

# QuSOP: A QUANTUM LEARNING–OPTIMISATION FRAMEWORK FOR ROBUST DOSE SELECTION IN EARLY CLINICAL DEVELOPMENT

**Kuan-Cheng Chen<sup>1,3</sup>**   **Ross Grassie<sup>2</sup>**  
**Lily Lee<sup>3</sup>**   **Felix Burt<sup>3</sup>**   **Hasan Almatrouk<sup>3</sup>**   **Sofia Moliner Bobo<sup>3</sup>**

<sup>1</sup> Jij Europe, London, United Kingdom

<sup>2</sup> Jij Inc., Tokyo, Japan

<sup>3</sup> Imperial College London, London, UK

## ABSTRACT

We introduce **QuSOP**, a hardware-agnostic, regulator-aligned framework that couples a *quantum long short-term memory* predictor (QLSTM) with *quantum Monte Carlo* (QMC) to deliver uncertainty-aware, chance-constrained dose selection for PK/PD decision-making. QLSTM learns steady-state biomarker trajectories over 24 h and 168 h windows from sparse, irregular data, while QMC propagates covariate priors (e.g., body weight, concomitant medication) to compute attainment probabilities  $p(d) = \Pr[\max_t \text{PD}(t) \leq \tau]$  on discrete dose grids; the minimal dose is then chosen to satisfy prespecified success levels (e.g.,  $\geq 90\%$  or  $\geq 75\%$ ). Quantum resources are confined to training/estimation subroutines (e.g., variational circuits, amplitude-estimation-style samplers); inference is deterministic on CPU/GPU with full audit trails. On a Phase 1-like dataset, the QLSTM+QMC pipeline yields well-calibrated sigmoidal dose–success curves and internally consistent once-daily/once-weekly recommendations; relaxing the attainment criterion from 90% to 75% reduces the mean minimal doses by  $\sim 25\%$  for once-daily (19.0→14.3 mg) and  $\sim 13\%$  for once-weekly (116.9→101.3 mg), with analogous trends when concomitant medication is excluded. QuSOP supports covariate-shift stress tests and multi-endpoint extensions, and is immediately deployable in sovereign GPU environments while remaining compatible with emerging QPUs. By combining QLSTM’s sample-efficient temporal modeling with QMC’s calibrated uncertainty propagation, QuSOP provides a pragmatic pathway to quantum-enhanced, transparent, and clinically relevant dose policies in early development.

## 1 INTRODUCTION

The drug development pipeline is divided into five key stages, which highlight the checkpoints required to transform a single molecule from an idea in the laboratory to a product for patients. Doing so requires extensive collection and analysis of data describing the action of a molecule and the effects of the model in the human body need to be characterised. A key model to assess the safety and efficacy of a drug is PK/PD, which integrates *pharmacokinetics* (PK) and *pharmacodynamics* (PD) to describe the drug’s effect in the body over a concentration range (Negus & Banks, 2016). While highly useful for identifying a drug’s optimal dose and dosing strategy, PK/PD models are confounded by factors which reflect the individuality of physiology (Standing, 2016). That is, a person’s weight, age, genetics and other factors can all impact how the drug metabolises; this inter-individual variability represents a significant challenge in both drug discovery and the clinical setting (Uno et al., 2024). To overcome this challenge, machine learning (ML) has been used to scale PK/PD models to the popula-

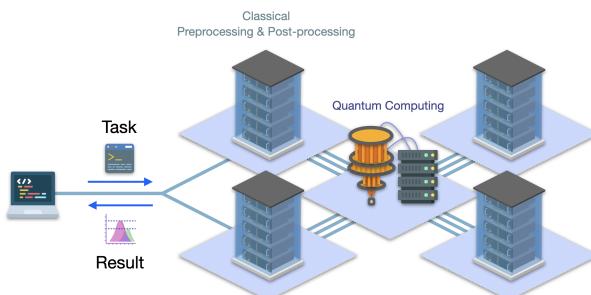


Figure 1: Quantum-centric supercomputing scheme.

tion level and help reduce the labor required to generate good models. However, many machine learning techniques suffer from sparse and noisy real-world data.

