

Abstract

- The generation of virtual populations (VPops) in quantitative systems pharmacology (QSP) modeling is a commonly applied approach to calibrate models to data from patient populations.
- Consisting of a collection of simulated individuals, a virtual population is a set of model parameters that can recapitulate both the central trends and variability of data.
- While multiple methods with different algorithms exist for virtual population generation, there is limited guidance for selecting and implementing virtual population generation methods.
- A systematic comparison of four existing methods – simulated annealing, approximate Bayesian computation, maximum *a posteriori* Bayes methods, and a generative adversarial network – showed there was a modest trade-off between virtual population fidelity and computation time.

Methods

Overview

- Simulate training and testing datasets by varying two key model parameters.
- Generate VPops using four methods: maximum *a posteriori* Bayes (MAP Bayes), sequential Monte Carlo approximate Bayesian computation (SMC-ABC), generative adversarial network (GAN), and simulated annealing (Allen method).
- Compare:
 - Estimated parameter distributions for sparse and dense datasets to the known parameter distributions.
 - Simulated response landmarks to known landmark values.
 - Computational cost and implementation considerations.

Synthetic Data Generation

- Base QSP Model: MAPK/ERK signaling pathway model simulating relative tumor size at days 28 and 56 [1].
- Varied Parameters: $d_{\max}$ (baseline death rate) and w_{OR} (MAPK pathway reliance, bounded [0,1]) sampled from a multivariate normal distribution; remaining parameters fixed at medians.
- Simulated Datasets:
 - Sparse calibration: $N = 10/\text{arm}$ (typical low-sample QSP trial)
 - Dense calibration: $N = 1000/\text{arm}$ (performance upper bound)
 - Validation test set: $N = 1000/\text{arm}$ (held-out ground truth)
- Treatment Arms ($N = 4$): Placebo, MEKi mono, EGFRi + ERKi, and MEKi + ERKi.
- Observation Noise: Multiplicative exponential error ($\sigma^2 = 0.01$).

