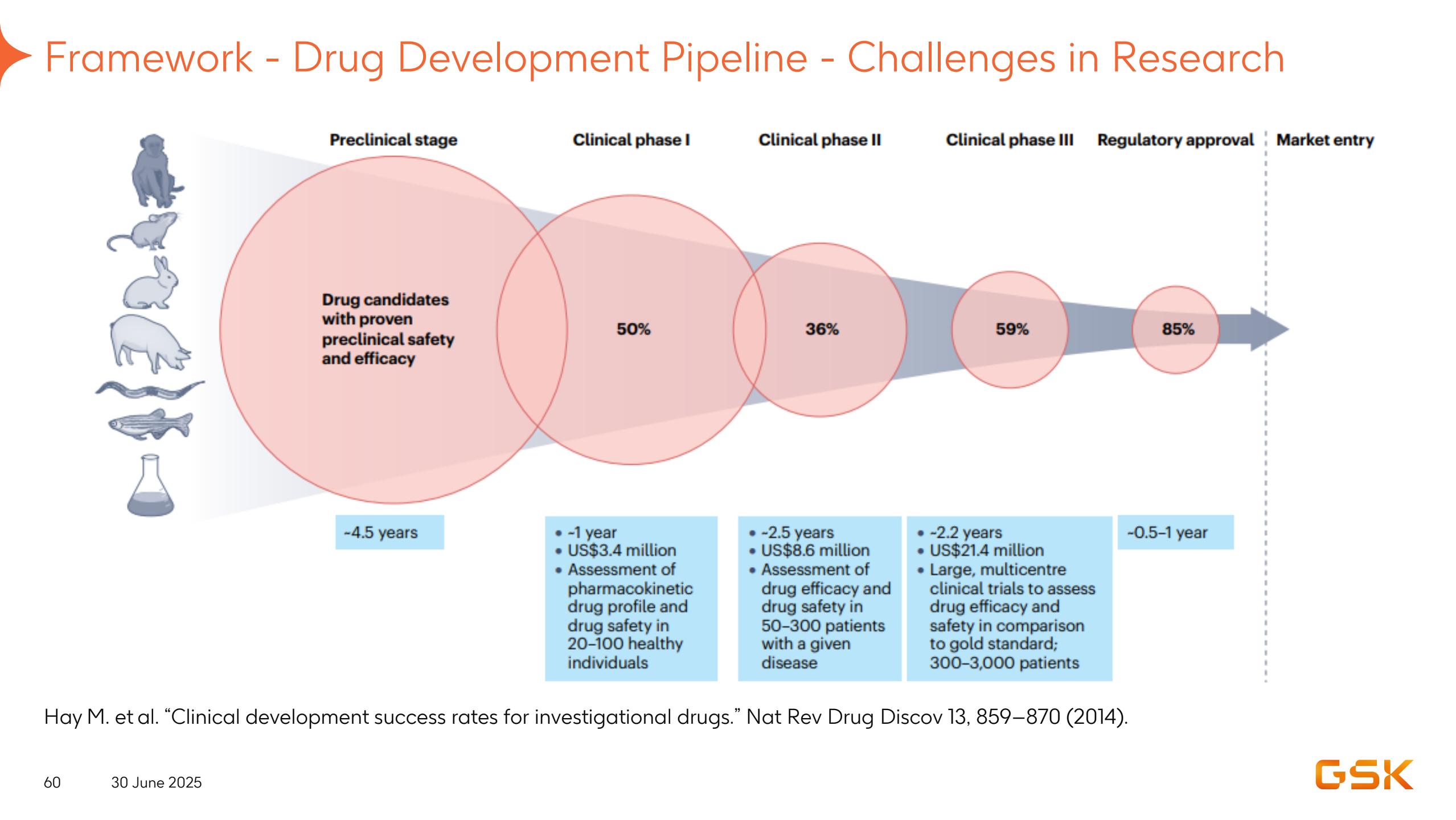
*** Top 10% most downloaded papers published by Wiley



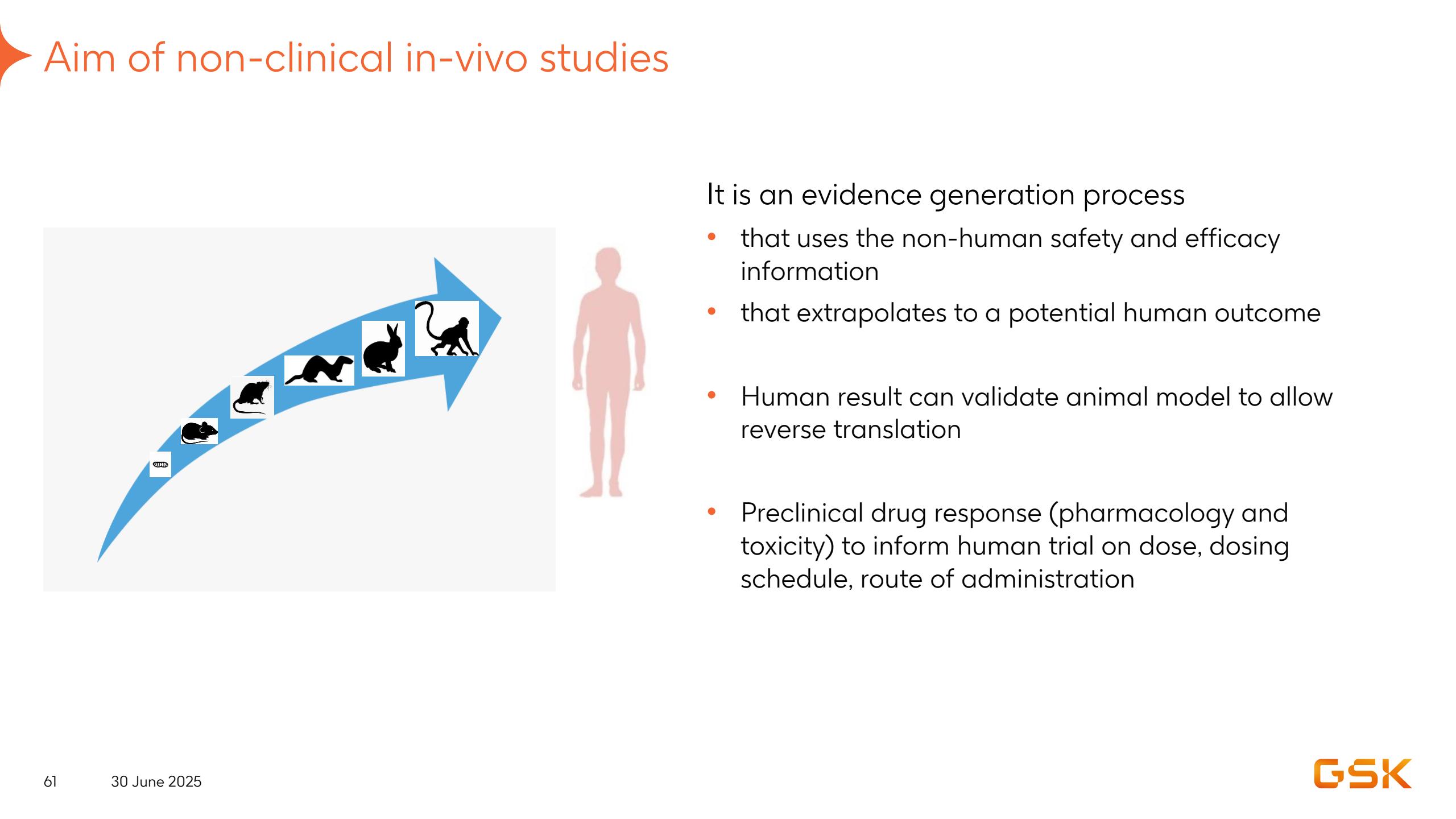


Non-clinical animal study





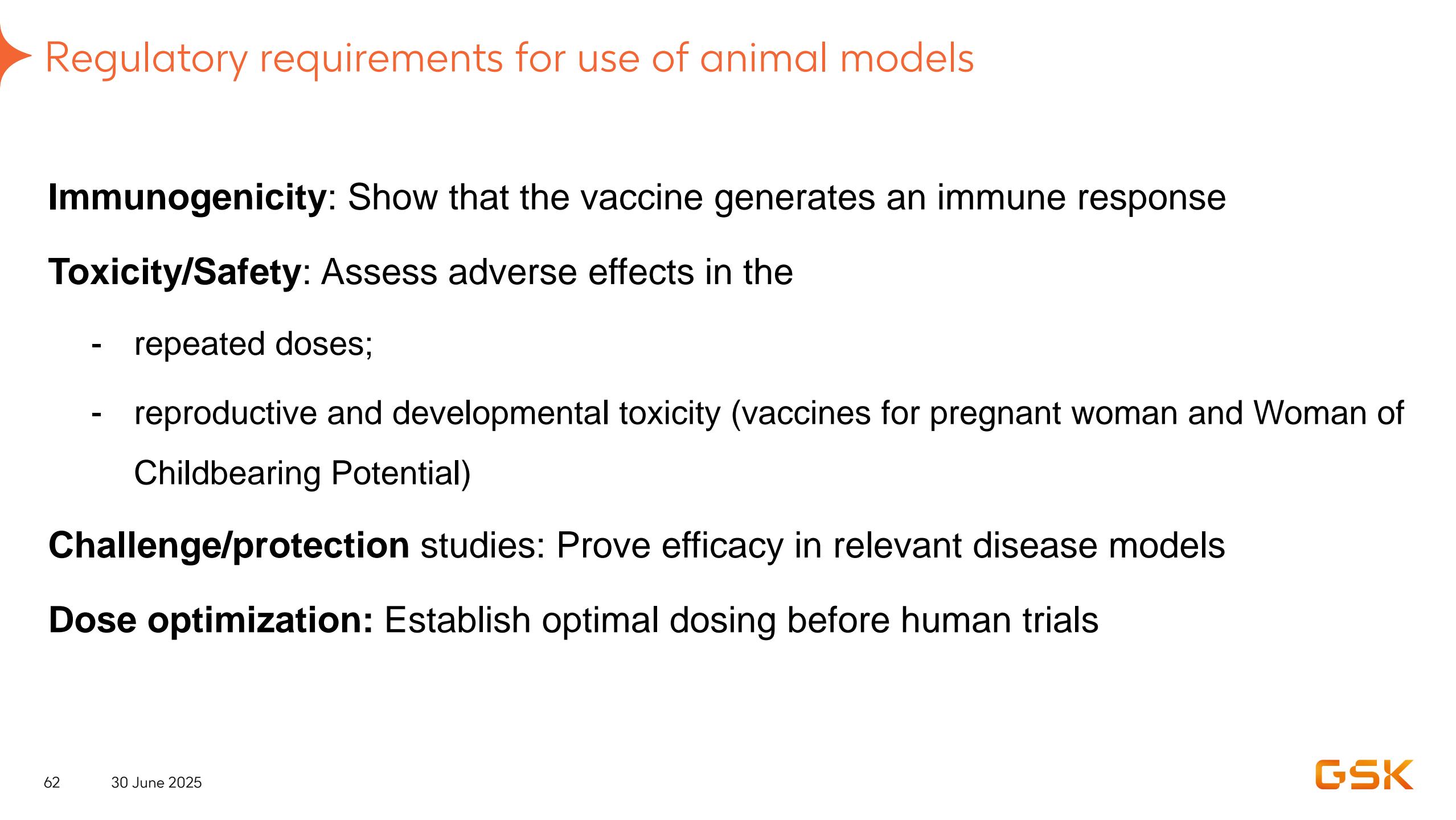
Hay M. et al. "Clinical development success rates for investigational drugs." Nat Rev Drug Discov 13, 859–870 (2014).



Aim of non-clinical in-vivo studies

It is an evidence generation process

- that uses the non-human safety and efficacy information
- that extrapolates to a potential human outcome
- Human result can validate animal model to allow reverse translation
- Preclinical drug response (pharmacology and toxicity) to inform human trial on dose, dosing schedule, route of administration



Regulatory requirements for use of animal models

Immunogenicity: Show that the vaccine generates an immune response

Toxicity/Safety: Assess adverse effects in the

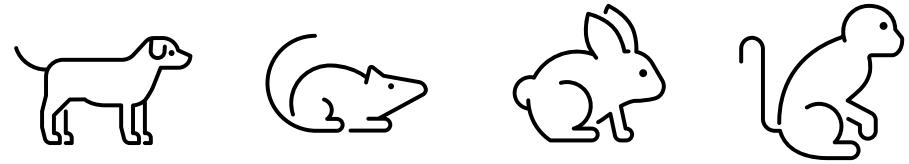
- repeated doses;
- reproductive and developmental toxicity (vaccines for pregnant woman and Woman of Childbearing Potential)

Challenge/protection studies: Prove efficacy in relevant disease models

Dose optimization: Establish optimal dosing before human trials

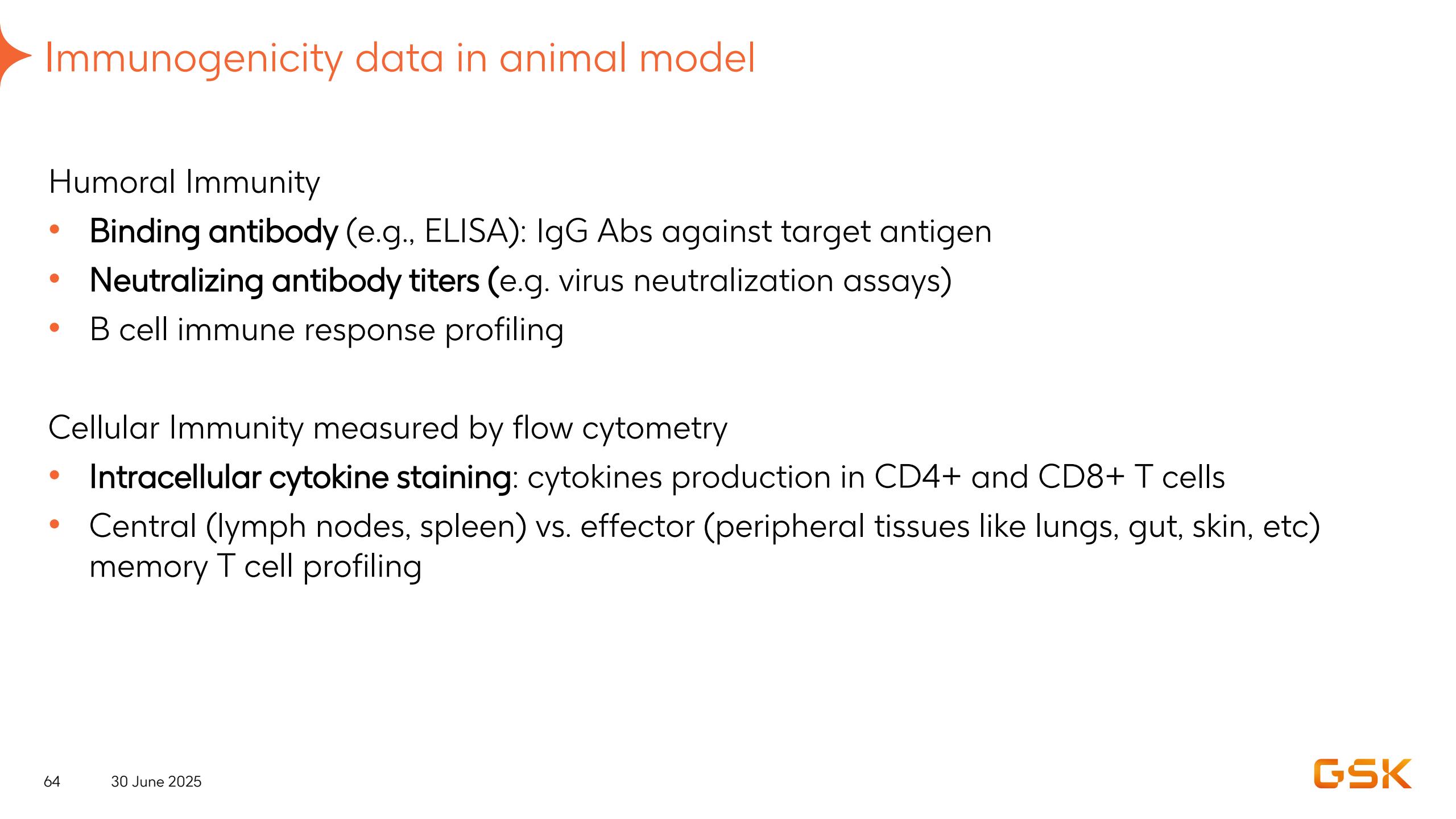
Who/where/what

Animal selection



➤ The selection of a particular animal species should consider:

- the vaccine antigen (capability to mount an immune response),
- the adjuvant (e.g., TLR),
- the desired immune response (e.g., availability of immuno-tools; humoral vs cellular expected immune response),
- the route of administration (e.g., administration of full human dose),
- availability of historical control data,
- 3Rs (Replacement-Reduction-Refinement) principles must be followed,
- In addition to the criteria mentioned above, studies should also be performed in a species permissive to replication of the pathogen. Species selection is determined on a case-by-case basis.



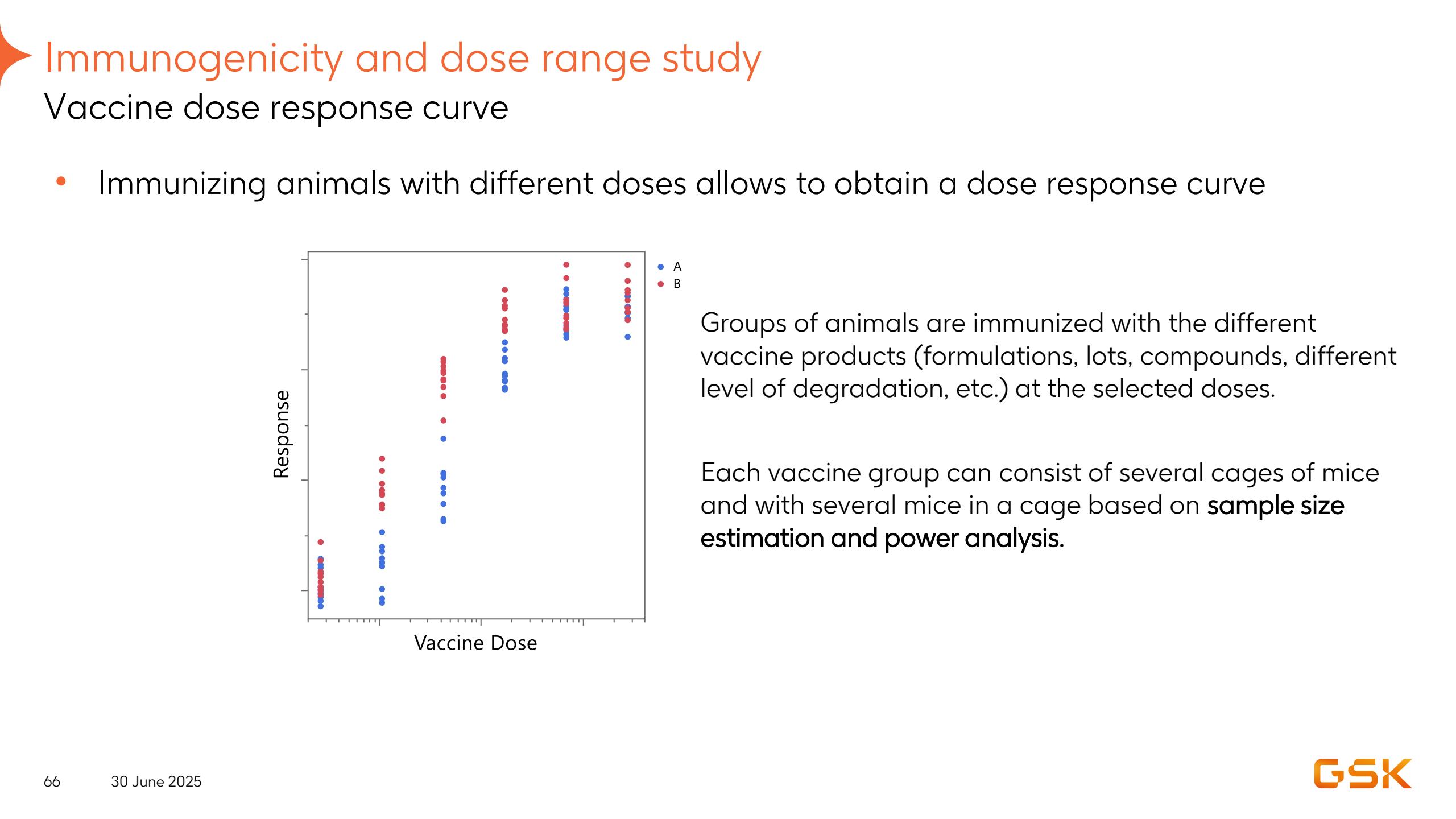
Immunogenicity data in animal model

Humoral Immunity

- **Binding antibody** (e.g., ELISA): IgG Abs against target antigen
- **Neutralizing antibody titers** (e.g. virus neutralization assays)
- B cell immune response profiling