Quantum computing is poised to address the data limitations of ML, since quantum and quantum-enhanced algorithms can efficiently embed and process biological data using less input (Doga et al., 2024). As such, identifying the best algorithms and applications of quantum computing within the pharmaceutical industry is a developing field of research.

**Quantum-centric supercomputing for PK/PD.** Given the current capabilities of quantum hardware, we envision a quantum-centric supercomputing method to develop a robust PK/PD model. This approach divides the complex learning-and-decision workloads into modules that execute where they fit best – classical accelerators for high-volume data wrangling and differentiable training; quantum hardware for combinatorial search and sampling over vast state spaces offers a pragmatic route to performance and scalability. The quantum-centric supercomputing vision (Bravyi et al., 2022; Gambetta, 2022) operationalises this division of labour by tightly coupling CPUs/GPUs with quantum processors. This approach is especially attractive in the context of PK/PD model development with limited quantum resources, where we must (i) learn exposure→response dynamics from sparse, heterogeneous time series data; (ii) propagate uncertainty to population-level attainment probabilities; and (iii) optimise discrete dose policies under chance constraints.

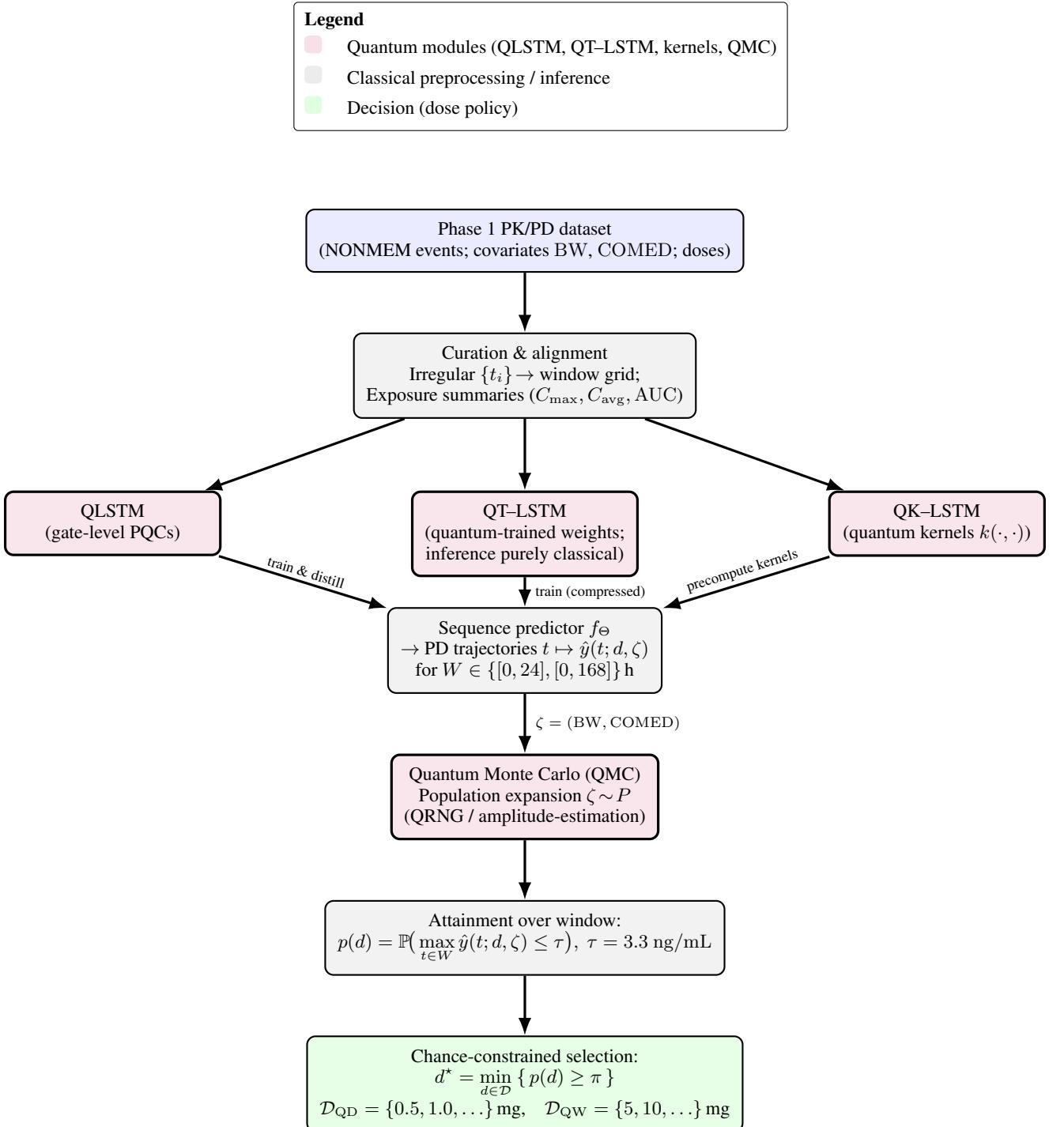
In contemporary quantum machine learning (QML), parametrised quantum circuits (PQCs) act as quantum neural networks (QNNs) (Chen et al., 2020). Data are embedded via dedicated encodings (Pérez-Salinas et al., 2020; Schuld et al., 2021), while parameters are updated on classical hardware within hybrid loops (Mari et al., 2020). Such hybridization aligns with the NISQ (noisy intermediate-scale quantum) era (Preskill, 2018) and extends to tightly coupled stacks where QNNs complement classical sequence models as pre- or post-processing blocks (Mari et al., 2020; Liu et al., 2021a). Despite promising signals of advantage in specific regimes (Cerezo et al., 2022; Huang et al., 2022; Biamonte et al., 2017; Caro et al., 2022; Huang et al., 2021), practical hurdles remain: realistic encodings can be depth- or qubit-hungry, and quantum-at-inference requirements hinder deployability when low-latency predictions are mandatory.