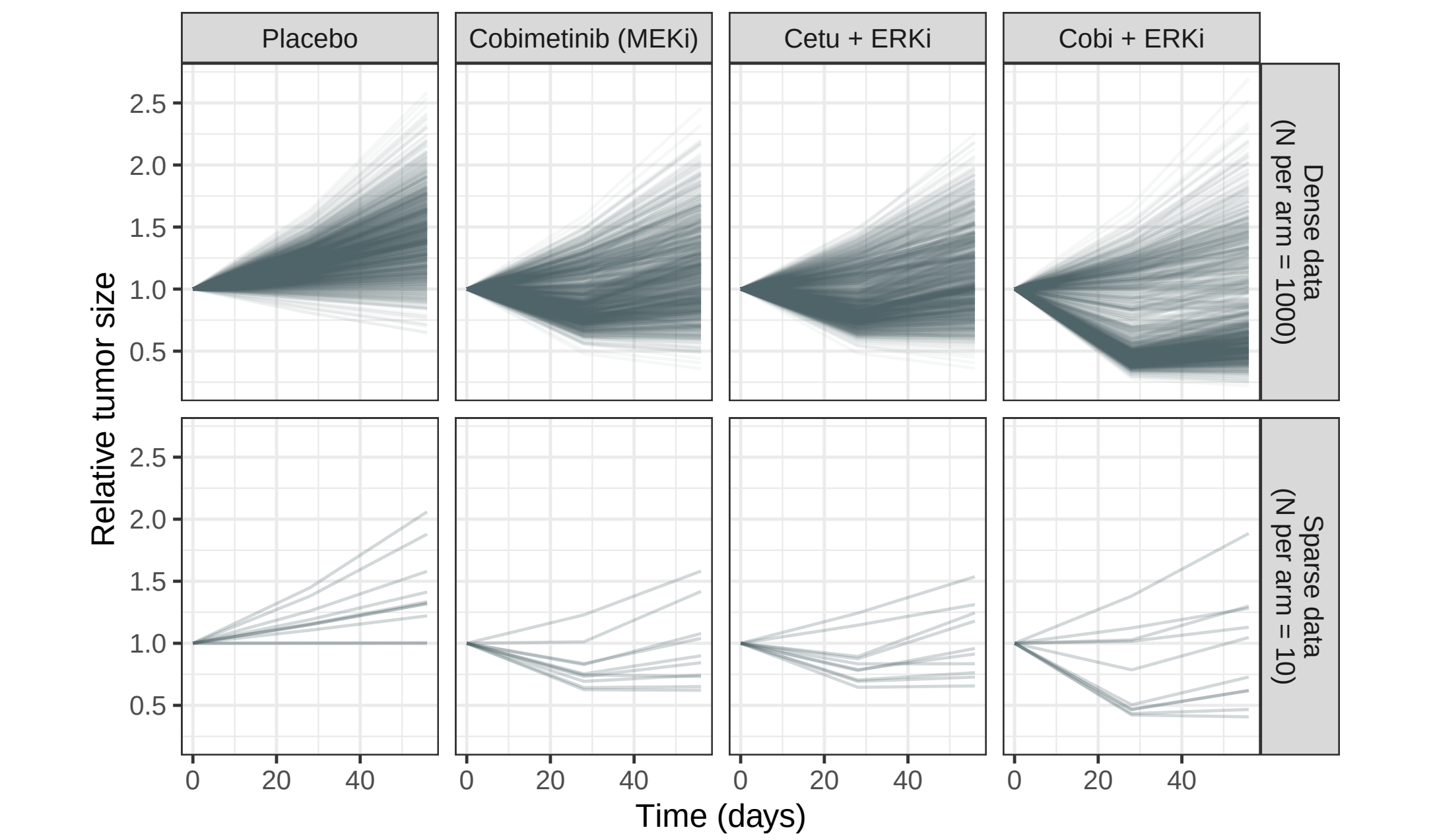


Figure 1. Simulated tumor trajectories under different treatment regimens. These simulated data were used to calibrate/fit model parameters.

VPop Generation Methods

- MAP Bayes (mapbayr in R):** Derives individual empirical Bayes estimates (EBEs) from calibration data and population estimates, then constructs VPop hyperparameters from the sample mean vector and covariance matrix of those EBEs.
- SMC-ABC (custom R Framework):** Estimates population hyperparameters (log-normal $d_{\max}$, beta w_{OR}) by adaptively sampling particles and matching simulated vs. observed 10th/50th/90th response percentiles via Euclidean distance.
- Conditional GAN (Flux.jl in Julia):** Trains a neural-network generator to produce joint parameter sets directly from noise, using a neural network ODE surrogate as a differentiable emulator to pass gradients through the QSP model [2].
- Allen Method (MATLAB translated to Julia):** Generates a plausible patient pool by minimizing output box-costs via simulated annealing (SA), then optimizes k -nearest neighbor (k -NN) density ratio inclusion probabilities to match target distributions [3].

VPop Generation and Comparison

Generated 1000 virtual patients/arm per scenario, simulated day 28 and 56 tumor responses without observation error (RUV), and calculated 1D Wasserstein-1 (W_1) distance against true test set outcomes.

- MAP Bayes Sampling: Drew individual random effects from a multivariate normal distribution $MVN(\hat{\mu}, \hat{\Sigma})$ fit to calibration EBEs, then combined with typical values.
- SMC-ABC Sampling: Resampled hyperparameter posteriors by particle weight, then drew individual $d_{\max}$ and w_{OR} values from the corresponding log-normal and beta distributions.
- GAN Sampling: Fed standard normal noise vectors through the trained generator to generate parameters transformed to native parameter scales ($d_{\max} > 0$, $w_{\text{OR}} \in [0, 1]$).
- Allen Method Sampling: Weighted plausible pool patients by k -NN inclusion probabilities, then drew individual $d_{\max}$ and w_{OR} parameter sets to match target distributions.

Results

- All methods successfully captured central tendencies.** Allen and SMC-ABC consistently captured the extreme tails of the parameter distributions. MAP Bayes and the GAN showed under-dispersion, particularly when trained on sparse data (Figure 2).
- Allen and SMC-ABC produced the most accurate tumor responses.** Consistent with their better recovery of parameter-distribution tails, Allen and SMC-ABC captured the bimodal tumor-response distributions more closely (Figure 3 and Table 1).
- Greater response-distribution fidelity was generally associated with higher computational cost.** Allen and SMC-ABC required hundreds of core-minutes for sparse and rich data scenarios, whereas MAP Bayes and GAN required less than 24 core-minutes across the evaluated scenarios (Table 1). Parallelization reduced wall-clock time for the more computationally intensive approaches (approximately 45 minutes for SMC-ABC with 16 cores).

