A Multi-Stage Calibration Framework for Digital Twins to Simulate Alternative ADC Dosing Regimens in NSCLC

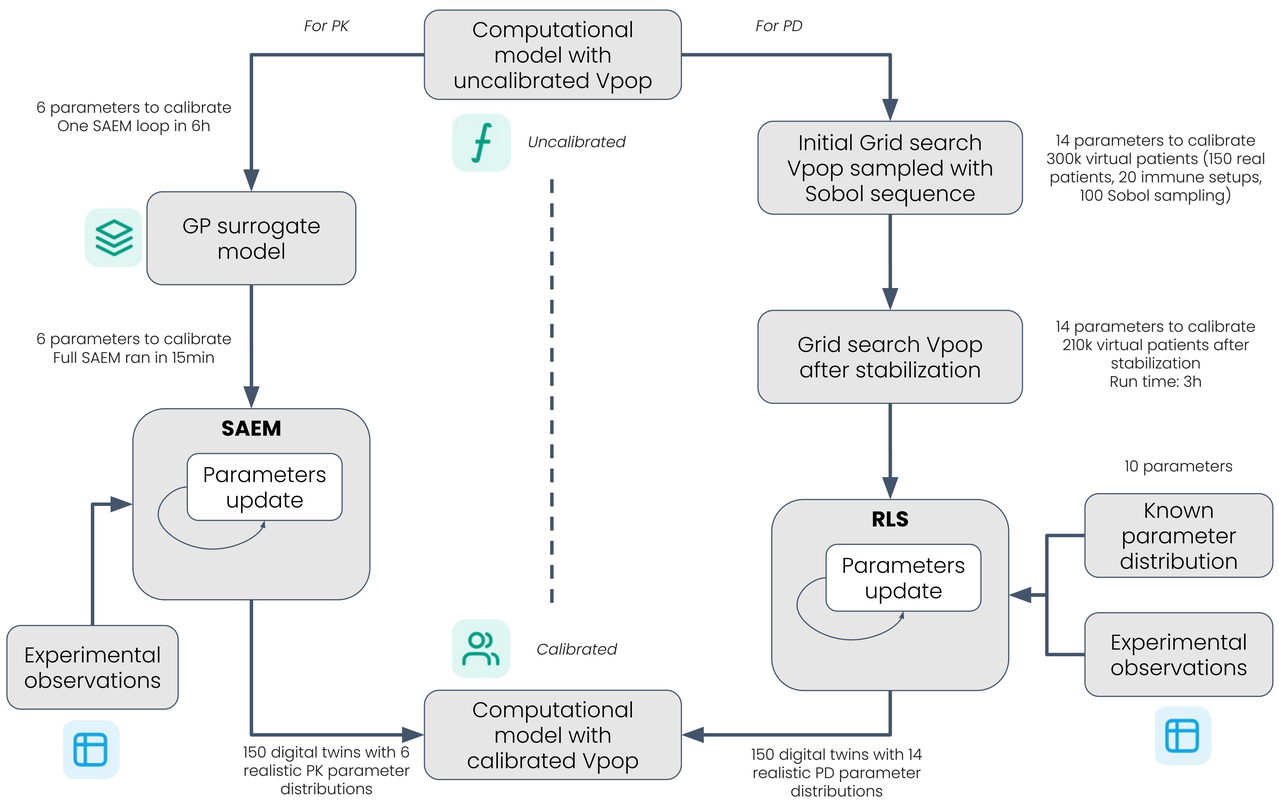
Théo Galland[1], Yann Meunier[1], Claire Couty[1], Louis Philippe[1], Sheetal Panday[2], Loveleena Bansal[2], Carolina De Miranda Silva[2], Azher Hussain[2], Anshu Marathe[2], Jim Bosley[1], Simon Arsène[1]

[1] Nova In Silico SA, Lyon, France [2] Merck & Co., Inc., Rahway, NJ, USA

A hierarchical calibration strategy makes **large QSP models estimable** — sequencing SAEM and a Sobol-informed grid search to deliver **high-accuracy digital twins** from heterogeneous patient-level data.