**Our approach: QuSOP for dose selection.** To address these constraints, we introduce **QuSOP** (Quantum Stochastic optimisation with Sequence Prediction), a quantum-centric, industry-oriented framework that couples classical/quantum sequence models with quantum (or quantum-inspired) optimisation and uncertainty quantification. We instantiate a *model family* with three interoperable variants:

- (Q1) **QLSTM**: an LSTM whose gates are augmented with PQCs for increased expressivity (Chen et al., 2022a);
- (Q2) **QT-LSTM**: quantum modules synthesise subsets of LSTM weights *during training only*, keeping inference purely classical (Liu et al., 2024a; Lin et al., 2024; Liu et al., 2024b);
- (Q3) **QK-LSTM**: quantum kernels provide geometry-aware embeddings that are precomputed and amortized across training (Hsu et al., 2025c;a;b).

In all cases the predictor’s inference path is classical and low-latency; quantum resources are reserved for the subroutines where they offer the highest leverage—expressive gating, compact parameter generation at train time, or geometry-aware similarity. This design directly mitigates expensive encodings and removes strict dependence on quantum hardware at inference (Cerezo et al., 2022; Huang et al., 2022). While many QML stacks target cryogenic gate-based devices, our formulation is hardware-agnostic: PQCs can run on superconducting, trapped-ion, silicon, or photonic backends; kernel estimation can be performed on simulators or QPUs and reused.

Integrated within QuSOP, learned PK→PD maps produce subject-level biomarker trajectories; Monte Carlo sampling (quantum or classical) converts these to attainment probabilities (e.g., probability that  $PD(t) \leq$  a clinically relevant threshold over 24–168 h at steady state); and a discrete optimiser (e.g., QAOA/annealing or classical integer baselines) selects once-daily or once-weekly doses that satisfy population-level targets (e.g.,  $\geq 90\%$  or  $\geq 75\%$  success) under covariate shifts (e.g., body-weight distributions) and policy constraints (e.g., no concomitant medication). Figure 2 summarises the overall scheme.



**Figure 2: QuSOP pipeline.** End-to-end flow combining three quantum-enhanced sequence variants (QLSTM, QT-LSTM, QK-LSTM) with quantum Monte Carlo for uncertainty-aware attainment estimation and a discrete chance-constrained optimiser for dose selection. Quantum modules (purple) are confined to training/estimation; inference is deterministic on CPU/GPU.