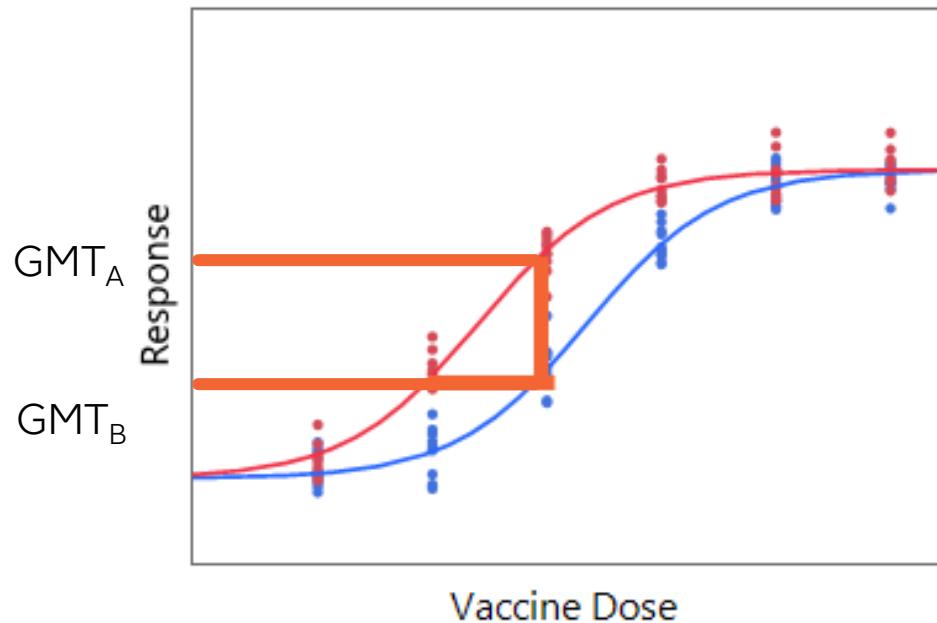
Cellular Immunity measured by flow cytometry

- **Intracellular cytokine staining**: cytokines production in CD4+ and CD8+ T cells
- Central (lymph nodes, spleen) vs. effector (peripheral tissues like lungs, gut, skin, etc) memory T cell profiling



Comparison of immune response between groups

GMR: geometric mean ratio



For dose i
 \log_{10} transformed immune response :
$$GMR_i = \text{GMT}_{Bi} / \text{GMT}_{Ai}$$

GMR is the fold change between mean responses of treatment A and B

Bayesian approach in estimation of GMR

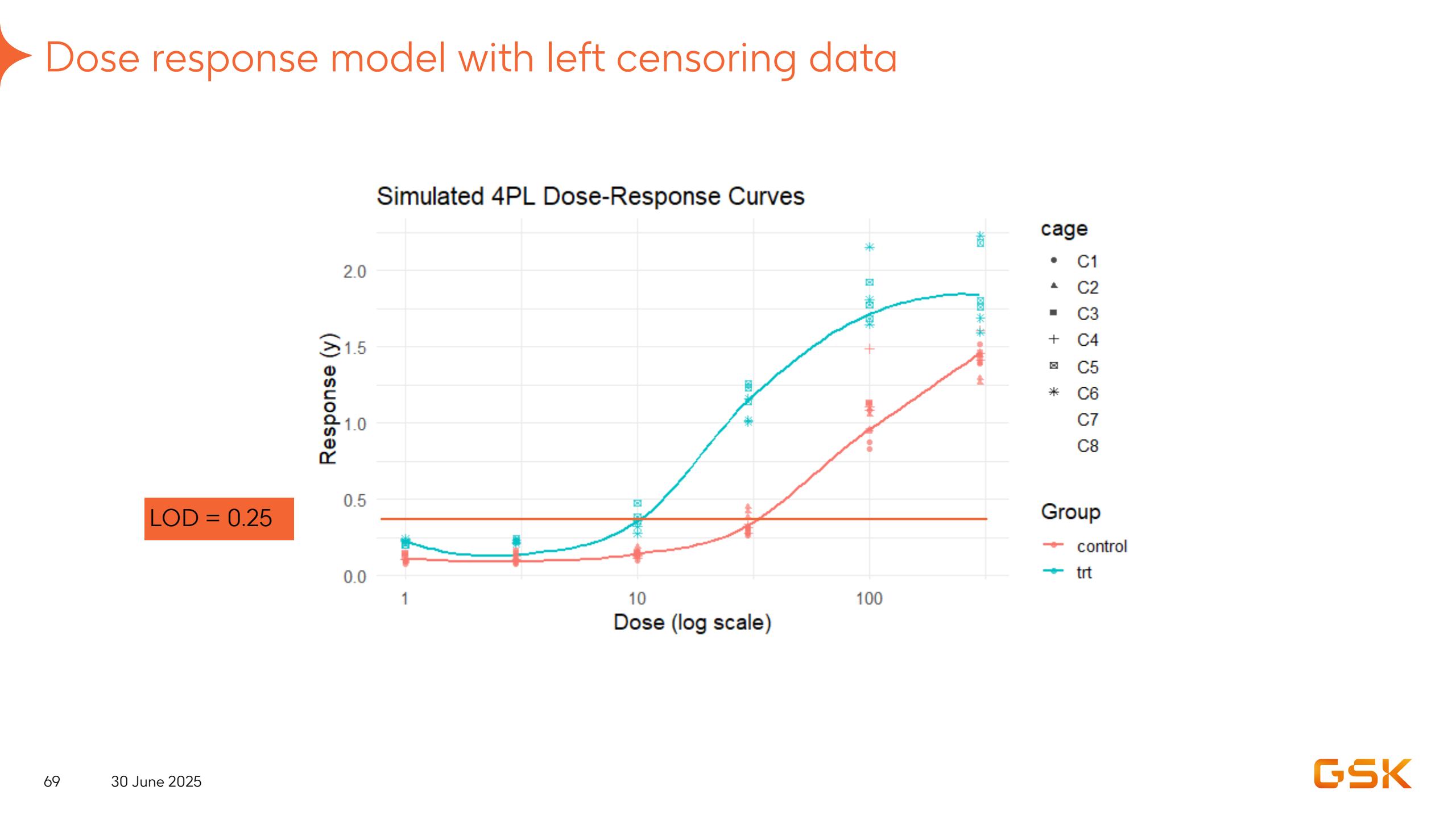
For observation i in group $g[i] \in \{\text{control}, \text{trt}\}$ and cage $c[i]$:

$$y_i \sim \mathcal{N}(\mu_i, \sigma^2) \quad , \quad \mu_i = A_{g[i],c[i]} + \frac{B_{g[i],c[i]} - A_{g[i],c[i]}}{1 + \left(\frac{x_i}{C_{g[i],c[i]}}\right)^D}$$

with each 4PL parameter decomposed as

$$\theta_{g,c} = \theta^{(g)} + b_{\theta,c} \quad , \quad b_{\theta,c} \sim N(0, \tau_\theta^2) \quad \text{for } \theta \in \{A, B, C, D\}.$$

- Dose-response modeling of trt and control for log transformed response using brms
- Obtain the posterior distribution of mean difference between trt A and B



Efficacy data in animal model

Examples

Protection from infection

- **Reduction in pathogen load:** measured in blood, lungs, nasal washes, tissues
- No detectable replication of the challenge virus/bacteria via CFU, qPCR, etc

Protection from Disease

- **Clinical signs:** Weight loss, fever, respiratory rate, mobility (RSV)
- **Efficacy/Survival rates:** staphylococcus infection model in mice
- **Histopathology:** Scoring of tissue inflammation/damage (HSV1/2)

Reactogenicity data in animal model

Examples

Local reactogenicity (injection site):

- **Swelling, redness (erythema), and induration**
- **Tissue damage or immune cell infiltration**

Systemic reactogenicity:

- Innate **cytokines** profiling (e.g., IL-6, IFN- γ , TNF- α in serum) measured within 1 day post vaccination
- **Weight loss, body temperature** change
- **Behavior** changes: less movement, reduced feeding, etc



Why are preclinical study results difficult to replicate?

- Inadequate experimental design: lack of randomization, incorrect Exp. Unit in analysis

Principle of study design

Experimental Unit

The **smallest entity** receiving a **single treatment** such that each entity has an equal and independent chance of receiving each of the treatments.

Experimental Design (according to objectives):

Co-mingling

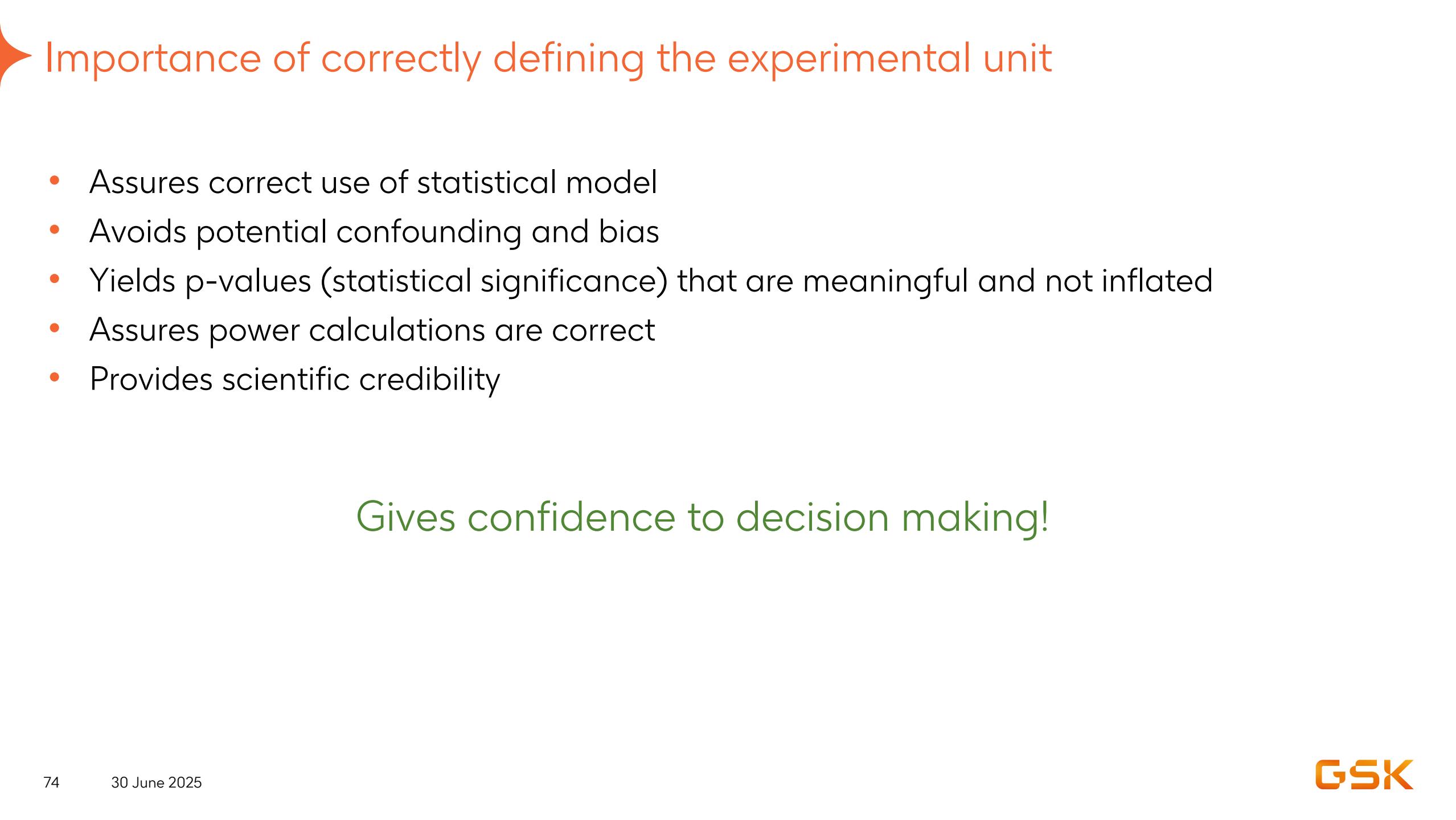


Rats randomized to treatments in a cage

Cage based



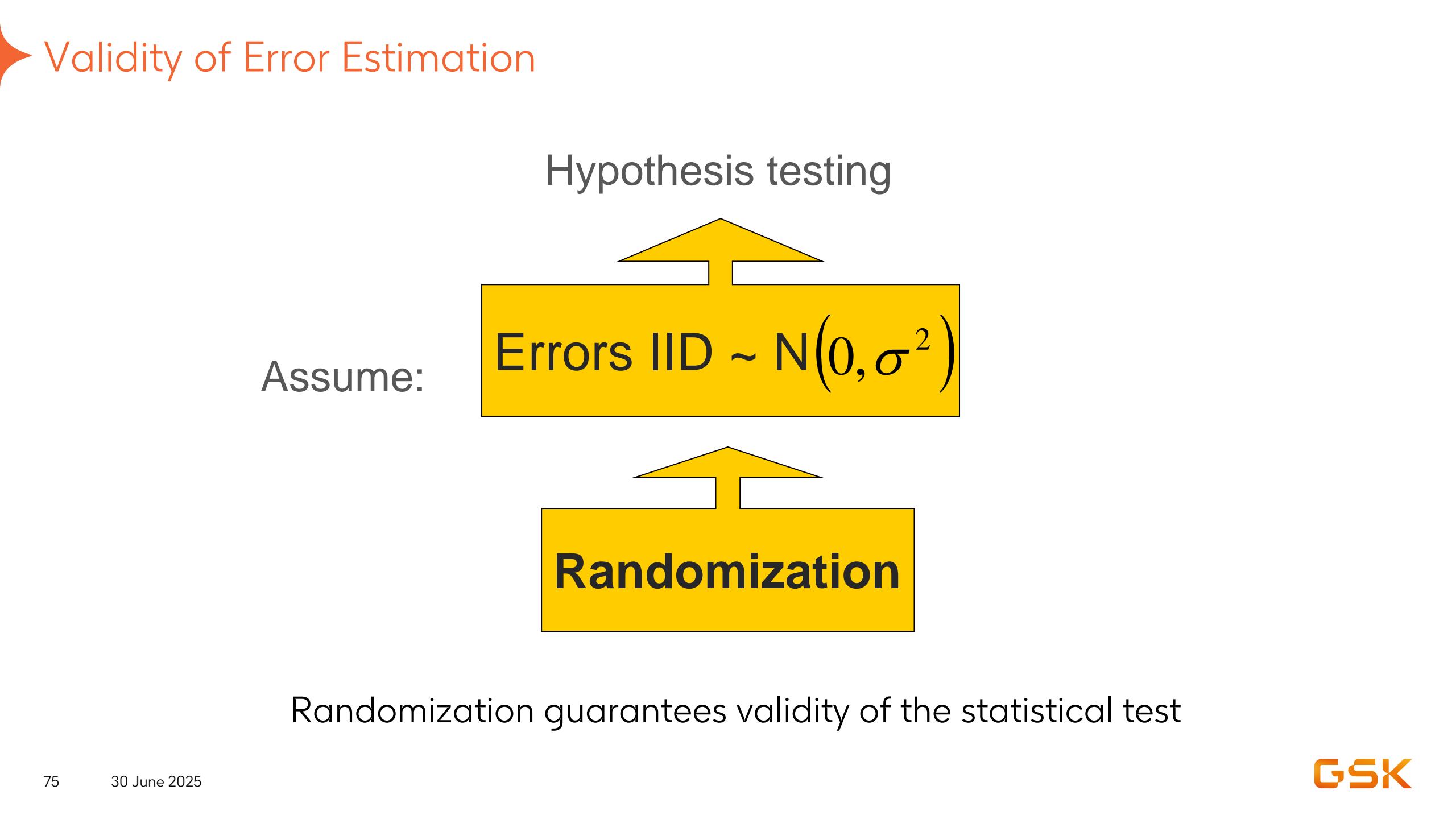
One treatment group in a cage



Importance of correctly defining the experimental unit

- Assures correct use of statistical model
- Avoids potential confounding and bias
- Yields p-values (statistical significance) that are meaningful and not inflated
- Assures power calculations are correct
- Provides scientific credibility

Gives confidence to decision making!





Why are preclinical study results difficult to replicate?

- Inadequate experimental design: lack of randomization, incorrect Exp. Unit in analysis
- Differences in **animal strains**, microbiota, age, sex, housing conditions, or immune status can alter outcomes

Allergic rhinitis to House Dust Mite - Mouse Strain

Response BL/6 vs Balb/c



Mouse Strains

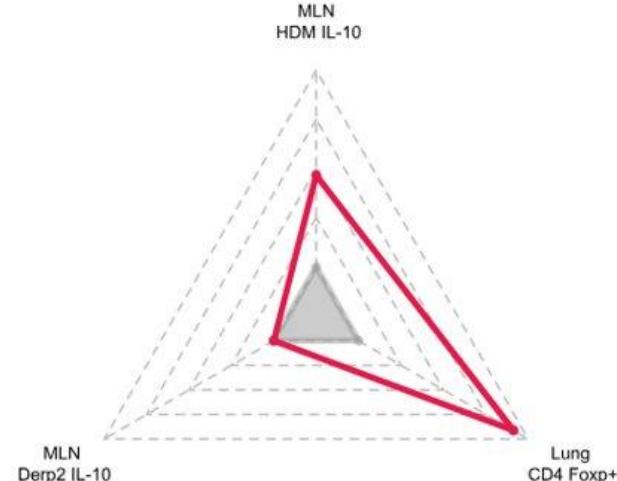
- C57BL/6



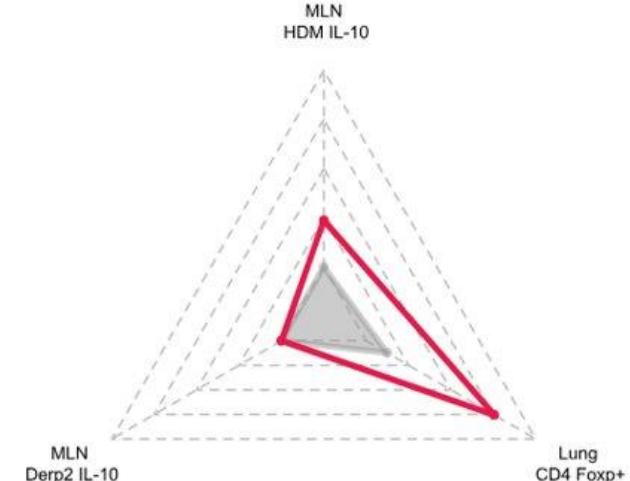
- Balb/c



C57BL/6 - T regulatory compartment



Balb/c - T regulatory compartment



- **A,G**: PBS / PBS / Saline
- **C,I**: HDM / HDM / Saline



Why are preclinical study results difficult to replicate?

- Inadequate experimental design: lack of randomization, incorrect Exp. Unit in analysis
- Differences in animal strains, microbiota, age, sex, housing conditions, or immune status can alter outcomes
- Low **standardization between labs:**
 - protocols often differ: different lab instruments can result in different ways of working
 - dosing schedule, route of administration: e.g., injection of 2 doses at D0 & D14 -> D21
 - assay timing, etc.: e.g., d7PII, d14PII, d21PII, d28PII



Why are preclinical study results difficult to replicate?