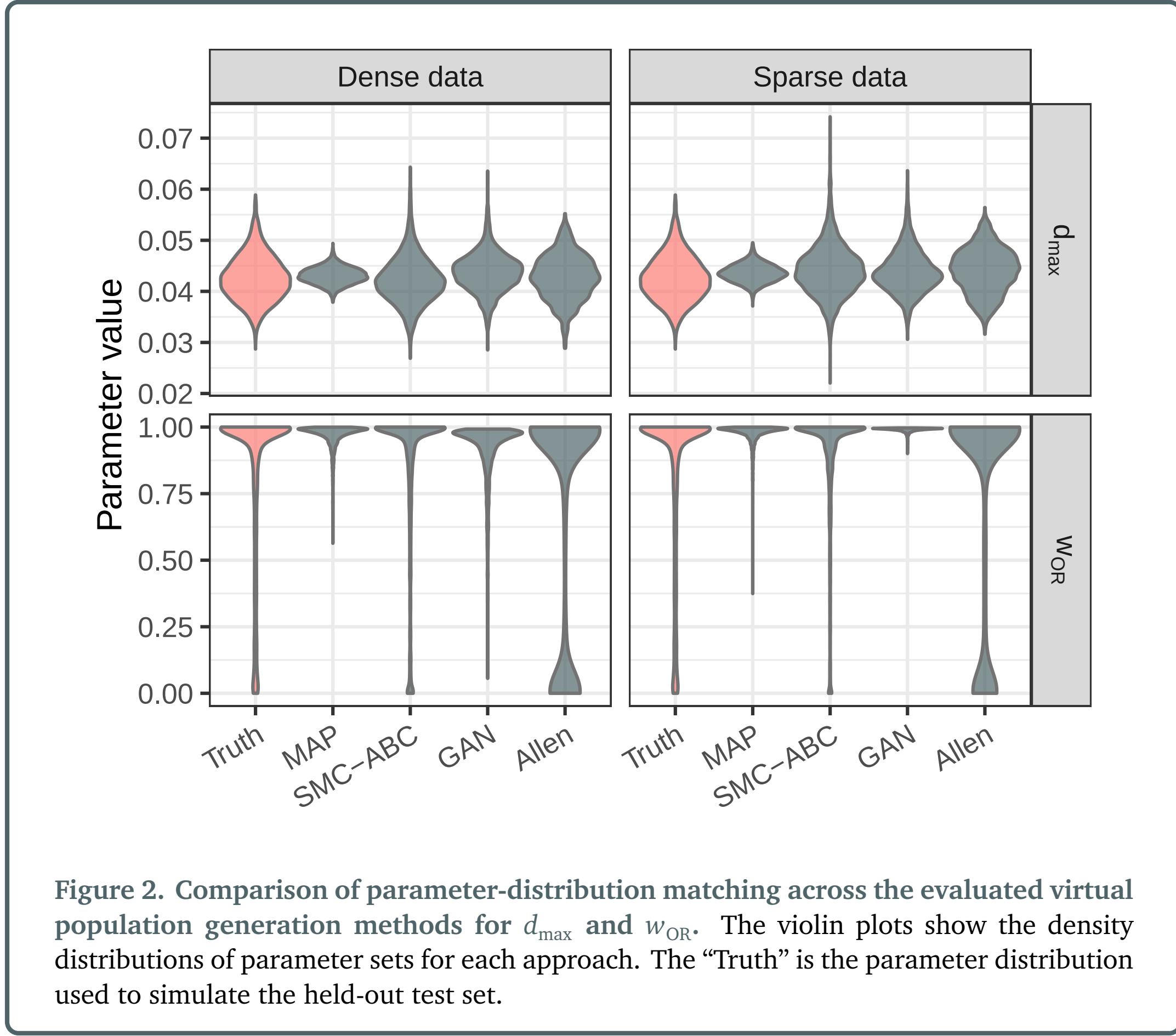


Figure 2. Comparison of parameter-distribution matching across the evaluated virtual population generation methods for $d_{\max}$ and w_{OR} . The violin plots show the density distributions of parameter sets for each approach. The “Truth” is the parameter distribution used to simulate the held-out test set.

Method	W_1 [Raw (Relative Error %)]		Compute Time (core-min)
	Day 28	Day 56	
Dense data			
MAP	0.0665 (6.98%)	0.106 (8.89%)	23.5
SMC-ABC	0.00511 (0.648%)	0.0169 (1.55%)	781
GAN	0.0441 (6.23%)	0.0506 (5.45%)	2.50
Allen	0.0304 (3.72%)	0.0455 (4.83%)	462
Sparse data			
MAP	0.0949 (10.8%)	0.147 (13.3%)	0.187
SMC-ABC	0.0397 (4.33%)	0.0831 (8.30%)	698
GAN	0.143 (18.2%)	0.201 (21.6%)	1.01
Allen	0.0395 (3.82%)	0.0939 (7.71%)	515

Table 1. Accuracy and computational cost by estimation method. W_1 is the 1D Wasserstein-1 distance; core-minutes represent total CPU processing time, distinct from wall-clock time achieved via parallelization. Lower W_1 values indicate greater similarity between the VPop-generated and known test-set response distributions.