## Contributions.

- **Data-efficient, deployment-ready sequence prediction.** Three interoperable variants (QLSTM, QT-LSTM, QK-LSTM) maintain a classical inference path while leveraging quantum resources for expressivity, compact parameter generation, or kernel geometry.
- **Uncertainty-aware, chance-constrained dose optimisation.** We estimate suppression probabilities over full dosing intervals (24–168 h) and solve discrete dose selection to meet specified attainment targets (e.g.,  $\geq 90\%$  or  $\geq 75\%$ ).
- **Robustness to covariate shifts and constraints.** The pipeline quantifies sensitivity to body-weight distributions and to restrictions such as prohibition of concomitant medication, supporting transparent scenario analyses for clinical and regulatory decision-making.
- **Hardware-agnostic, NISQ-compatible realisation.** All components run today on CPUs/GPUs with quantum simulators; optional QPUs or photonic backends can be used for PQC or kernel steps without changing the deployed inference interface.

## 2 TEAM MEMBERS INTRODUCTION:

Our team combines quantum algorithms, compilers, AI/ML, pharmacometrics-facing validation, HPC engineering, and programme delivery—a profile tailored to this PK/PD challenge.

- **Dr Louis Chen (Algorithm & Pipeline Lead)** — Quantum computing researcher at Jij Europe; M.Sc. & Ph.D. from Imperial (QuEST). Winner of the 2024 Deloitte Quantum Challenge, 2<sup>nd</sup> place in the Xanadu Software Competition. IEEE QCE 2025 Best Paper (Application & Photonics) plus Distinguished Technical Paper Award. Focused on distributed QC, (quantum) ML/AI, and (quantum) optimisation; leads algorithm/pipeline design and rigorous SOTA benchmarking.
- **Dr Ross Grassie (Programme & Engineering Lead)** — Global Presales Lead at Jij Inc. Bridges product, research, and enterprise stakeholders to scope, architect, and deliver high-impact quantum/AI solutions. Owns work-plan structure, technical scoping, risk/resource management, and delivery governance; ensures clear success criteria, reproducibility, and stakeholder alignment from PoC to scale-up.
- **Lily Lee (Domain & Validation Lead)** — 3<sup>rd</sup>-year Ph.D. at Imperial in quantum computing for pharmaceutical problems and simulation, with prior pharma internship experience. Provides PK/PD domain grounding, experimental design, and statistical analysis of clinical-style data. Co-develops and validates models, ensuring results are interpretable and reported to pharmacometric expectations. Winner of the 2024 Deloitte Quantum Climate Challenge.
- **Felix Burt (Scalability, Simulation & Resource Modelling)** — 3<sup>rd</sup>-year Ph.D. at Imperial researching distributed quantum computing. Specialises in quantum circuit design for QML, optimisation and error management. Implements proof-of-principle quantum amplitude estimation for Monte Carlo. Winner of the 2024 Deloitte Quantum Climate Challenge and best paper award in IEEE QCE 2025 photonics track.
- **Hasan Almatrouk (HPC & MLOps)** — 1<sup>st</sup>-year Ph.D. at Imperial focused on quantum applications with strong systems integration skills. Builds reproducible HPC pipelines (containerisation, scheduling, experiment tracking) and performance-tunes ODE/simulation and hybrid QML workloads. Supports clean comparative reporting and articulates operational impact and future work directions.
- **Sofia Moliner Bobo (Quantum algorithms)** — 2<sup>nd</sup>-year Ph.D. at Imperial researching quantum algorithms for financial applications. Strong background in partial differential equations, stochastic optimisation and Monte Carlo simulation. Supports implementation of quantum Monte Carlo and amplitude estimation.

### 3 PRELIMINARIES: SEQUENCE MODELS FOR PK/PD AND THEIR QUANTUM EXTENSIONS

We model a dosing window as a length- $T$  sequence of covariates and responses. For once-daily regimens ( $\tau_{\text{qd}} = 24$  h) we use per-time-step inputs  $x_t = (t/\tau_{\text{qd}}, d/\text{BW}, \text{COMED}) \in \mathbb{R}^3$  with daily dose  $d$  (mg), body weight BW (kg), and COMED  $\in \{0, 1\}$ ; the output is the biomarker  $y_t$  (ng/mL). The learning task is to fit a sequence predictor  $f_\Theta : \mathbb{R}^{T \times 3} \rightarrow \mathbb{R}^T$  that maps  $\{x_t\}$  to the entire PD trajectory  $\{\hat{y}_t\}$ ; inverse scaling returns physical units for downstream chance-constrained dose optimisation in QuSOP.