- Inadequate experimental design: lack of randomization, incorrect Exp. Unit in analysis
- Differences in **animal strains**, microbiota, age, sex, housing conditions, or immune status can alter outcomes
- Low **standardization between labs**:
 - protocols often differ
 - dosing schedule, route of administration
 - assay timing, etc.
- Reagent, assay (e.g., cell line) drift

Example of reagents in an ELISA test for measuring antibody titer

Reagent	Purpose
Coating Antigen	The target protein (e.g., viral spike, bacterial toxin) coated onto the plate to capture antibodies from the sample
Blocking Buffer	Prevents nonspecific binding to the plate (e.g., BSA, casein, skim milk)
Sample/Serum	Contains the antibodies to be measured (e.g., diluted human or animal serum)
Secondary Antibody	Enzyme-linked antibody (e.g., anti-human IgG-HRP) that binds to the sample antibodies
Wash Buffer	Removes unbound materials between steps (commonly PBS or TBS + Tween 20)
Substrate (e.g., TMB)	Reacts with enzyme (e.g., HRP) to produce a color change measurable by absorbance
Stop Solution	Stops the enzymatic reaction (e.g., sulfuric acid or phosphoric acid)
Standard/Calibrator (if quant.)	A reference antibody or known titer to quantify unknowns (optional in relative titration)
Dilution Buffer	Used to dilute samples and antibodies (e.g., PBS with BSA)



Why are preclinical study results difficult to replicate?

- Inadequate experimental design: lack of randomization, incorrect Exp. Unit in analysis
- Differences in **animal strains**, microbiota, age, sex, housing conditions, or immune status can alter outcomes
- Low **standardization between labs**:
 - protocols often differ
 - dosing schedule, route of administration
 - assay timing, etc.
- Reagent, assay (e.g., cell line) drift
- Negative or null results often go unpublished (**publication bias**)

Translational statistics

GSK

Ambition - Advancing translational research is critical to the future of R&D

Classical translation

Translational work takes place mostly at preclinical and clinical stage

Near term future

Integration of translational work more broadly
- better decision making around target programs and molecules to progress

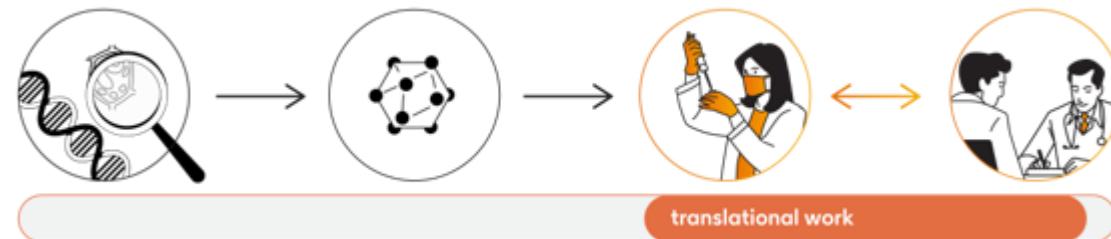
Long term :

Deep phenotyping

Translational underpins entire

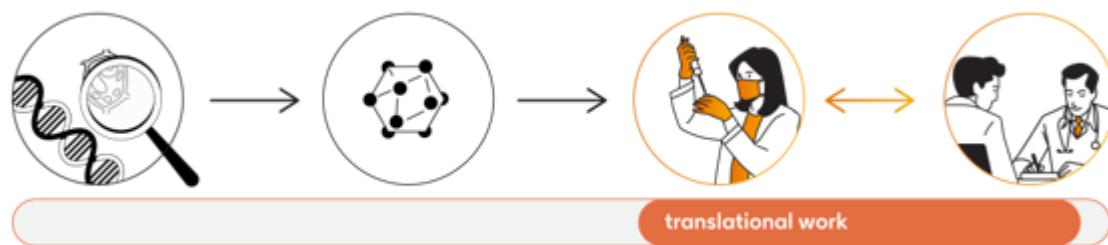
R&D process in an end-to-end manner

Target ID & Validation Molecule Design Pre-Clinical Clinical



Ambition - Advancing translational research is critical to the future of R&D

Target ID & Validation Molecule Design Pre-Clinical Clinical



Target - greater translatability enhances biological understanding & target selection



Molecule design - screening large number of compounds on translatable models



Clinical - systematic and scaled integration of patient derived models and data across R&D

Surrogacy and CoP in Vaccine development

Surrogate Endpoint: humoral/functional immune response

True Endpoint: vaccine efficacy (VE)

CoP: a surrogate of protection is a correlate of risk (as an immune response) that predicts accurately the level of VE

- A CoP is mechanistic if it is mechanistically and **causally responsible** for protection
- Non-mechanistic CoPs

To prove surrogate endpoint as mechanistic CoP is not easy

- Association \neq Causation
- Complexity of Immune Protection
- Trial Design Constraints
- Pathogen and Host Heterogeneity
- Changing Correlates Over Time

GMR	VE (Red Dots)	VE (Black Dots)
25	0.87	-
31	0.83	-
35	0.88	-
36	0.95	-
39	-	0.95
40	-	0.97
43	-	0.96

85

30 June 2025

Systems vaccinology

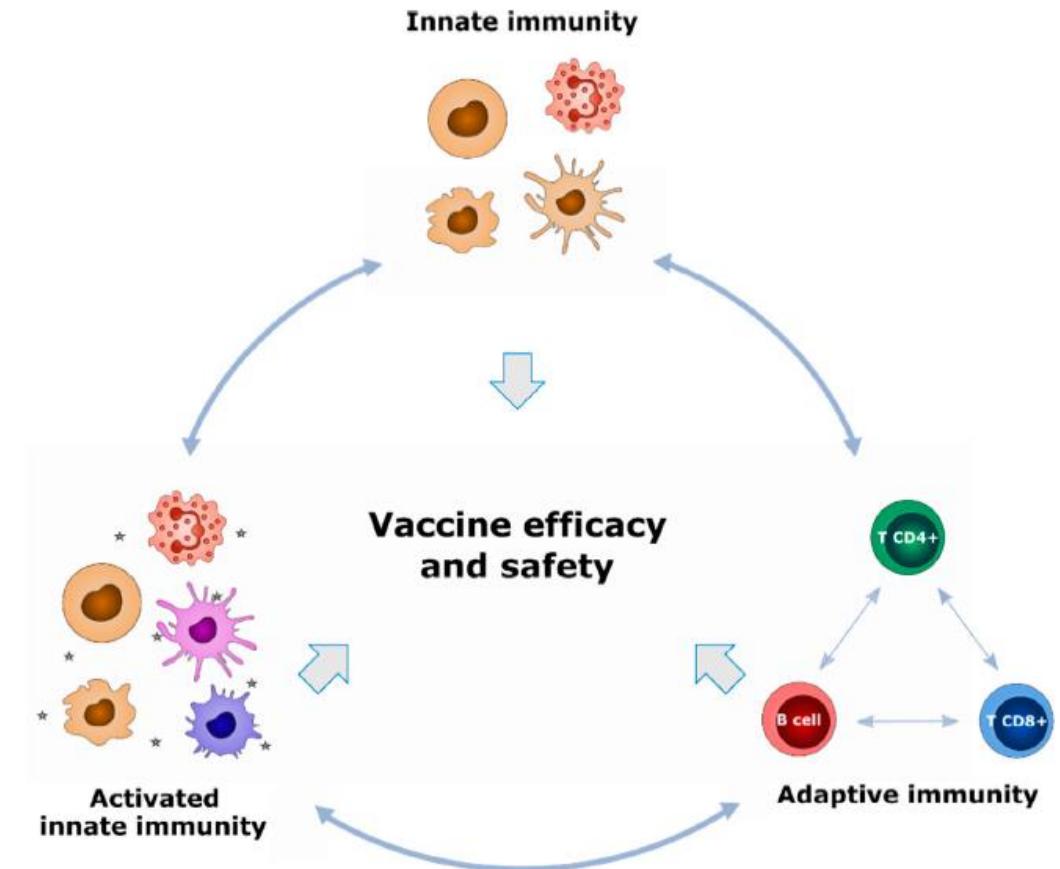
Based on preclinical model and human cohort

Classical predictive biomarkers:

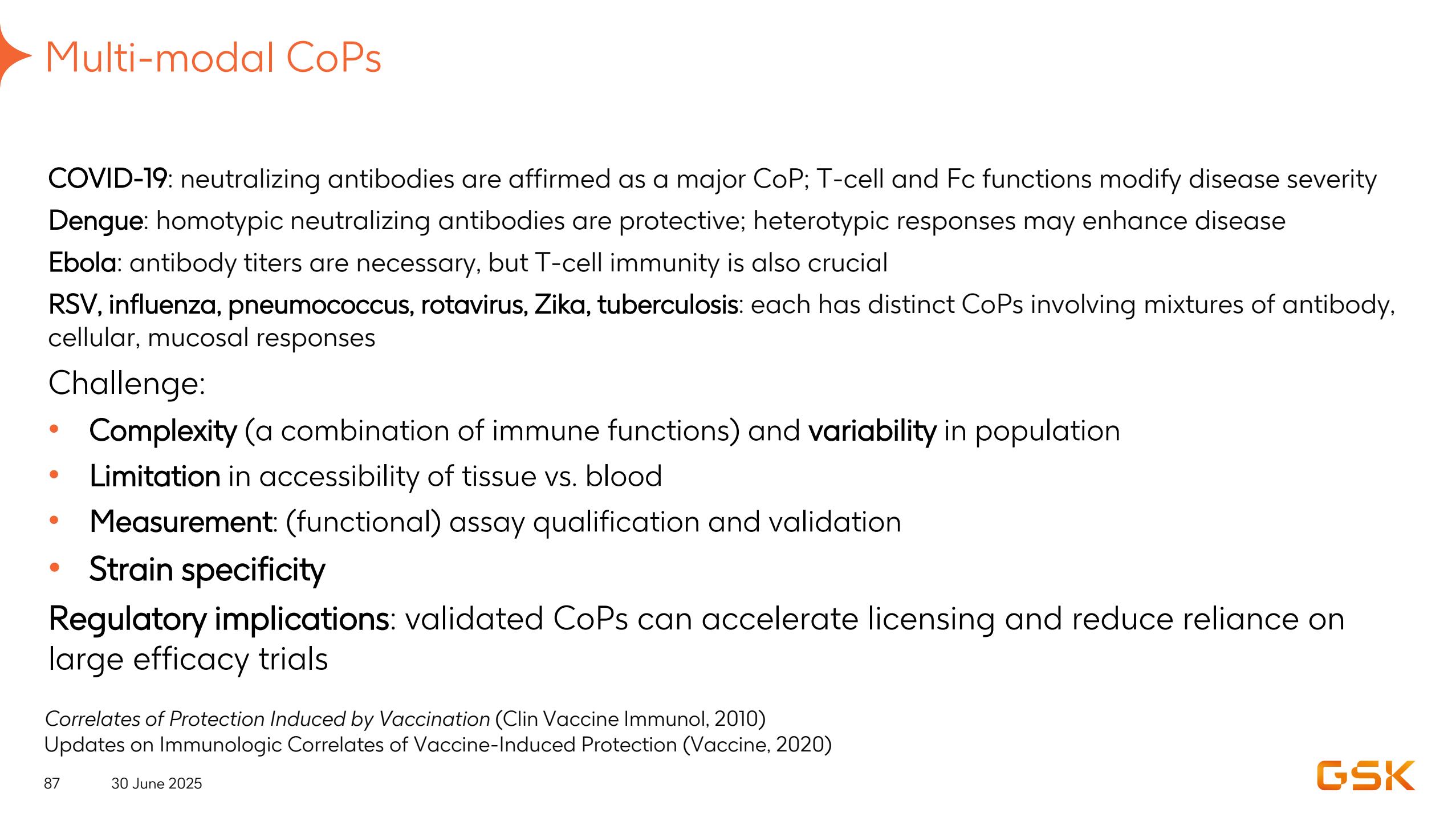
- Antibody responses:
 - Humoral & Functional
 - Cellular immunity: T/B-cell responses
 - Immune memory: effector and central

Influence from host and vaccine factors

- age, sex
- genetics
- microbiome
- vaccine type



Tilbeurgh et al., Predictive Markers of Immunogenicity and Efficacy for Human Vaccines



Multi-modal CoPs

COVID-19: neutralizing antibodies are affirmed as a major CoP; T-cell and Fc functions modify disease severity

Dengue: homotypic neutralizing antibodies are protective; heterotypic responses may enhance disease

Ebola: antibody titers are necessary, but T-cell immunity is also crucial

RSV, influenza, pneumococcus, rotavirus, Zika, tuberculosis: each has distinct CoPs involving mixtures of antibody, cellular, mucosal responses

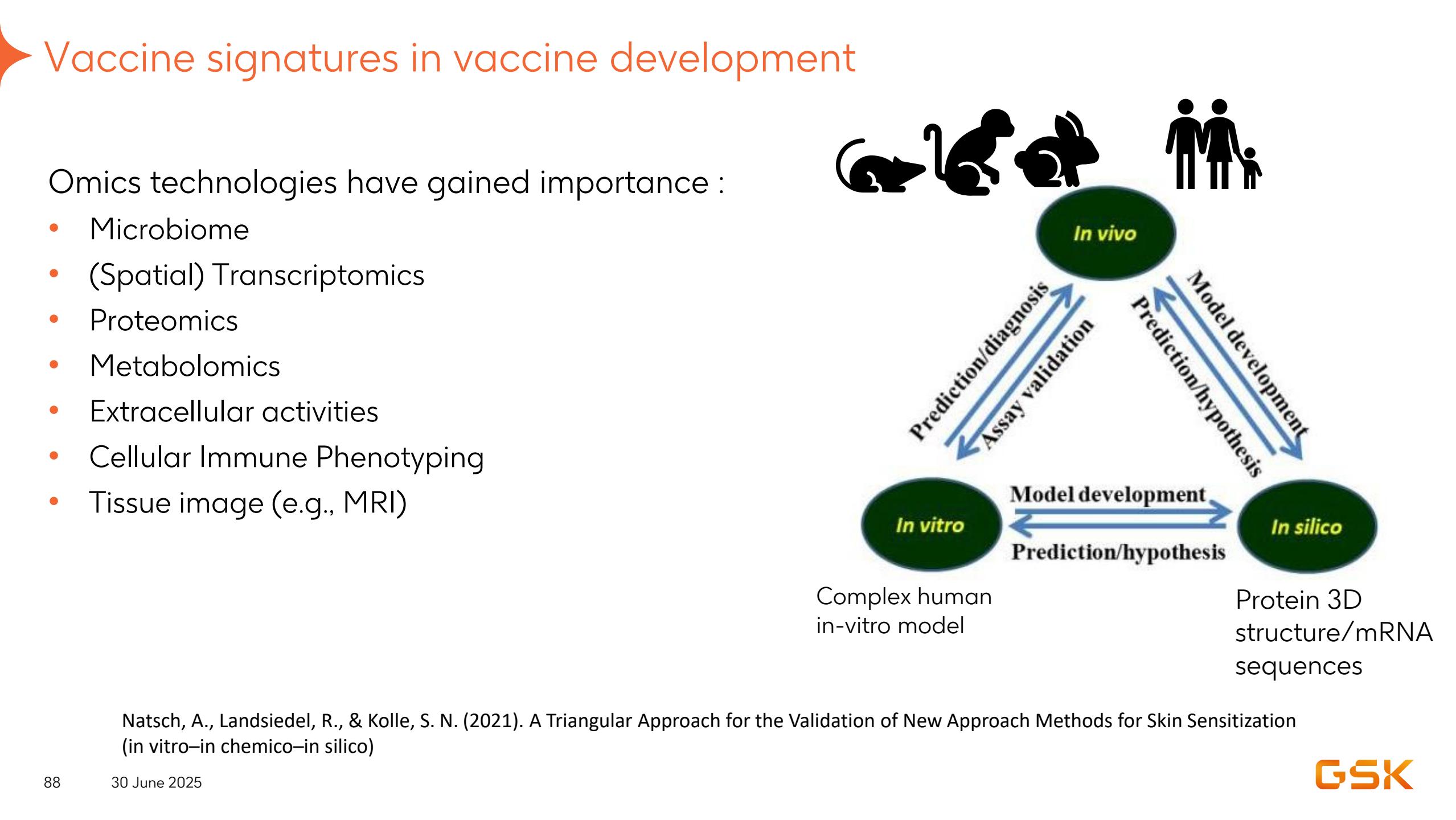
Challenge:

- **Complexity** (a combination of immune functions) and **variability** in population
- **Limitation** in accessibility of tissue vs. blood
- **Measurement:** (functional) assay qualification and validation
- **Strain specificity**

Regulatory implications: validated CoPs can accelerate licensing and reduce reliance on large efficacy trials

Correlates of Protection Induced by Vaccination (Clin Vaccine Immunol, 2010)

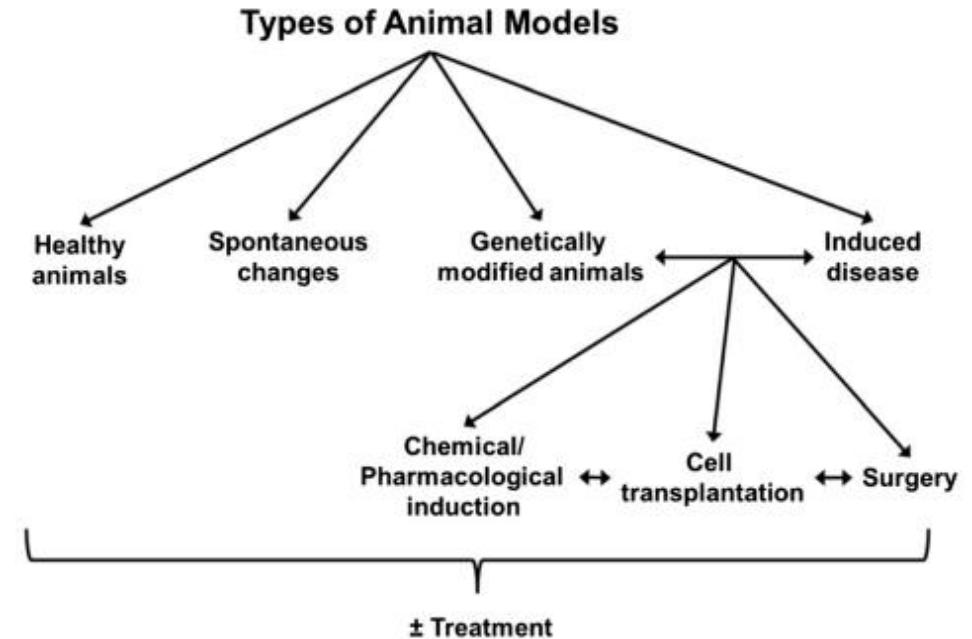
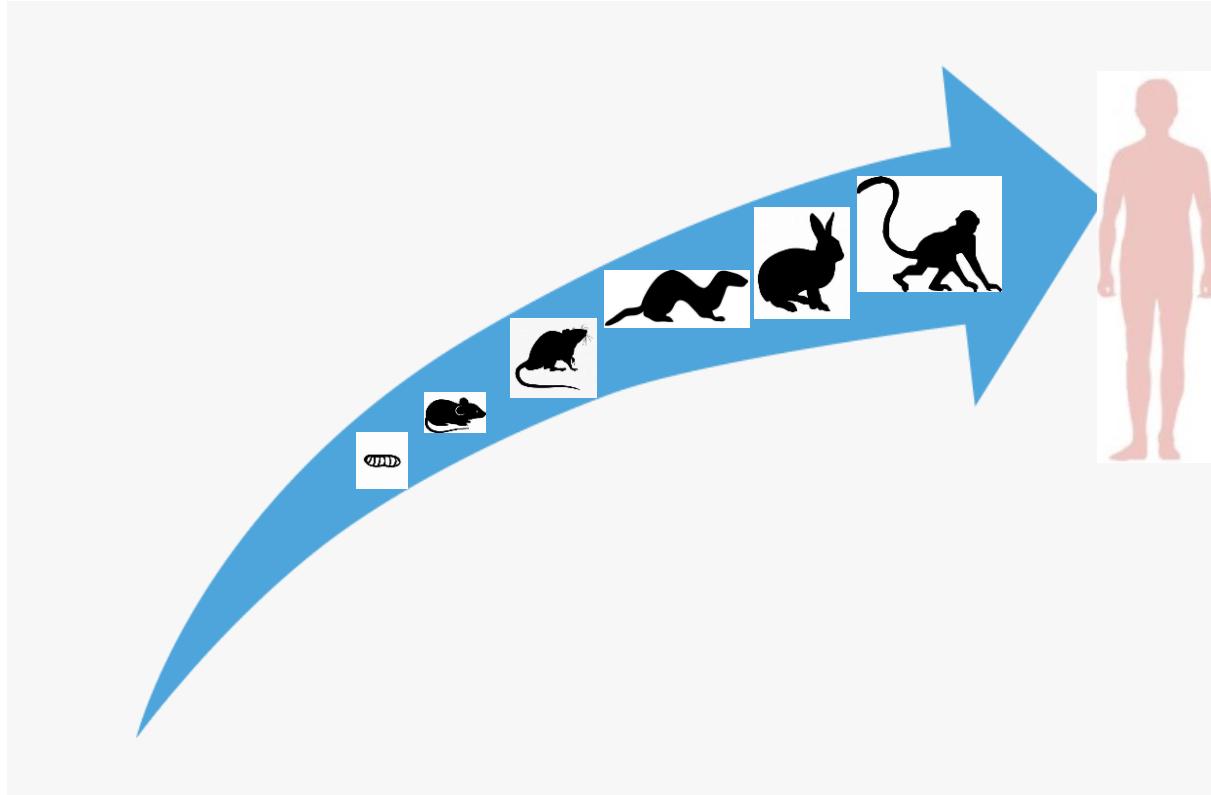
Updates on Immunologic Correlates of Vaccine-Induced Protection (Vaccine, 2020)