1 PK calibration (GP surrogate + SAEM) and PD calibration (Sobol grid search + RLS)

Population fixed effects & inter-individual variability for PK, toxicity (neutropenia) and efficacy (RECIST)



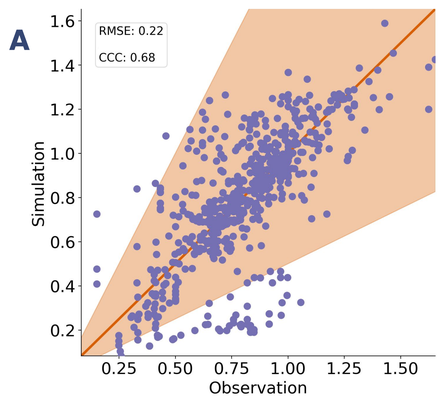
2 Digital twin generation

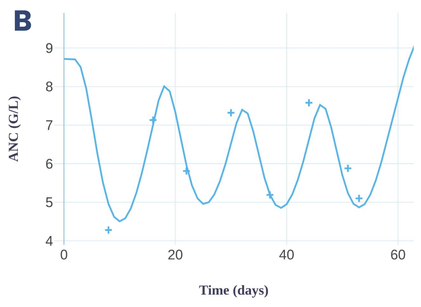
Each calibrated model instance is re-initialized with baseline biomarkers and real-world dosing history.

A: Observation vs prediction on tumor burden for the 150 calibrated patients. The underlying goal is to reproduce population heterogeneity rather than exact individual tumor burden prediction;

B: Individual digital twin fit on neutrophil count;

C: Individual digital twin fit on tumor burden.





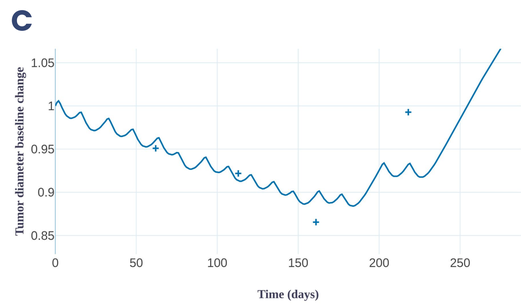


Fig. 1 — Hierarchical calibration strategy. Stage 1 estimates PK with a GP surrogate driving SAEM, while toxicity and efficacy PD are calibrated by Sobol-informed grid search plus regularized least squares under population constraints; stage 2 turns the resulting individual parameter sets into simulation-ready digital twins.

OBJECTIVE

QSP models integrate heterogeneous oncology data, but usual estimation methods fail to converge on models of this size. We set out to build a rigorous, hierarchical calibration process for an internal antibody-drug conjugate (ADC) platform integrated with established immune-oncology models [1][2], capturing the mechanistic interactions between drug exposure, hematologic toxicity and tumour burden — and ensuring individual-level accuracy while maintaining population-level constraints.

METHOD

A multi-scale QSP model couples ADC mechanisms (TMDD, bystander effect) with an immune-axis ICI framework; neutropenia is modelled semi-mechanistically, driven by free payload. Calibration on NSCLC patients uses clinical PK, neutrophil counts and RECIST data, with in vitro priors from an already-calibrated mode-of-action model. PK parameters and IIV are estimated by SAEM on a GP surrogate [3]; PD components are calibrated by Sobol-informed grid search under strict population constraints, which absorbs the patient-specific dosing schedules that block joint estimation.