#### 3.1 LONG SHORT-TERM MEMORY (LSTM)

LSTMs [Hochreiter & Schmidhuber \(1997\)](#) mitigate vanishing/exploding gradients in vanilla RNNs by introducing gated state updates (Fig. 3). Given input  $x_t$ , hidden  $h_{t-1}$ , and cell  $C_{t-1}$ , the forget, input, and output gates and the candidate update are

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f), \quad i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \\ \tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c), \quad o_t = \sigma(W_o[h_{t-1}, x_t] + b_o),$$

with

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \quad h_t = o_t \odot \tanh(C_t), \quad \hat{y}_t = g(h_t). \quad (1)$$

We train with mean-squared error on standardised  $y_t$  using BPTT and early stopping; at inference, inputs are normalised as in training and predictions are inverse-scaled to ng/mL. This baseline already captures nonlinear PK→PD mappings while remaining low-latency and deployment-ready [Liu et al. \(2021b\)](#).

#### 3.2 QUANTUM LSTM (QLSTM)

To enrich the gating nonlinearity, QLSTM replaces selected affine gates with variational quantum circuits (VQCs), keeping the LSTM state update intact [Chen et al. \(2022b\)](#). With  $v_t = [h_{t-1}, x_t]$ ,

$$f_t = \sigma(\text{VQC}_f(v_t)), \quad i_t = \sigma(\text{VQC}_i(v_t)), \quad (2a)$$

$$\tilde{C}_t = \tanh(\text{VQC}_c(v_t)), \quad o_t = \sigma(\text{VQC}_o(v_t)), \quad (2b)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \quad h_t = o_t \odot \tanh(C_t). \quad (2c)$$

Each VQC<sub>•</sub> comprises data encoding  $U(v_t)$  and a trainable block  $V(\theta)$  (single-qubit rotations and entanglers), with measurements yielding gate activations; gradients are obtained via the parameter-shift rule and combined with BPTT. Figure 4 sketches the cell and a representative VQC.

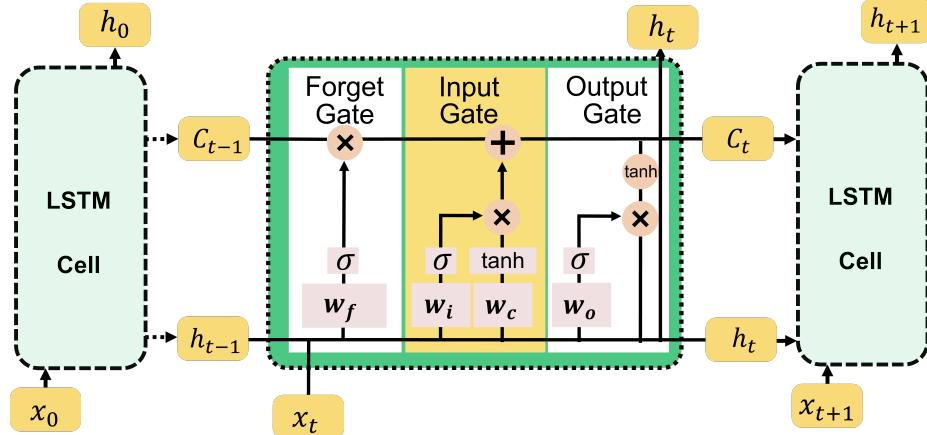


Figure 3: Standard LSTM cell. Input, forget, and output gates regulate the cell state  $C_t$ , enabling long- and short-term dependency tracking.