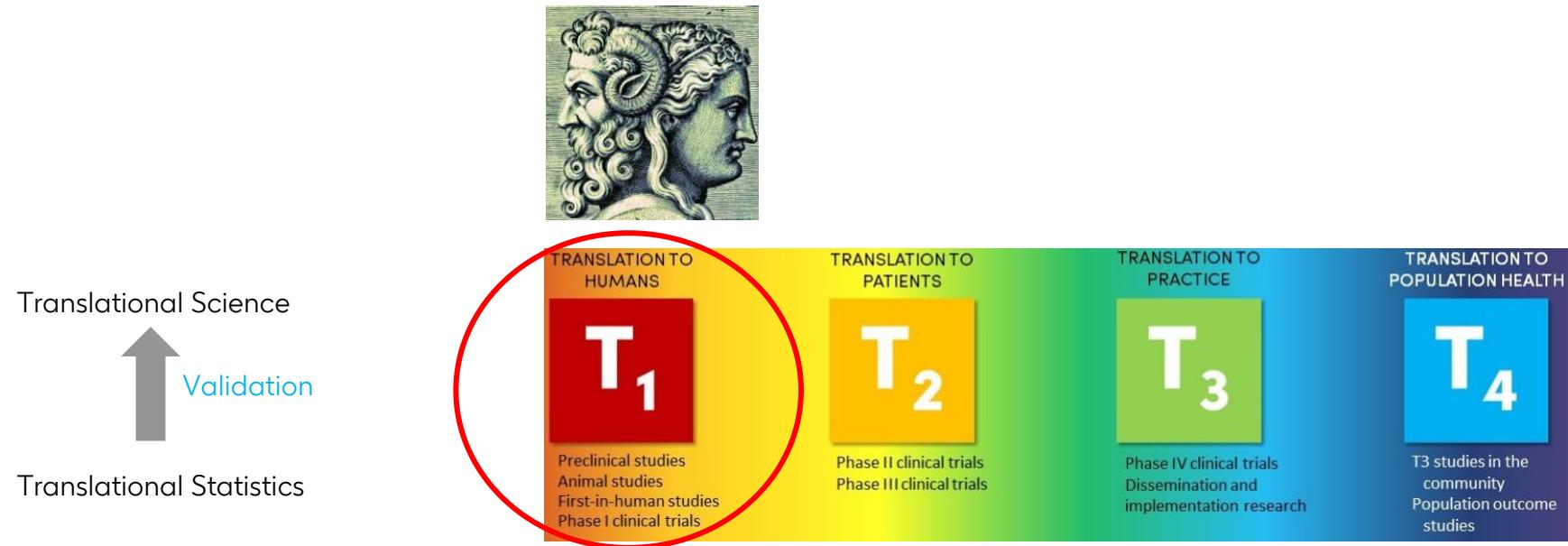
« In life », meaning the study takes place in a living organism

In vivo (Latin for « within the living ») is experimentation using a whole organism as opposed to a partial or dead organism or an in vitro controlled environment.

Animal testing and clinical trials are two forms of in vivo research.



Translational stat – what it is !



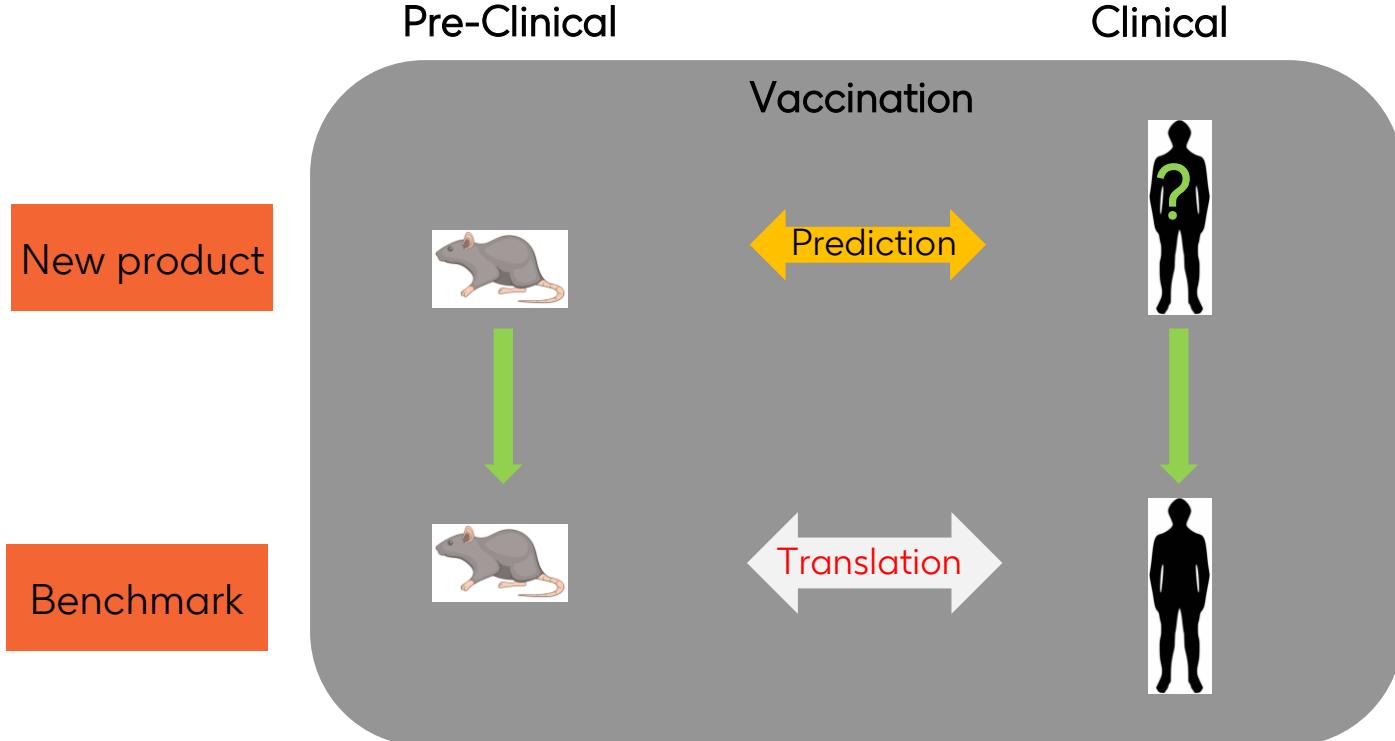
Statistics plays an important role in translational medicine to ensure that the translational process is accurate and reliable with certain statistical assurance.

It is covering :

- Statistical inference for the applicability of an animal model to a human model
- Strategies for selection of clinical study endpoints
- (e.g. absolute changes, relative changes, meaningful threshold).

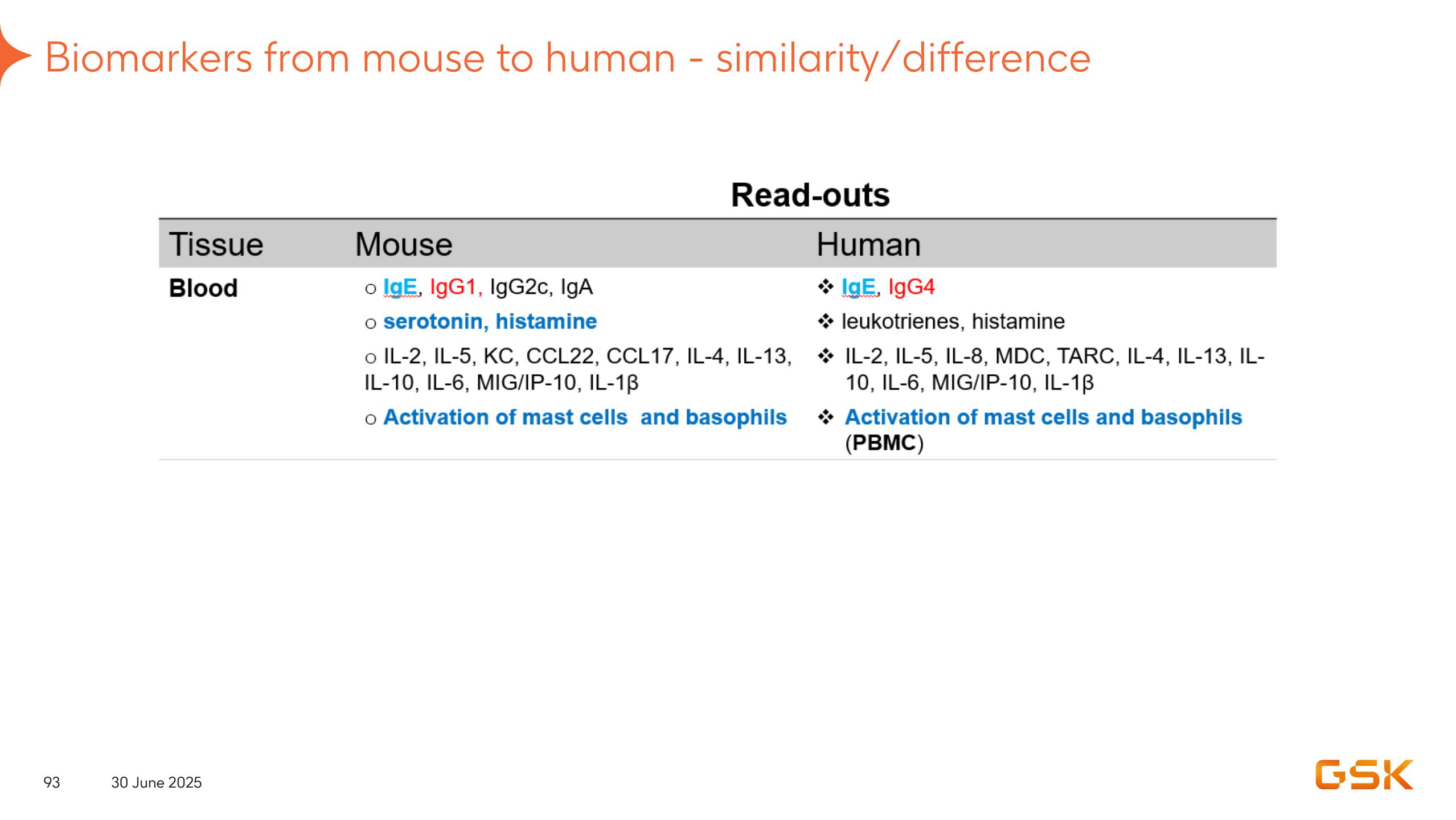
Preclinical to clinical translation

Between species



Assumptions for translation:

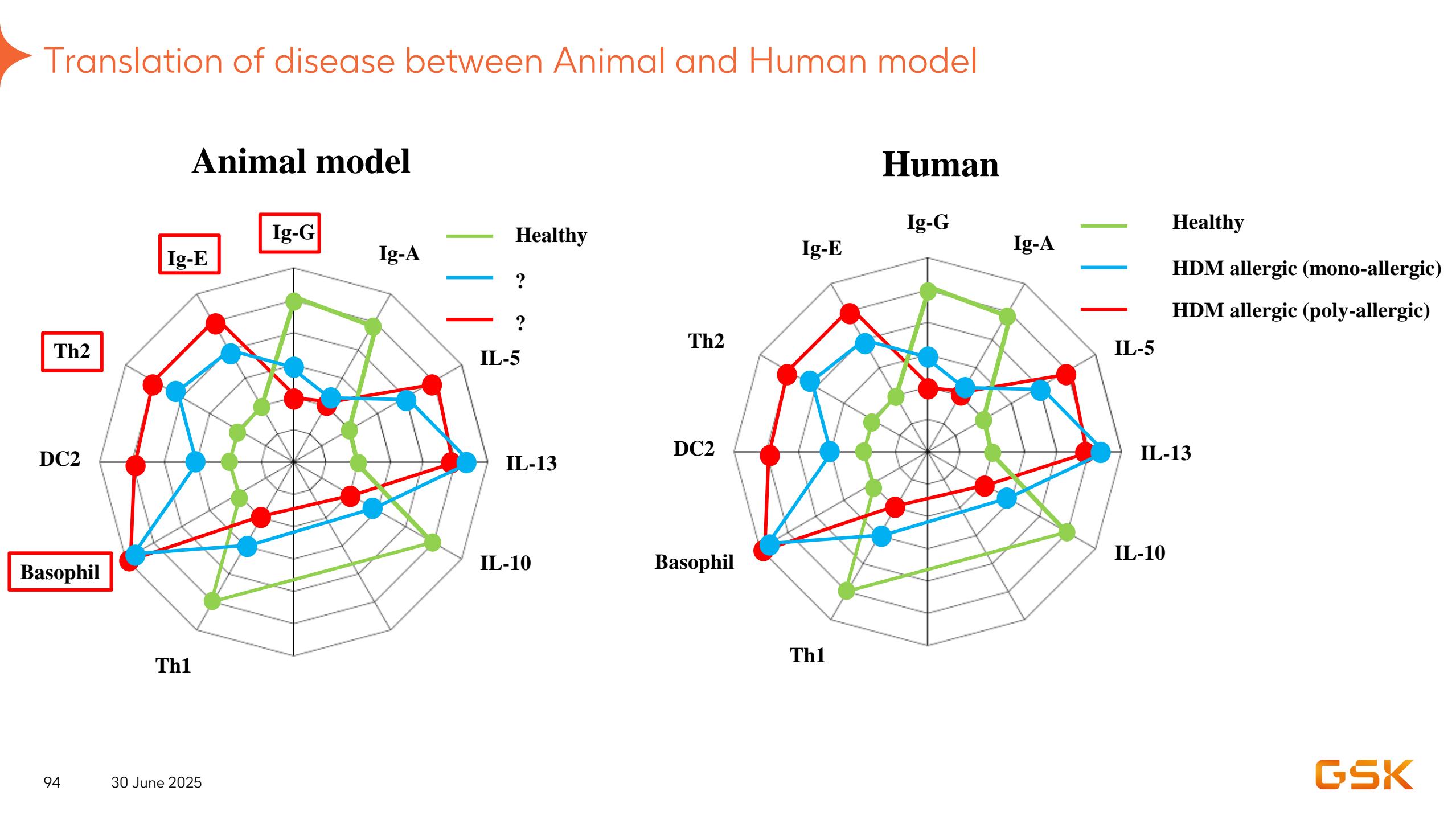
- MoA of Vaccine is similar in animal model and human
- Immune response measured in both species under similar experimental conditions
- In vitro assays for testing animal and human samples generates similar immunogenicity in both species

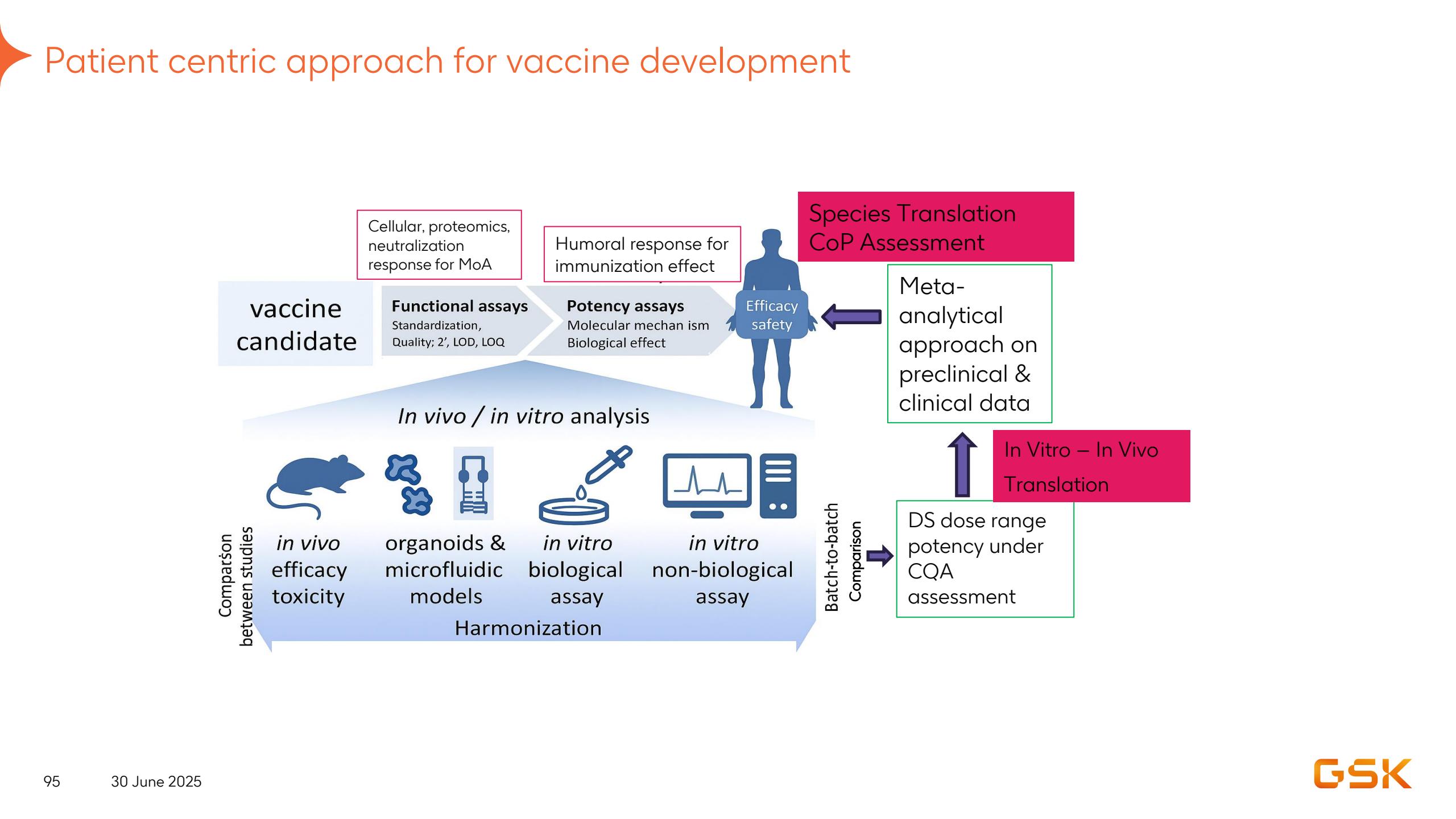


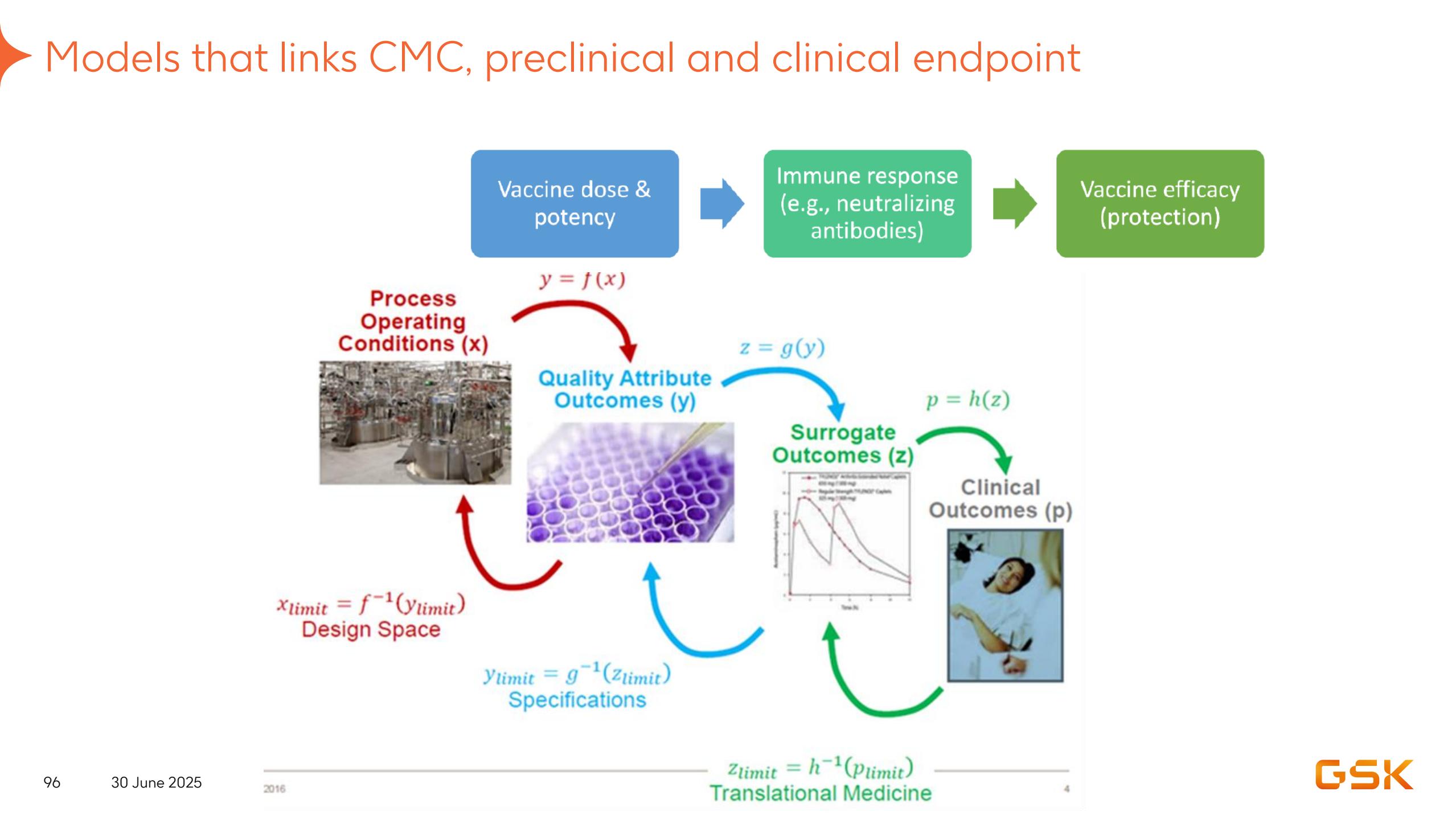
Biomarkers from mouse to human - similarity/difference