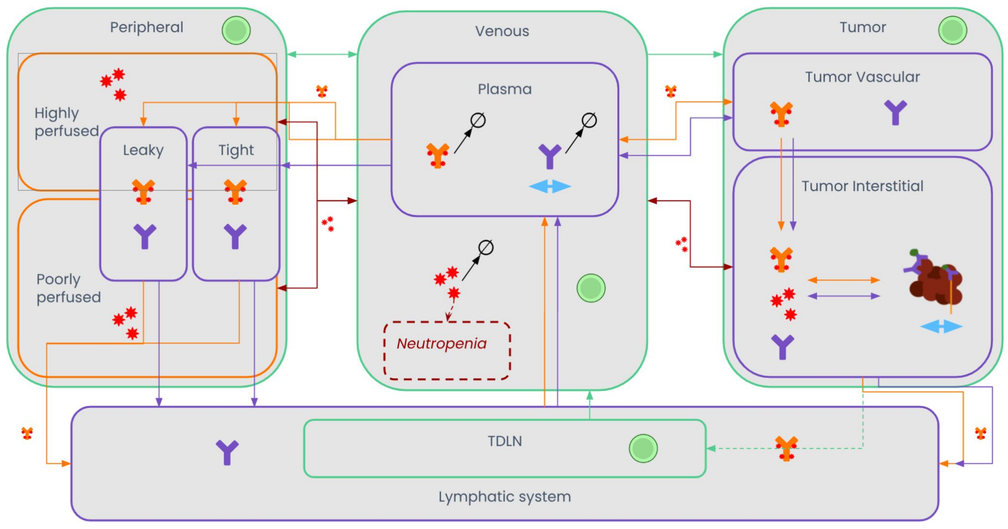


Fig. 2 – QSP model structure including both ADC and ICI dynamics across plasma, lymph and tumor site. The QSP model includes also mechanism of actions of both treatments as well as an hematotoxicity axis.

RESULTS

The step-wise calibration holds across every modelled variable

Population PK agreed with observed plasma concentrations; the toxicity module reproduced individual neutrophil nadirs and recovery kinetics, matching clinical neutropenia grade distributions; efficacy calibration reproduced individual tumour-diameter trajectories over time.

The grid search buys efficiency and identifiability at once

Sobol-informed exploration plus RLS identified stable parameter sets that stayed consistent with the initial in vitro data package while reflecting the clinical heterogeneity of advanced NSCLC patients. Resulting digital twins captured ADC and free payload clearance as well as dose-limiting neutropenia, and established mechanistic links between individualized PK, toxicity and efficacy outcomes — including patient-specific dose thresholds for transitions to severe neutropenia.

Why grid search + Regularized Least Squares (RLS) for the PD stage?

Accommodates Dosing Heterogeneity: Decouples patient-specific dosing regimens from model instantiation by scoring individual patient trajectories against a shared parameter grid.

Efficient Parameter Space Coverage: Sobol sequences sample the 4-parameter space efficiently (512 reusable configurations), replacing dense factorial sweeps.

Biologically Anchored Identifiability: RLS fits clinical trajectory data while penalizing deviations from in vitro priors to maintain biological fidelity.

High Computational Scalability: Fully parallelizable grid evaluations (~800k simulations across 18 independent batches) make cohort-scale digital twin generation practical.

CONCLUSIONS

Sequencing SAEM for pharmacokinetics and a Sobol-informed grid search with RLS for pharmacodynamics addresses the computational and identifiability challenges of calibrating a large oncology QSP model on heterogeneous data and diverse dosing regimens. Structurally separating PK and PD estimation, while enforcing population-level biological constraints, yields stable individual parameters that reconcile drug-specific kinetics with patient-specific tumour biology — and ultimately accurate digital twins suitable for in silico evaluation of alternative regimens and ADC + ICI combinations in NSCLC.

REFERENCES

[1] Wang CY et al. Development and Evaluation of a Quantitative Systems Pharmacology Model for Mechanism Interpretation and Efficacy Prediction of Atezolizumab in Combination with Carboplatin and Nab-Paclitaxel in NSCLC. *Pharmaceuticals (Basel)*. 2024;17(2):238.

[2] Li TR et al. Pivotal Dose of Pembrolizumab: A Dose-Finding Strategy for Immuno-Oncology. *Clin Pharmacol Ther*. 2021;110(1):200–209.

[3] Lemarre P et al. A Machine Learning–Driven SAEM Framework for Scalable Parameter Estimation in QSP. *QSPC 2026 poster*.