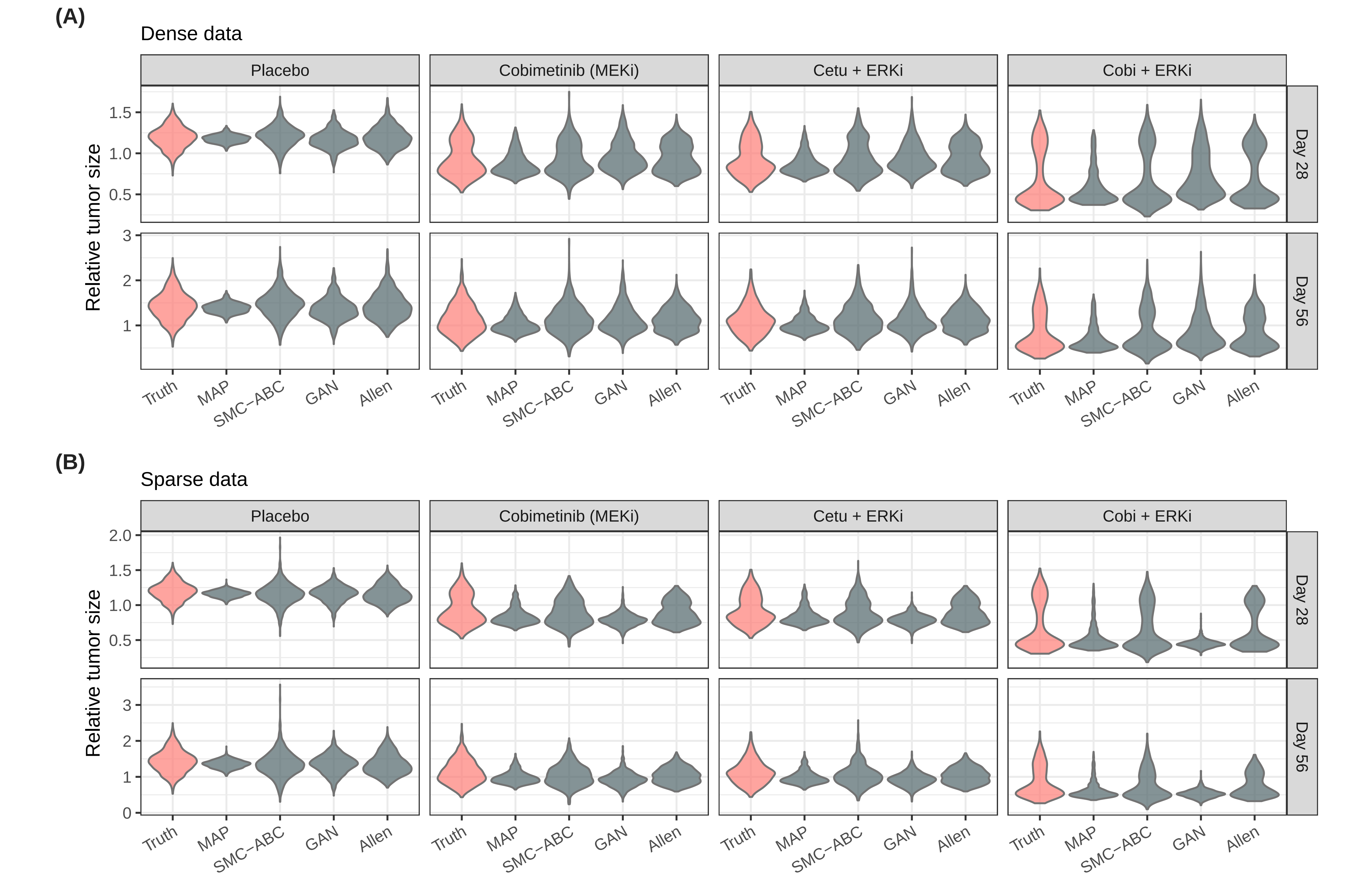


Figure 3. Comparison of relative tumor size distributions across the evaluated virtual population generation methods applied to the dense (A) and sparse (B) datasets, stratified by treatment arm and simulation day. The violin plots show the density distribution of relative tumor size for each approach. The “Truth” is the simulated outcome from the held-out test set.

Conclusions

- No single method was optimal across all evaluated criteria.** Method selection should consider calibration-data density, desired distributional fidelity, computational resources, and implementation requirements.
- Allen and SMC-ABC most closely recovered the known distributions.** Both methods performed well for parameter and tumor-response distributions, particularly under sparse calibration, but required substantially greater computational resources.
- MAP Bayes provided a computationally efficient and accessible approach.** However, it showed η -shrinkage and under-dispersion, particularly with sparse calibration data.
- GANs enabled rapid VPop resampling.** Once trained, they rapidly generated additional VPop cohorts, although they showed under-dispersion, particularly with sparse data. GANs could offer advantages as data dimensionality and observation density increase.

Metrum Research Group Publications and Posters