Read-outs

Tissue	Mouse	Human
Blood	<ul style="list-style-type: none">○ IgE, IgG1, IgG2c, IgA○ serotonin, histamine○ IL-2, IL-5, KC, CCL22, CCL17, IL-4, IL-13, IL-10, IL-6, MIG/IP-10, IL-1β○ Activation of mast cells and basophils	<ul style="list-style-type: none">❖ IgE, IgG4❖ leukotrienes, histamine❖ IL-2, IL-5, IL-8, MDC, TARC, IL-4, IL-13, IL-10, IL-6, MIG/IP-10, IL-1β❖ Activation of mast cells and basophils (PBMC)

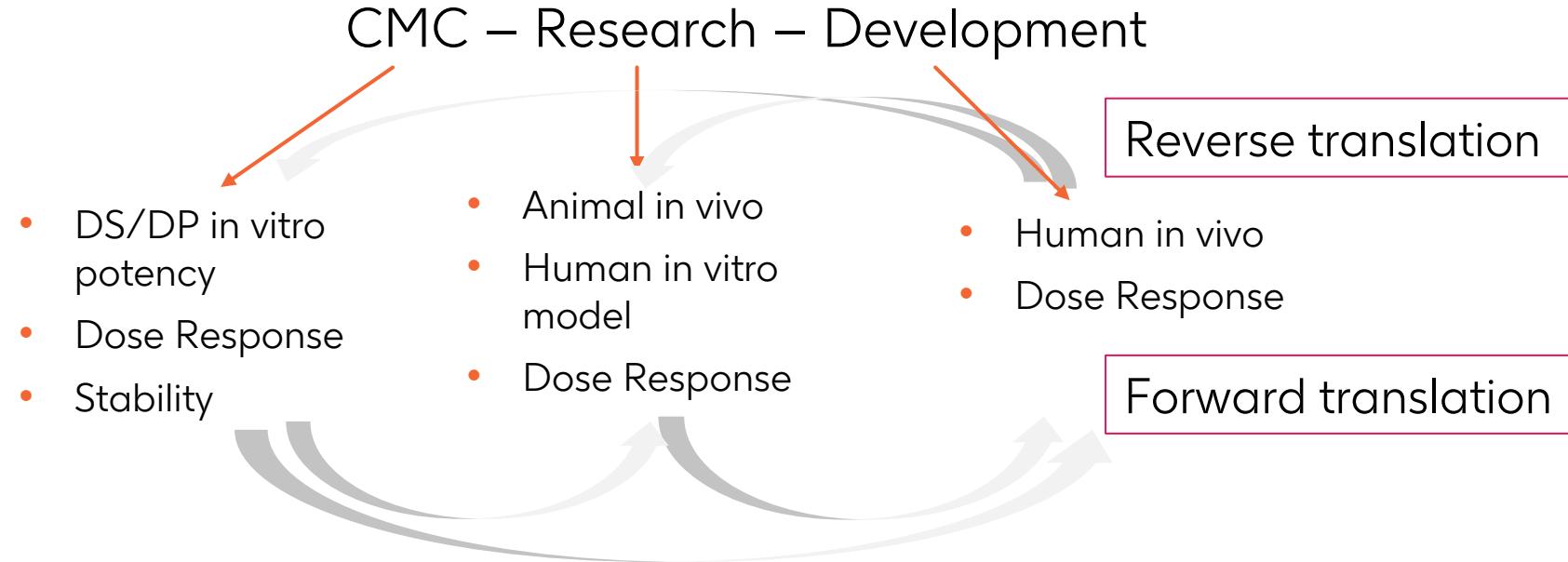






Translational framework

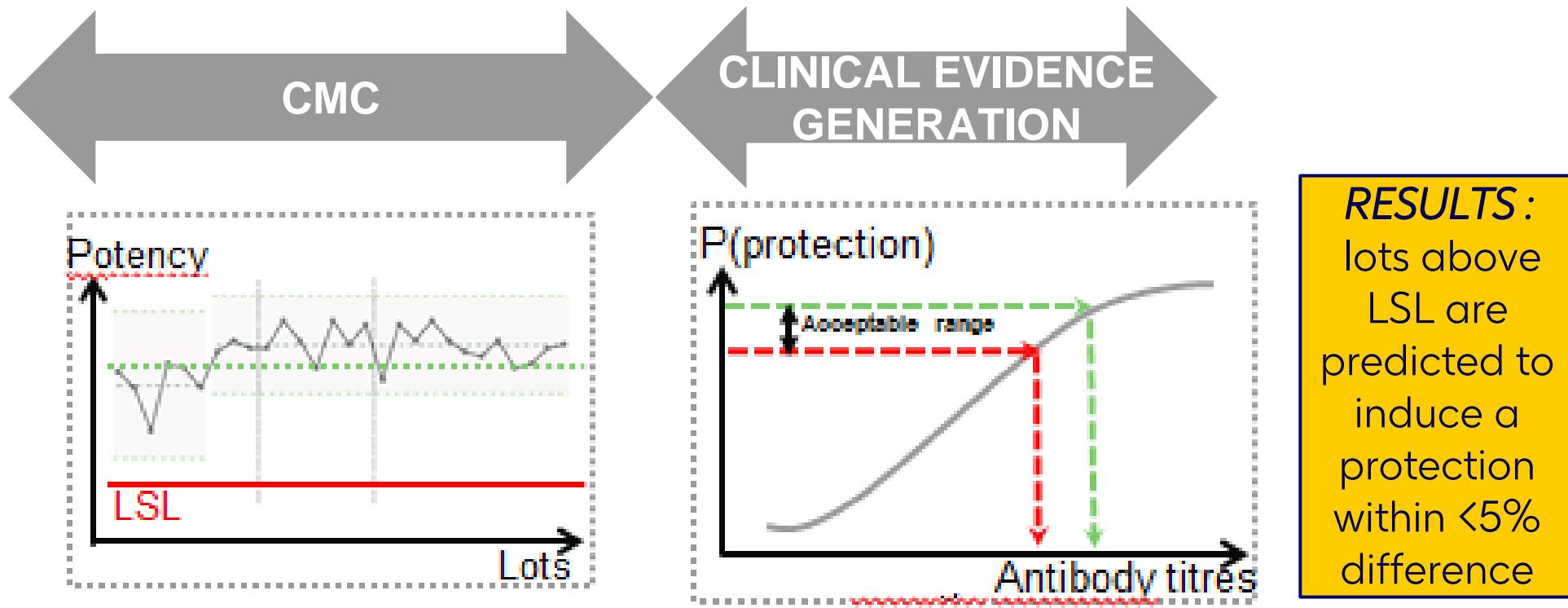
Acceleration of vaccine development



CoP can be used to define minimum on lot release specifications

Information on efficacy

Clinical justification for specifications of antigen Potency release test



Integrating evidences of CMC & preclinical to predict effect on human

CMC & preclinical to clinical translation

Summary of CQAs for vaccine product

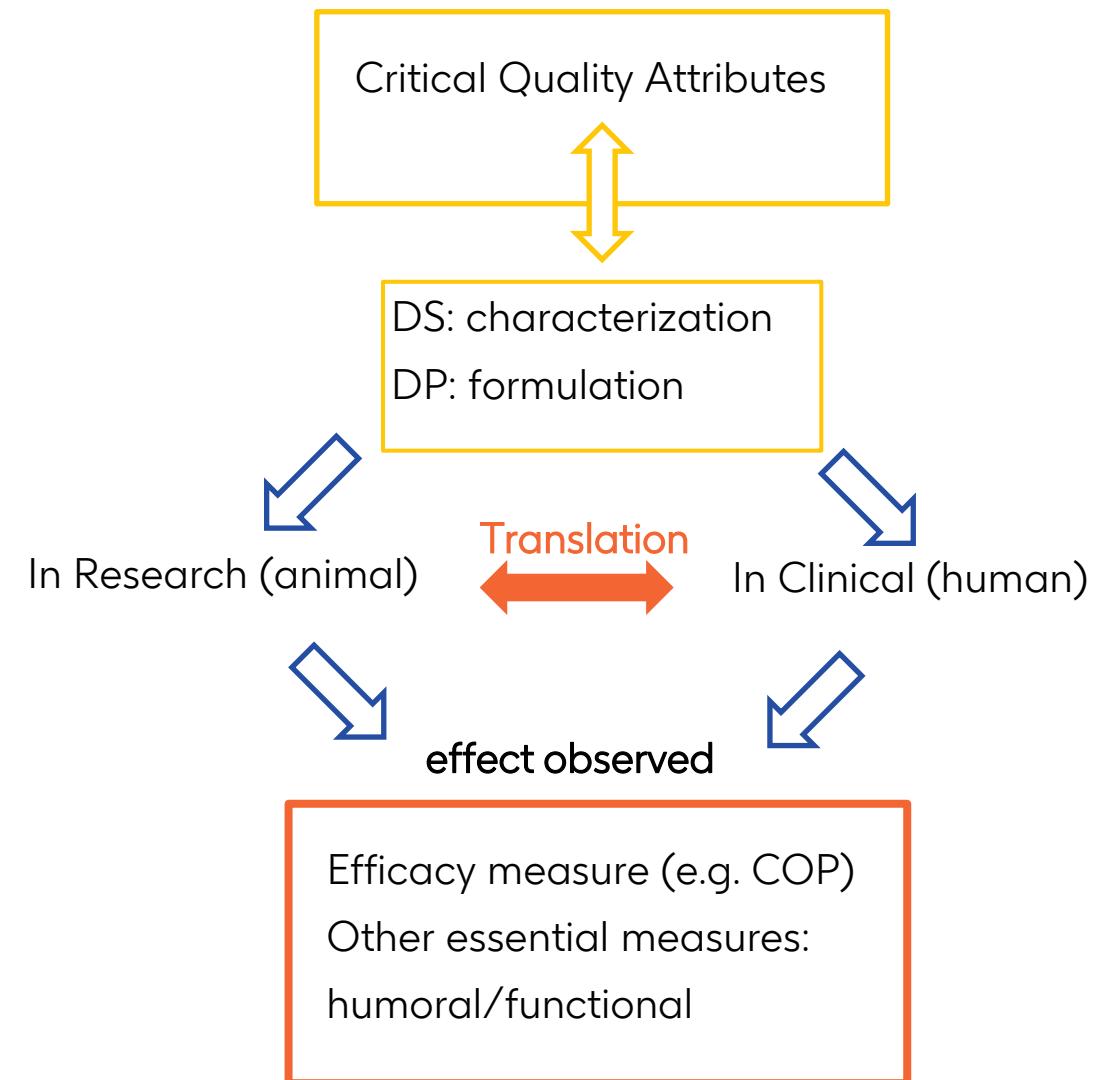
- Evaluated by dose-response models
- In relationship to vaccine potency

Bayesian meta-analysis of preclinical studies

- Historical data re-use for obtaining Meta-Analytic Predictive Prior
- Enhance future preclinical design

Translational relationship built on

- CoP and biomarkers between species
- Predictive modeling



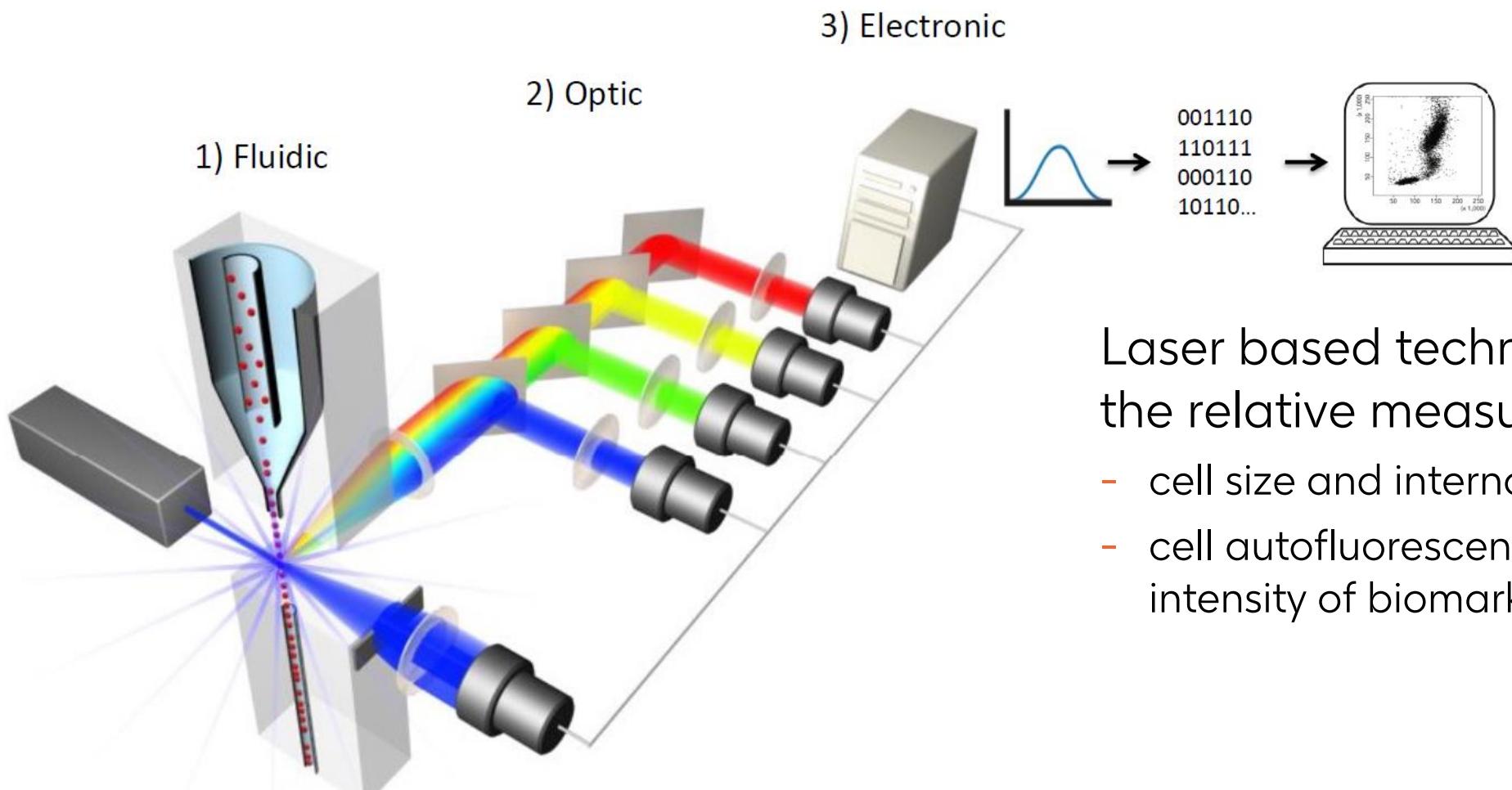
Translational topic

Dimensionality reduction and visualisation of single cell
flow cytometry data



Flow cytometry

Introduction



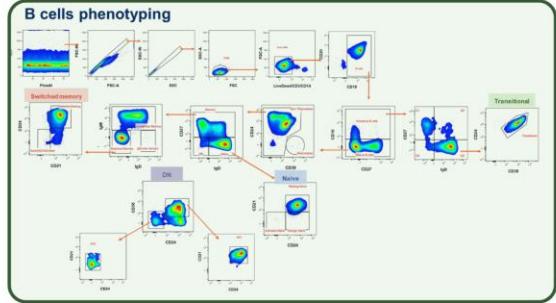
Laser based technology that allows the relative measurement of :

- cell size and internal content
- cell autofluorescence and fluorescence intensity of biomarkers

Introduction to Flow Cytometry: A Learning Guide, BD Biosciences, link: <https://www.bdbiosciences.com/en-us/learn/training/basic/flow-cytometry-introduction>

Flow cytometry in vaccine development

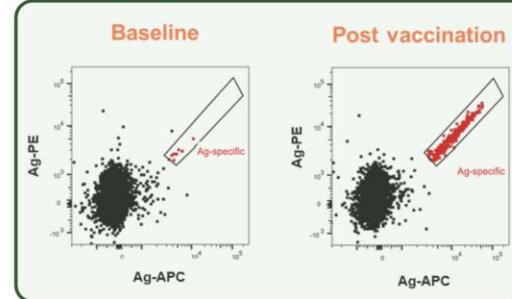
Innate/Adaptive immune Phenotyping



Biomarker Discovery

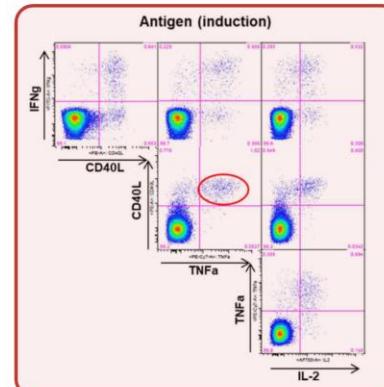
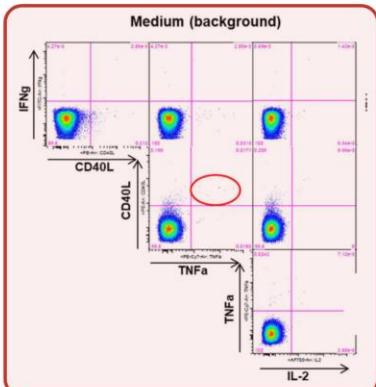


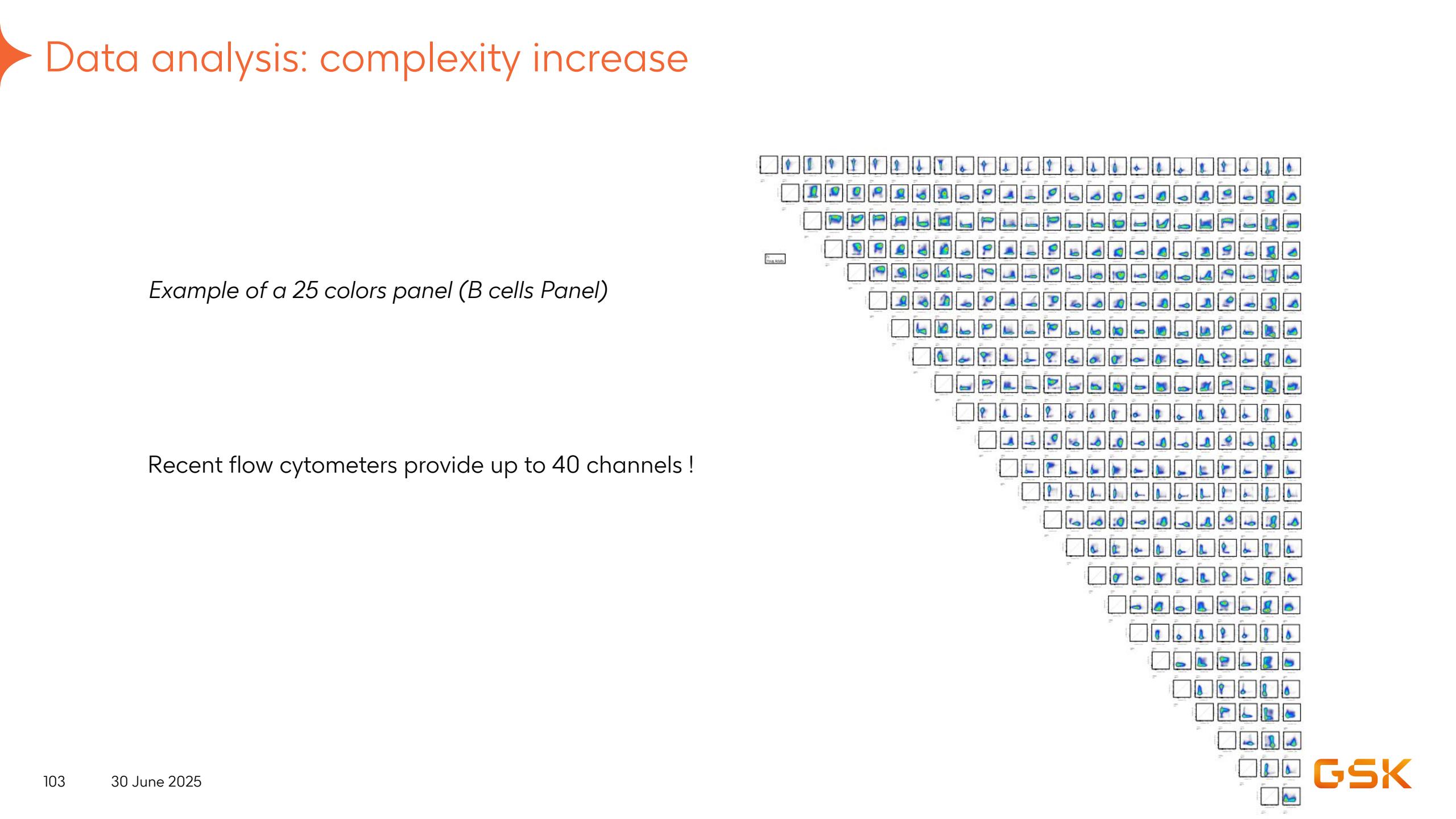
Antigen-specific T/B cells identification



Clinical endpoint

T cells functionality (ICS)

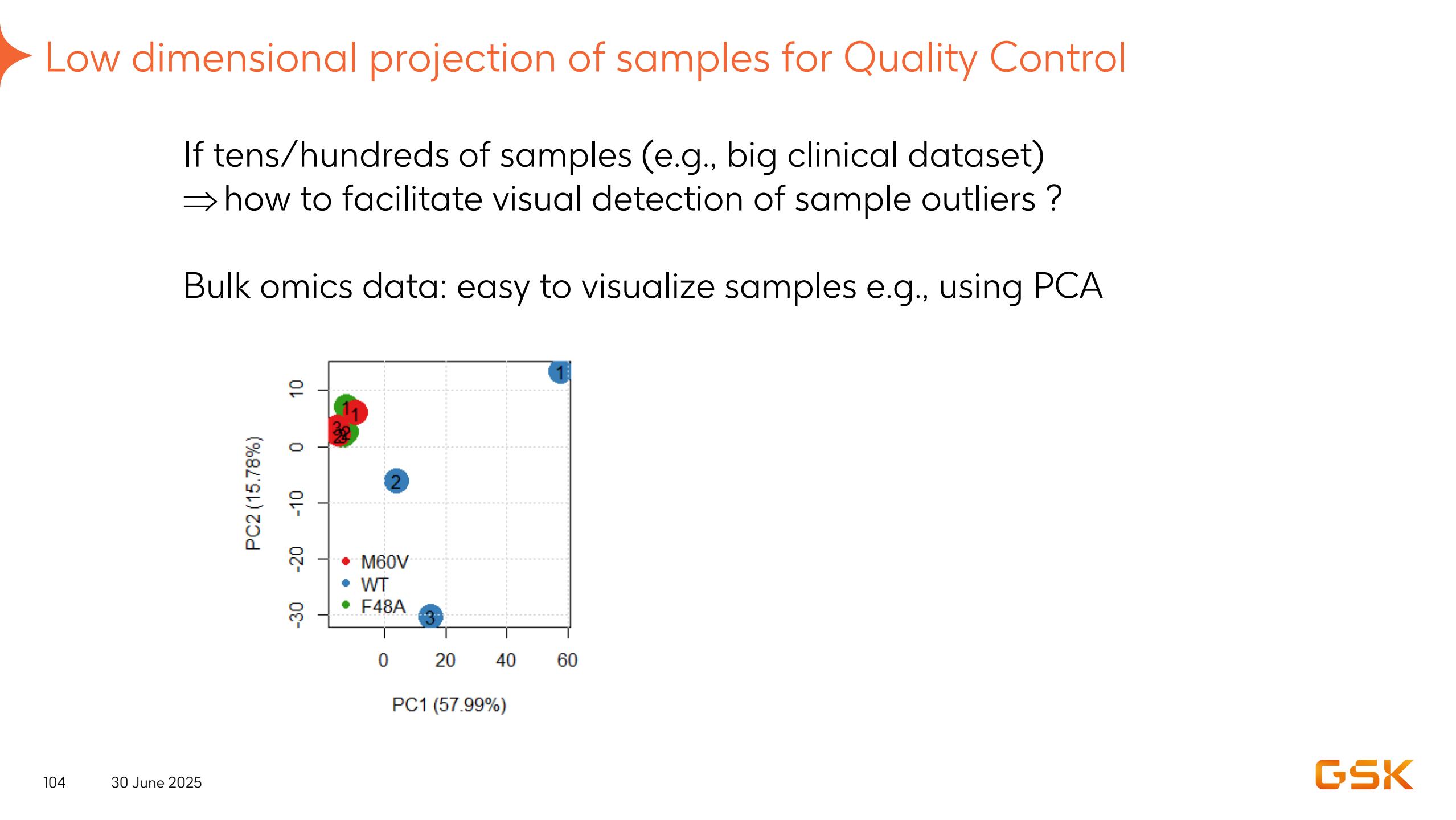


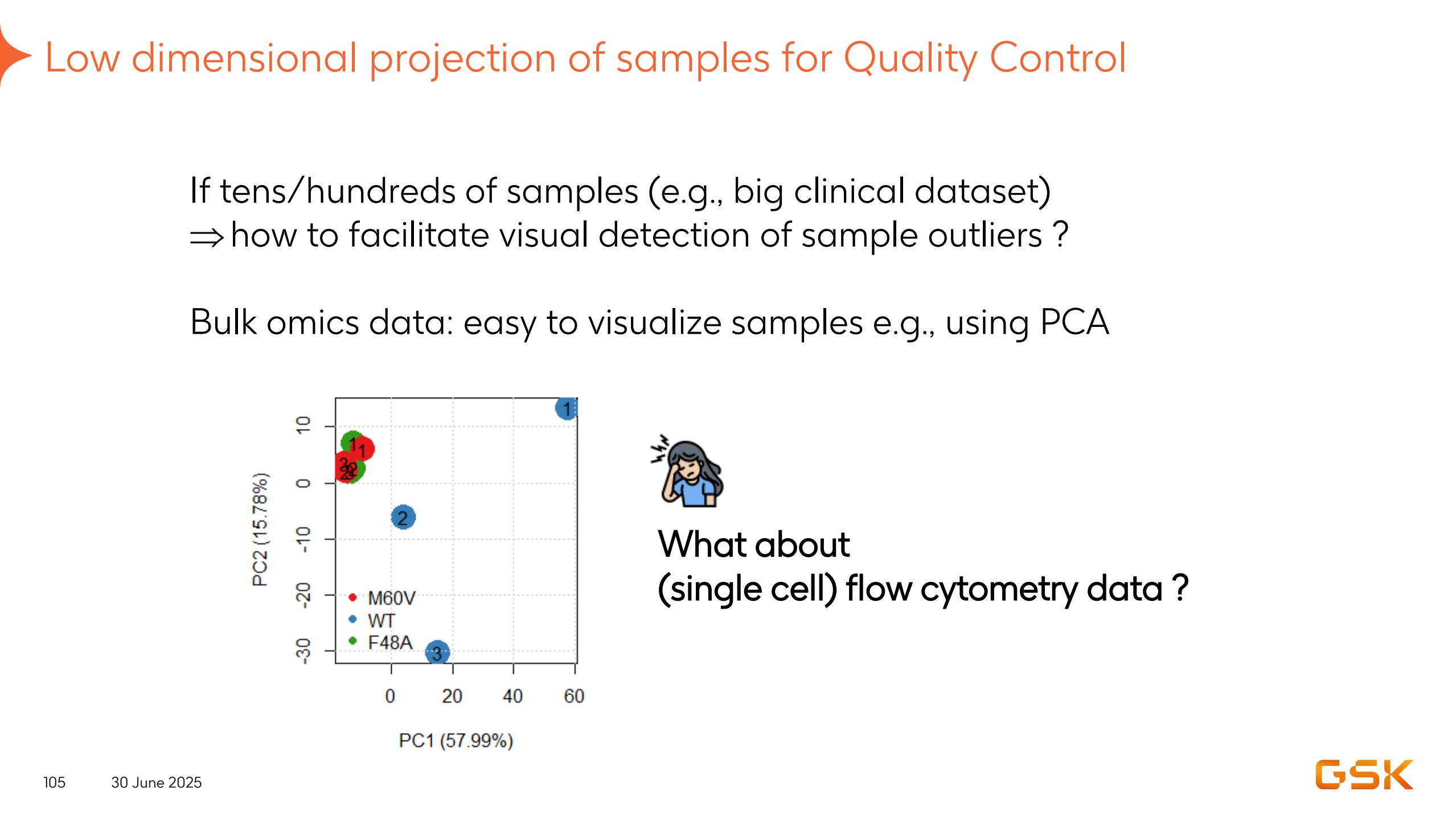


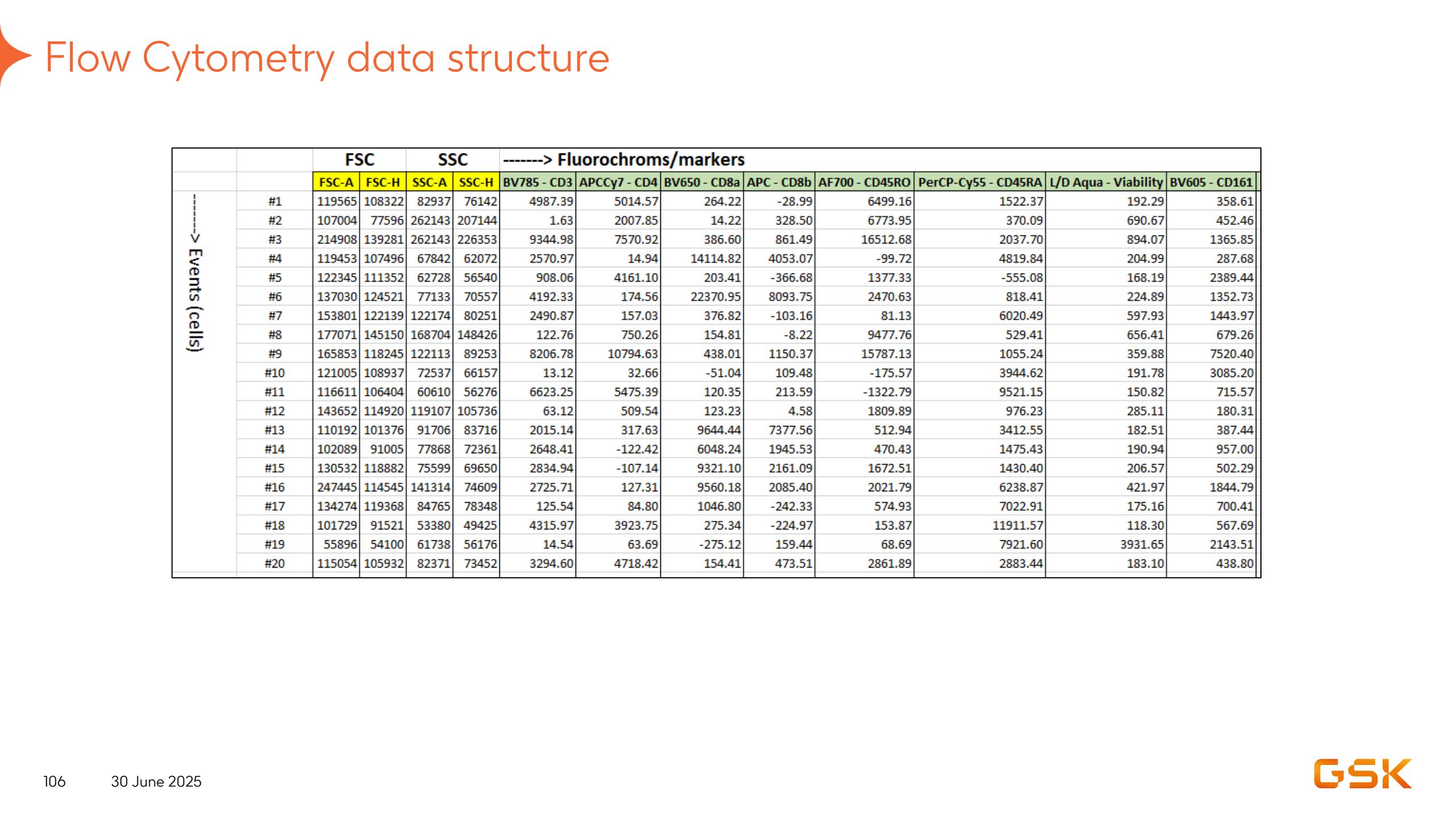
Data analysis: complexity increase

Example of a 25 colors panel (B cells Panel)

Recent flow cytometers provide up to 40 channels !



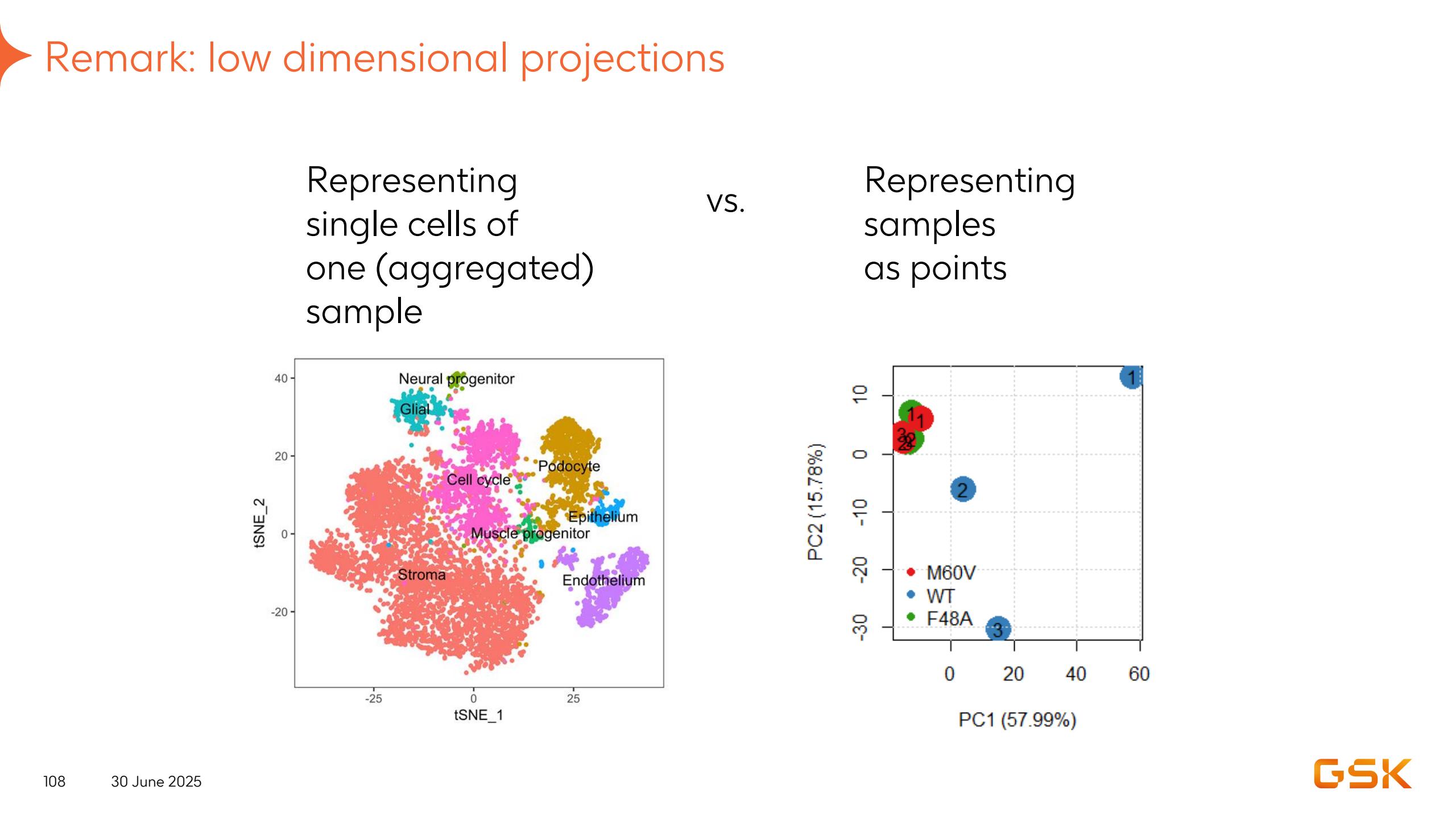




Flow Cytometry data structure

	FSC	SSC	Fluorochroms/markers											
	FSC	SSC	Fluorochroms/markers											
	FSC	SSC	Fluorochroms/markers											
	FSC	SSC	Fluorochroms/markers											
Events (cells)	FSC-A	FSC-H	SSC-A	SSC-H	BV785 - CD3	APCCy7 - CD4	BV650 - CD8a	APC - CD8b	AF700 - CD45RO	PerCP-Cy55 - CD45RA	L/D Aqua - Viability		BV605 - CD161	
	#1	119565	108322	82937	76142	4987.39	5014.57	264.22	-28.99	6499.16	1522.37	192.29	358.61	
	#2	107004	77596	262143	207144	1.63	2007.85	14.22	328.50	6773.95	370.09	690.67	452.46	
	#3	214908	139281	262143	226353	9344.98	7570.92	386.60	861.49	16512.68	2037.70	894.07	1365.85	
	#4	119453	107496	67842	62072	2570.97	14.94	14114.82	4053.07	-99.72	4819.84	204.99	287.68	
	#5	122345	111352	62728	56540	908.06	4161.10	203.41	-366.68	1377.33	-555.08	168.19	2389.44	
	#6	137030	124521	77133	70557	4192.33	174.56	22370.95	8093.75	2470.63	818.41	224.89	1352.73	
	#7	153801	122139	122174	80251	2490.87	157.03	376.82	-103.16	81.13	6020.49	597.93	1443.97	
	#8	177071	145150	168704	148426	122.76	750.26	154.81	-8.22	9477.76	529.41	656.41	679.26	
	#9	165853	118245	122113	89253	8206.78	10794.63	438.01	1150.37	15787.13	1055.24	359.88	7520.40	
	#10	121005	108937	72537	66157	13.12	32.66	-51.04	109.48	-175.57	3944.62	191.78	3085.20	
	#11	116611	106404	60610	56276	6623.25	5475.39	120.35	213.59	-1322.79	9521.15	150.82	715.57	
	#12	143652	114920	119107	105736	63.12	509.54	123.23	4.58	1809.89	976.23	285.11	180.31	
	#13	110192	101376	91706	83716	2015.14	317.63	9644.44	7377.56	512.94	3412.55	182.51	387.44	
	#14	102089	91005	77868	72361	2648.41	-122.42	6048.24	1945.53	470.43	1475.43	190.94	957.00	
	#15	130532	118882	75599	69650	2834.94	-107.14	9321.10	2161.09	1672.51	1430.40	206.57	502.29	
	#16	247445	114545	141314	74609	2725.71	127.31	9560.18	2085.40	2021.79	6238.87	421.97	1844.79	
	#17	134274	119368	84765	78348	125.54	84.80	1046.80	-242.33	574.93	7022.91	175.16	700.41	
	#18	101729	91521	53380	49425	4315.97	3923.75	275.34	-224.97	153.87	11911.57	118.30	567.69	
	#19	55896	54100	61738	56176	14.54	63.69	-275.12	159.44	68.69	7921.60	3931.65	2143.51	
	#20	115054	105932	82371	73452	3294.60	4718.42	154.41	473.51	2861.89	2883.44	183.10	438.80	

⇒ One data matrix per biological sample
 ⇒ Nb of events (cells) is variable from sample to sample



Sample QC visualization:

an innovative visualization tool to enhance sample quality control

Context:

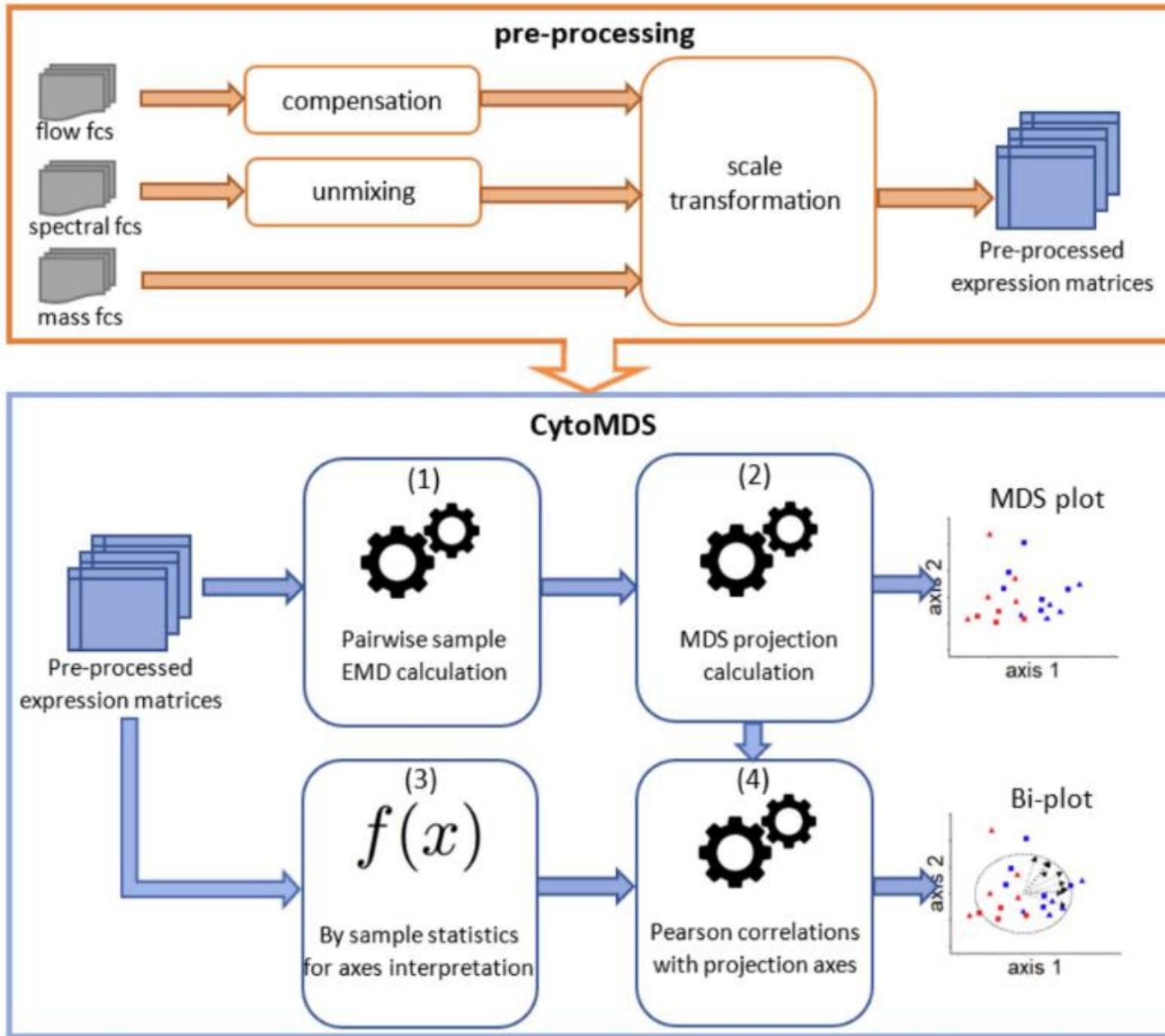
- medium to large study with several tens (or hundreds) of samples
- design of experiment (possibly with several controlled factors)
- possible batch effects (e.g., data acquisition spanning several days)
- Effect of factors (potential biomarkers)



Quality control (before and/or after preprocessing) of each sample, one by one, might be difficult and is time consuming



How to visualize all samples at once, on (a) simple plot(s), in order to evaluate ‘distances’ between each samples ? (=> identify outliers, batch effects, clusters corresponding to DoE, etc.)



Hauchamps et al. 2025

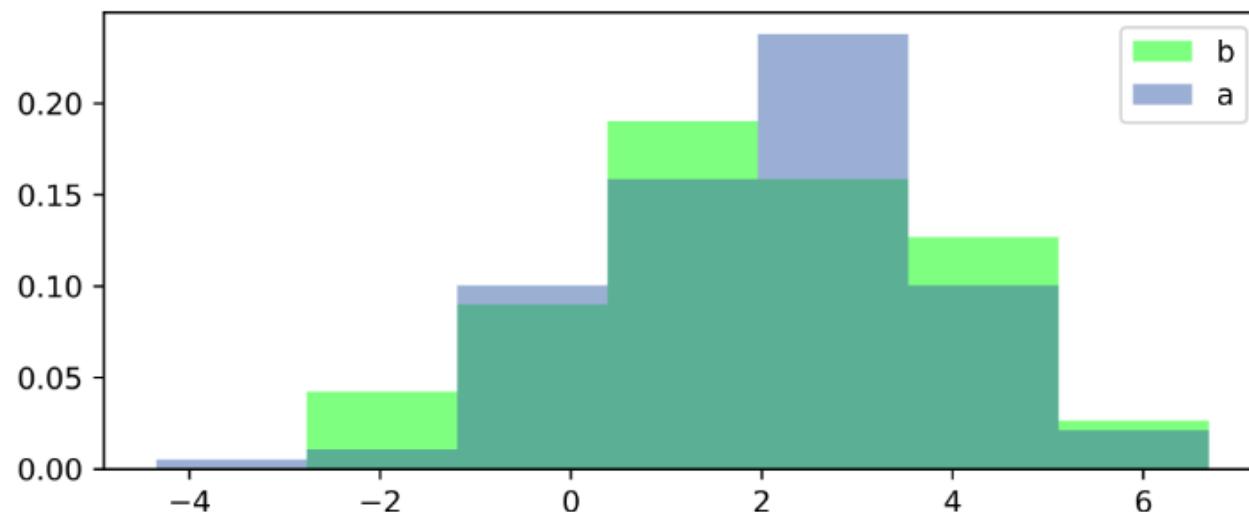


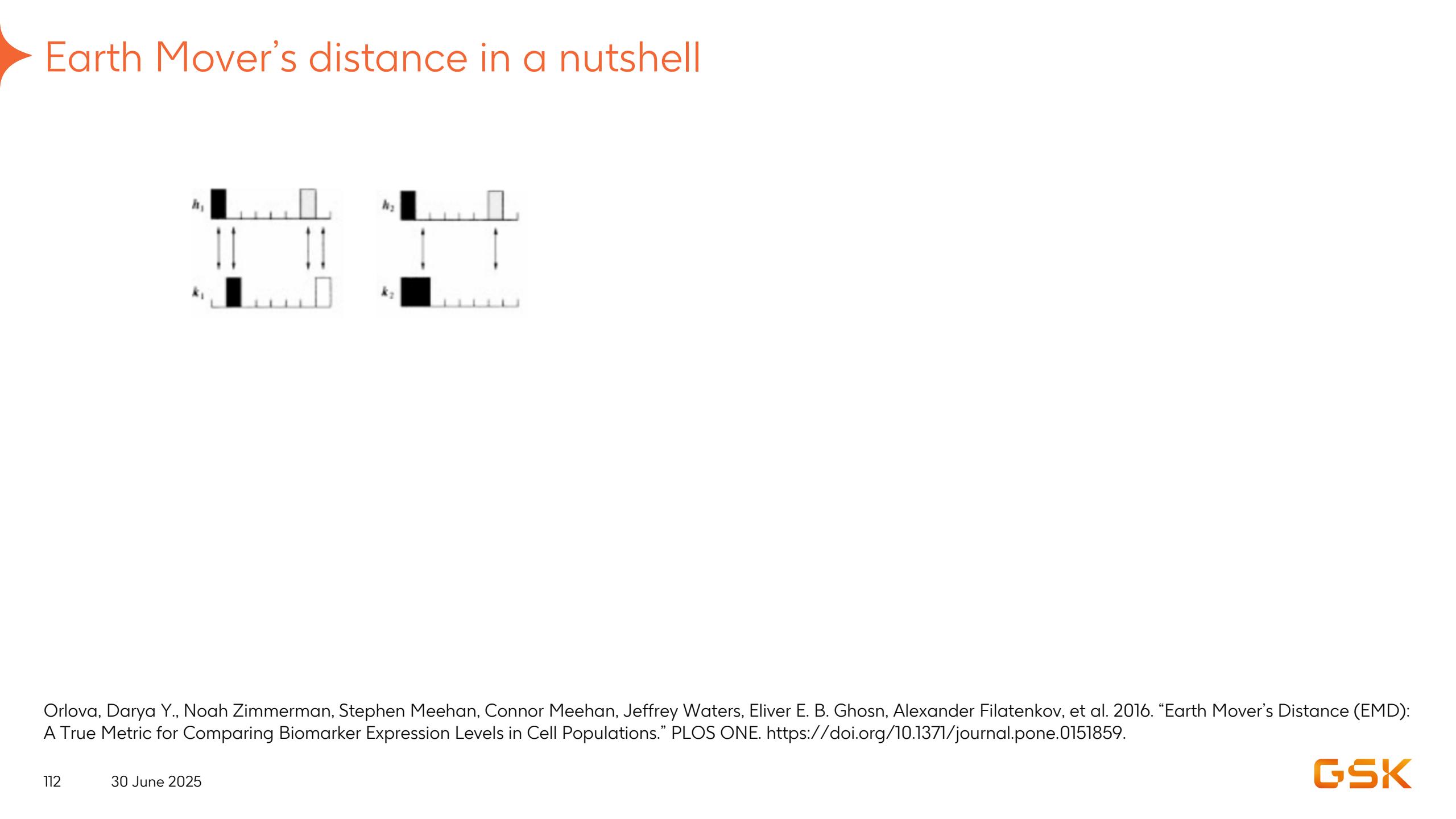
Probability binning distances

Existence of many possible mathematical definitions
of a metric between distributions

- ⇒ Natural to use distribution discretization by histograms
- ⇒ *Probability binning distances* sum bin-to-bin differences

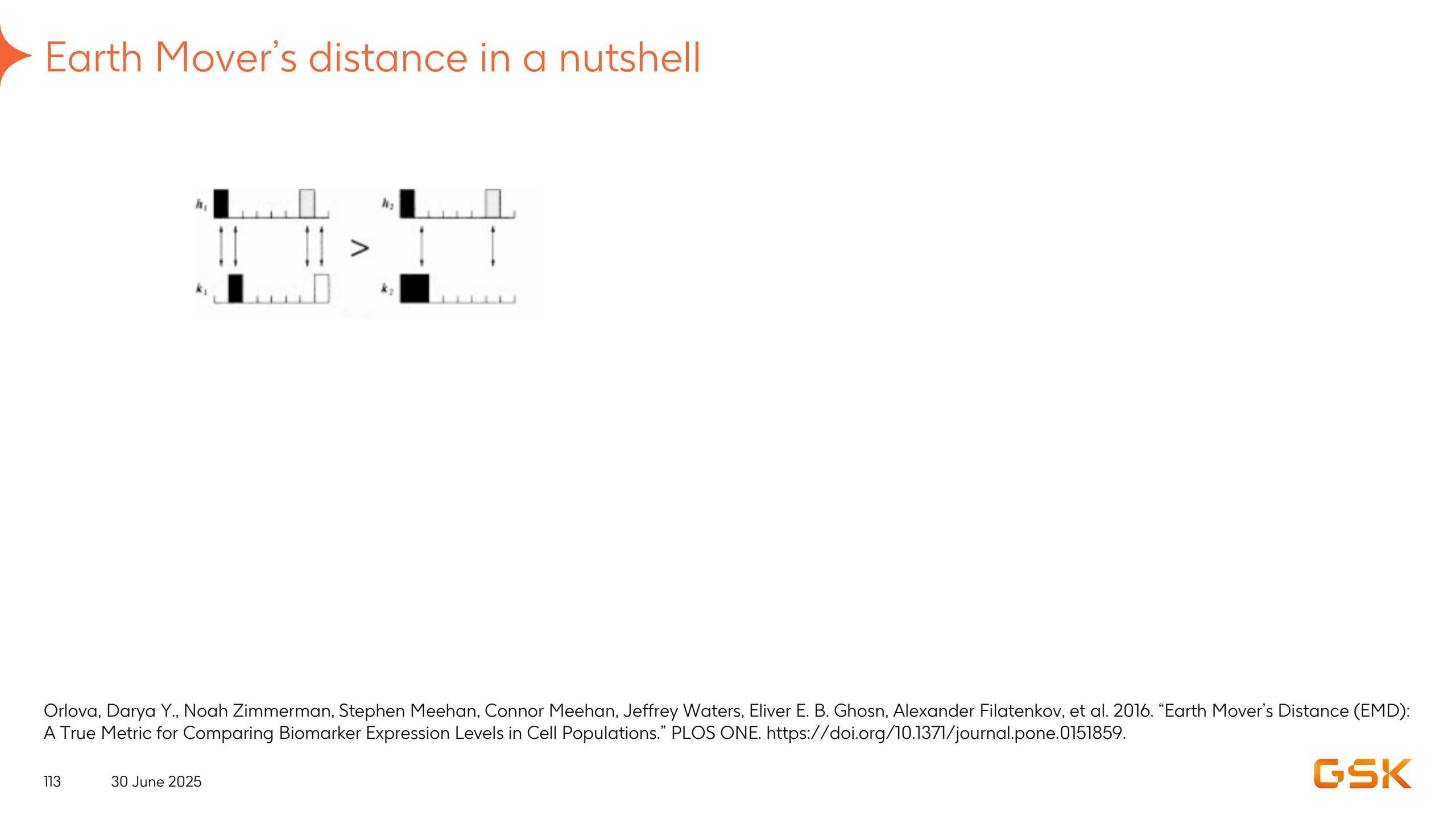
comparison of a and b





Earth Mover's distance in a nutshell

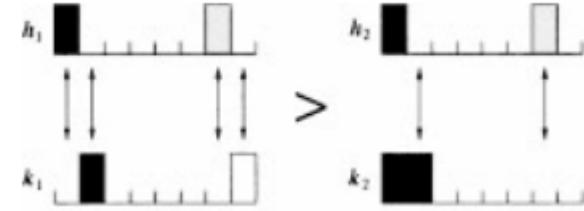
Orlova, Darya Y., Noah Zimmerman, Stephen Meehan, Connor Meehan, Jeffrey Waters, Eliver E. B. Ghosn, Alexander Filatenkov, et al. 2016. "Earth Mover's Distance (EMD): A True Metric for Comparing Biomarker Expression Levels in Cell Populations." PLOS ONE. <https://doi.org/10.1371/journal.pone.0151859>.



Earth Mover's distance in a nutshell

Orlova, Darya Y., Noah Zimmerman, Stephen Meehan, Connor Meehan, Jeffrey Waters, Eliver E. B. Ghosn, Alexander Filatenkov, et al. 2016. "Earth Mover's Distance (EMD): A True Metric for Comparing Biomarker Expression Levels in Cell Populations." PLOS ONE. <https://doi.org/10.1371/journal.pone.0151859>.

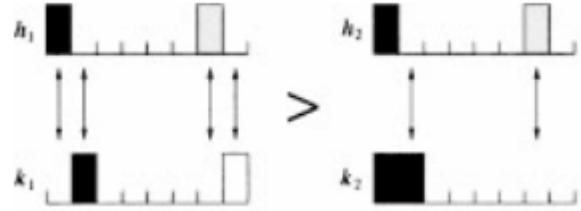
Earth Mover's distance in a nutshell



=> counter-intuitive!

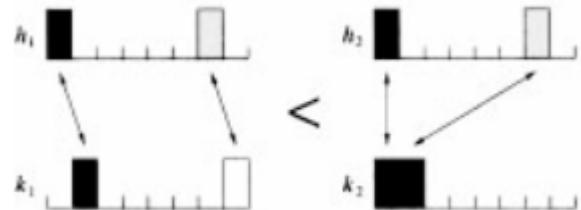
Orlova, Darya Y., Noah Zimmerman, Stephen Meehan, Connor Meehan, Jeffrey Waters, Eliver E. B. Ghosn, Alexander Filatenkov, et al. 2016. "Earth Mover's Distance (EMD): A True Metric for Comparing Biomarker Expression Levels in Cell Populations." PLOS ONE. <https://doi.org/10.1371/journal.pone.0151859>.

Earth Mover's distance in a nutshell



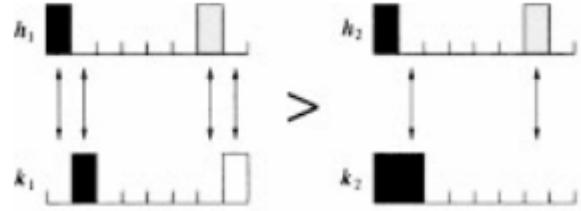
=> counter-intuitive!

Earth Mover's distance (aka 1-Wasserstein metric)
computes the effort/cost (= mass x distance) needed
to transport the mass of one distribution to obtain the other



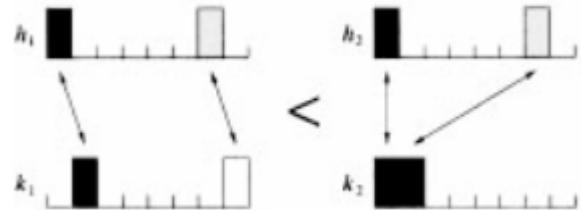
Orlova, Darya Y., Noah Zimmerman, Stephen Meehan, Connor Meehan, Jeffrey Waters, Eliver E. B. Ghosn, Alexander Filatenkov, et al. 2016. "Earth Mover's Distance (EMD): A True Metric for Comparing Biomarker Expression Levels in Cell Populations." PLOS ONE. <https://doi.org/10.1371/journal.pone.0151859>.

Earth Mover's distance in a nutshell



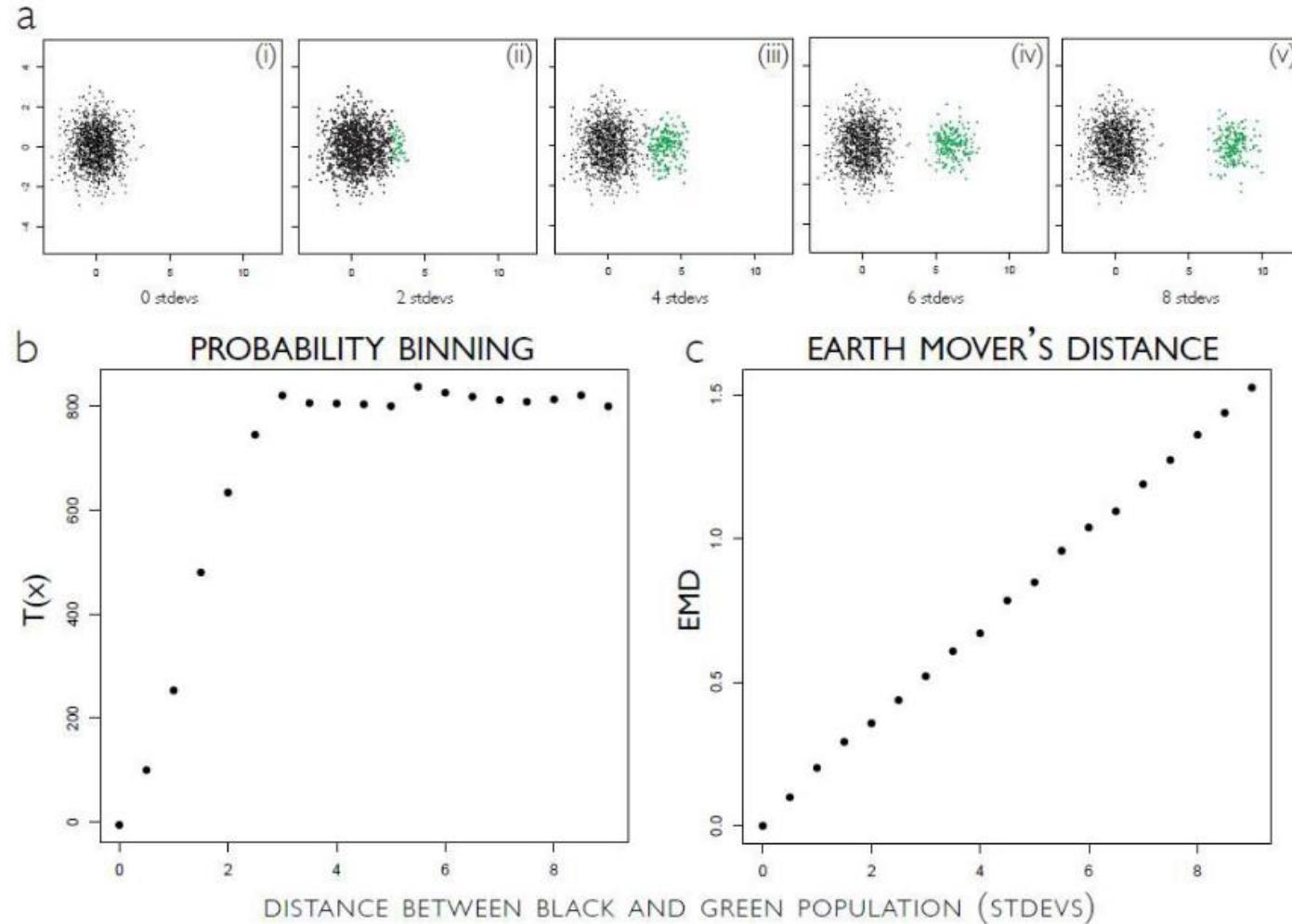
=> counter-intuitive!

Earth Mover's distance (aka 1-Wasserstein metric)
computes the effort/cost (= mass x distance) needed
to transport the mass of one distribution to obtain the other



Orlova, Darya Y., Noah Zimmerman, Stephen Meehan, Connor Meehan, Jeffrey Waters, Eliver E. B. Ghosn, Alexander Filatenkov, et al. 2016. "Earth Mover's Distance (EMD): A True Metric for Comparing Biomarker Expression Levels in Cell Populations." PLOS ONE. <https://doi.org/10.1371/journal.pone.0151859>.

Earth Mover's distance in a nutshell (2)

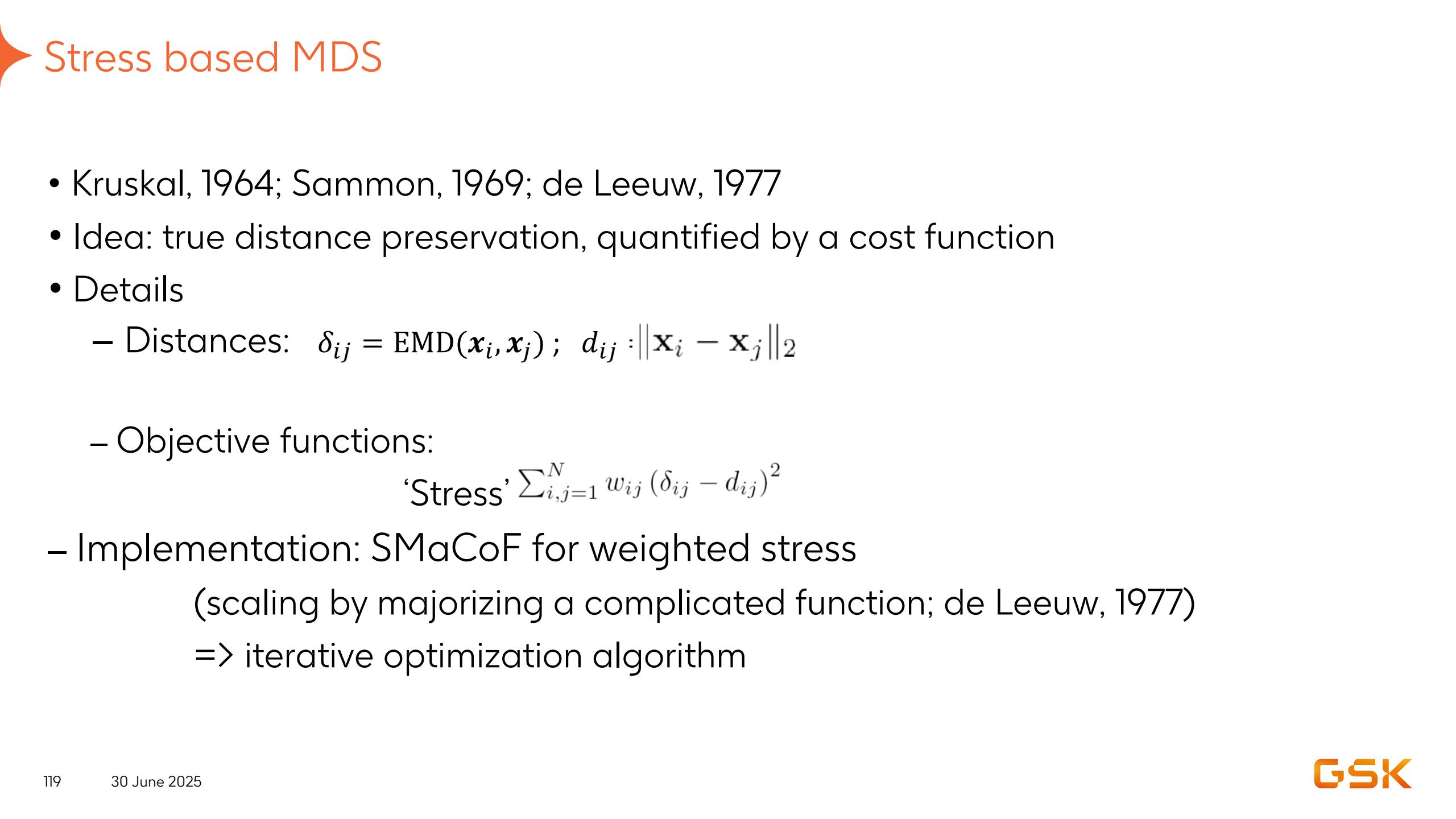


Orlova, Darya Y., Noah Zimmerman, Stephen Meehan, Connor Meehan, Jeffrey Waters, Eliver E. B. Ghosn, Alexander Filatenkov, et al. 2016. "Earth Mover's Distance (EMD): A True Metric for Comparing Biomarker Expression Levels in Cell Populations." PLOS ONE. <https://doi.org/10.1371/journal.pone.0151859>.



Multi Dimensional Scaling (MDS)

- When computing low dimensional projections for QC purpose, we are interested in distance preservation
⇒ favour ‘PCA alike’ method
- PCA needs a matrix of high dimensional coordinates as an input (but we only have a matrix of pairwise distance)
- If the metric is Euclidean, *classical metric MDS* gives the same solution as PCA (and it uses the gram matrix – obtained from distance matrix – as an input)
- Here the metric is not Euclidean => use *Stress Based MDS*



Stress based MDS

- Kruskal, 1964; Sammon, 1969; de Leeuw, 1977
- Idea: true distance preservation, quantified by a cost function
- Details
 - Distances: $\delta_{ij} = \text{EMD}(\mathbf{x}_i, \mathbf{x}_j)$; $d_{ij} : \|\mathbf{x}_i - \mathbf{x}_j\|_2$
 - Objective functions:
‘Stress’ $\sum_{i,j=1}^N w_{ij} (\delta_{ij} - d_{ij})^2$
 - Implementation: SMaCoF for weighted stress
(scaling by majorizing a complicated function; de Leeuw, 1977)
=> iterative optimization algorithm

Application of CytoMDS

ImmunoSenescence study

- Human Peripheral Blood Mononuclear Cell samples
- Young adults vs. older adults of healthy donors
- Two rounds of data acquisition (panel set-up phase), each time 5 vs. 5
- Some samples in common between the former and later data acquisition

Person Id	Group	Former Panel	Later Panel
71	Young Adult		L71
79	Young Adult		L79
81	Young Adult	F81	L81
82	Young Adult	F82	
85	Young Adult	F85	
86	Young Adult	F86	L86
92	Young Adult	F92	L92
97	Old Adult	F097	
285	Old Adult		L285
382	Old Adult	F382	L382
410	Old Adult	F410	L410
490	Old Adult		L490
559	Old Adult		L559
587	Old Adult	F587	
697	Old Adult	F697	

Application of CytoMDS

ImmunoSenescence study

- Pairwise sample EMD were calculated using the channels
 - FSC-A (forward scatter area): proportional to cell size
 - FSC-H (forward scatter height): peak height of the pulse to isolate singlets
 - SSC-A (side scatter area): proportional to internal granularity
 - Live/Dead

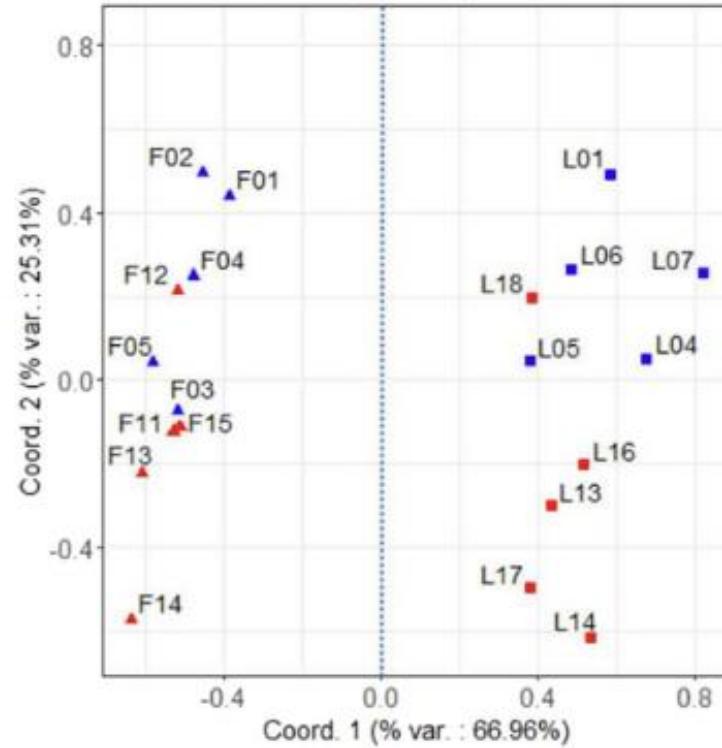
Person Id	Group	Former Panel	Later Panel
71	Young Adult		L71
79	Young Adult		L79
81	Young Adult	F81	L81
82	Young Adult	F82	
85	Young Adult	F85	
86	Young Adult	F86	L86
92	Young Adult	F92	L92
97	Old Adult	F097	
285	Old Adult		L285
382	Old Adult	F382	L382
410	Old Adult	F410	L410
490	Old Adult		L490
559	Old Adult		L559
587	Old Adult	F587	
697	Old Adult	F697	

Application of CytoMDS

ImmunoSenescence Human PBMC dataset

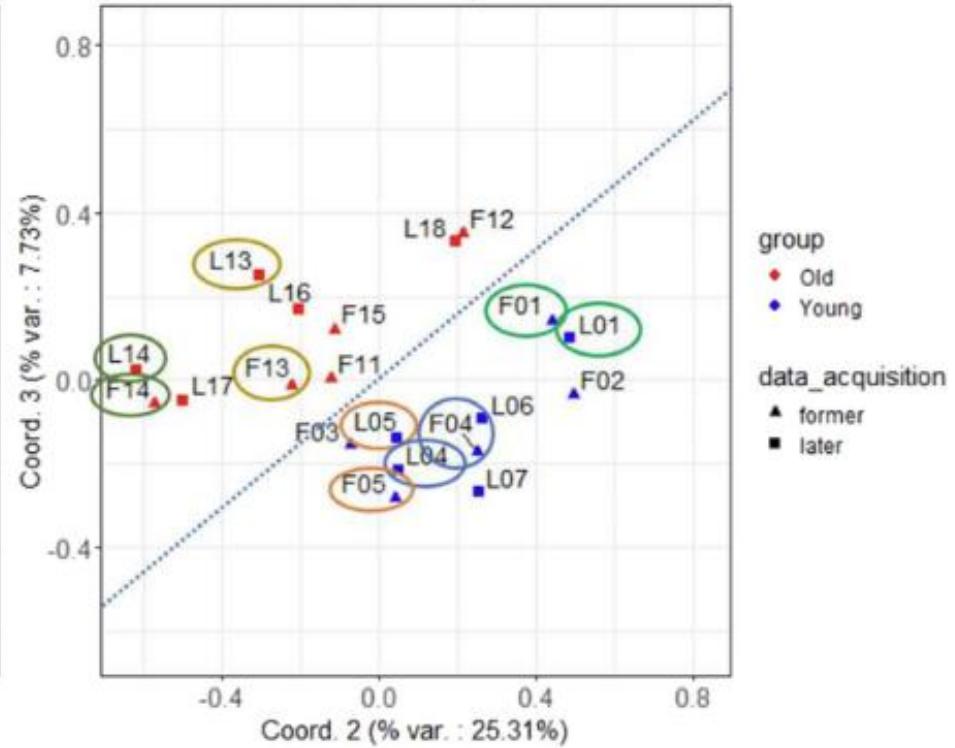
A. MDS - projection axes 1 and 2

(Pseudo R² = 0.9761; nDim = 3)



B. MDS - projection axes 2 and 3

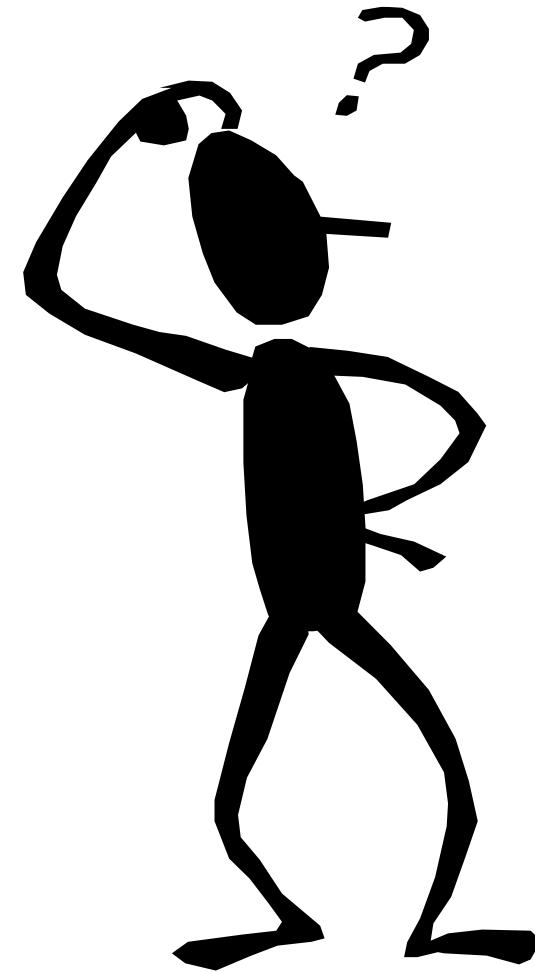
(Pseudo R² = 0.9761; nDim = 3)

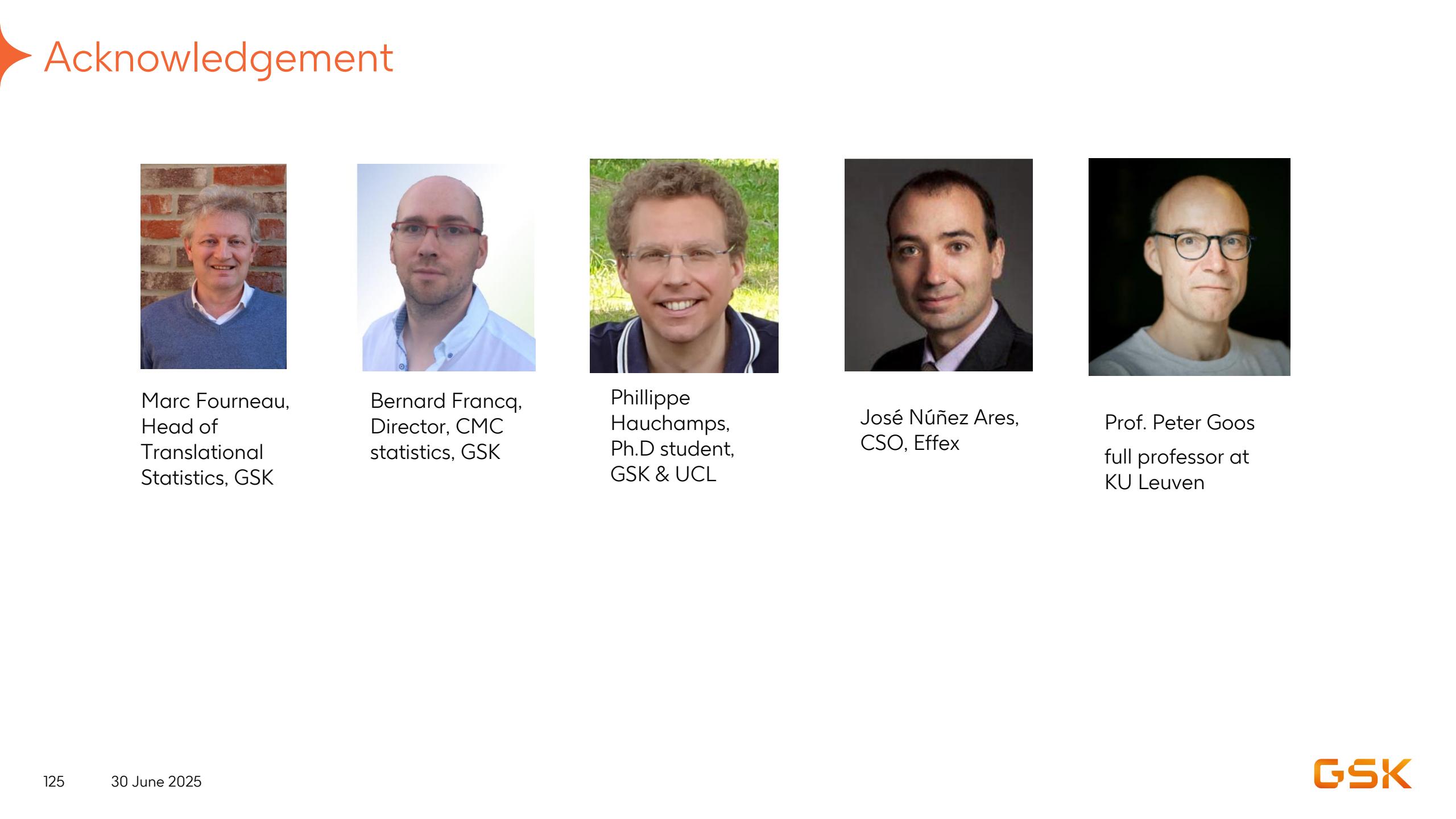


Hauchamps et al. 2025

Hauchamps, P., Delandre, S., Temmerman, S. T., Lin, D., & Gatto, L. (2025).
Visual Quality Control With CytoMDS, a Bioconductor Package for Low Dimensional Representation of Cytometry Sample Distances. *Cytometry Part A*, 2025 Mar;107(3):177-186
<https://pubmed.ncbi.nlm.nih.gov/40035132/>

Hauchamps, Vicini, Gatto, MDSvis, a Shiny application for interactive visualisation of Multi Dimensional Scaling projections, arXiv preprint.





gsk