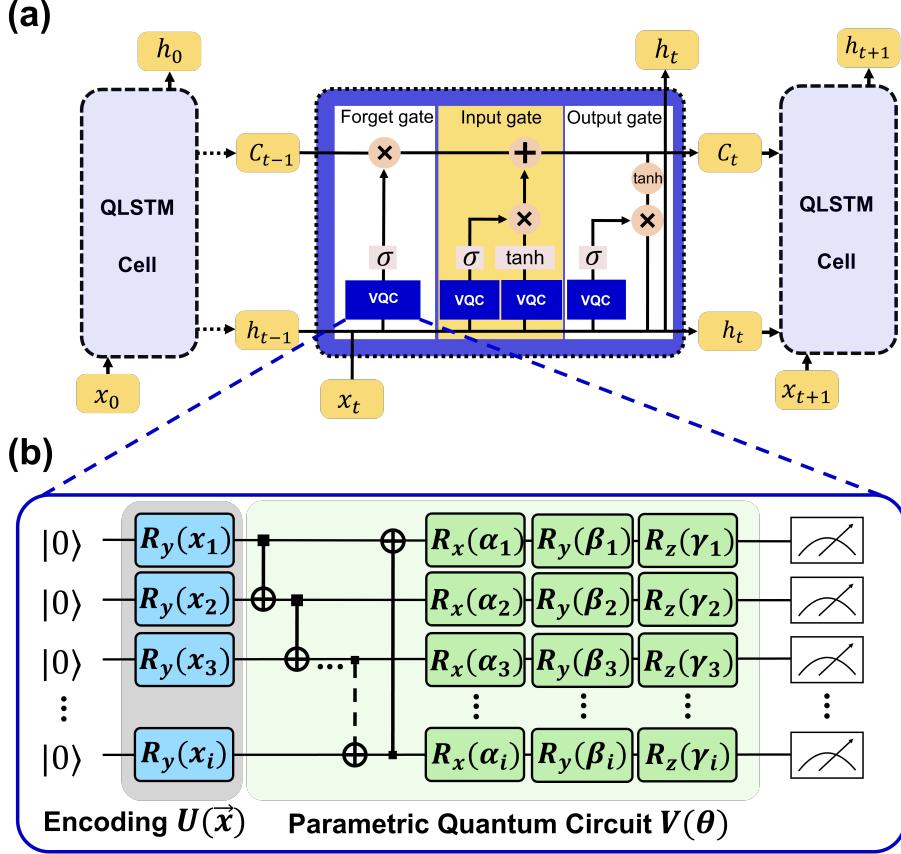


Figure 4: QLSTM: classical gate affine maps replaced by variational quantum circuits. (Left) QLSTM cell. (Right) VQC with data encoding  $U(\vec{x})$  and trainable block  $V(\vec{\theta})$ ; measurements feed gate activations.

### 3.3 QUANTUM-TRAIN LSTM (QT-LSTM)

QT-LSTM uses quantum resources only during *training* to synthesise a large set of classical LSTM weights from a compact PQC, keeping inference purely classical and fast(Liu et al., 2024a; Lin et al., 2024; Liu et al., 2024b). Let the LSTM have parameter vector  $\theta = (\theta_1, \dots, \theta_M)$  and let  $|\psi(\varphi)\rangle$  be an  $N$ -qubit state with  $N = \lceil \log_2 M \rceil$ . Measurement probabilities  $p_i = |\langle i|\psi \rangle|^2$  are mapped to weights via a small classical network  $M_\gamma$ :

$$\theta_i = M_\gamma(\text{bin}(i), p_i), \quad i = 1, \dots, M. \quad (3)$$

The PQC employs  $U_3(\mu, \phi, \lambda)$  and controlled- $U_3$  gates to build  $|\psi(\varphi)\rangle$  with parameter count  $O(\text{polylog } M)$ . Training minimises the same MSE loss,

$$\ell_{\text{MSE}} = \frac{1}{N_d} \sum_{n=1}^{N_d} (y_n - \hat{y}_n)^2, \quad (4)$$

while updating  $\varphi$  and  $\gamma$  (analytic or parameter-shift gradients in simulation). After convergence, the induced  $\theta$  is frozen and inference proceeds on classical hardware only.

### 3.4 QUANTUM-KERNEL LSTM (QK-LSTM)

QK-LSTM injects quantum feature geometry into the gates by replacing  $W[h_{t-1}, x_t] + b$  with kernel expansions evaluated in a quantum Hilbert space(Hsu et al., 2025c;a;b). For  $v_t = [h_{t-1}, x_t]$  and a reference set

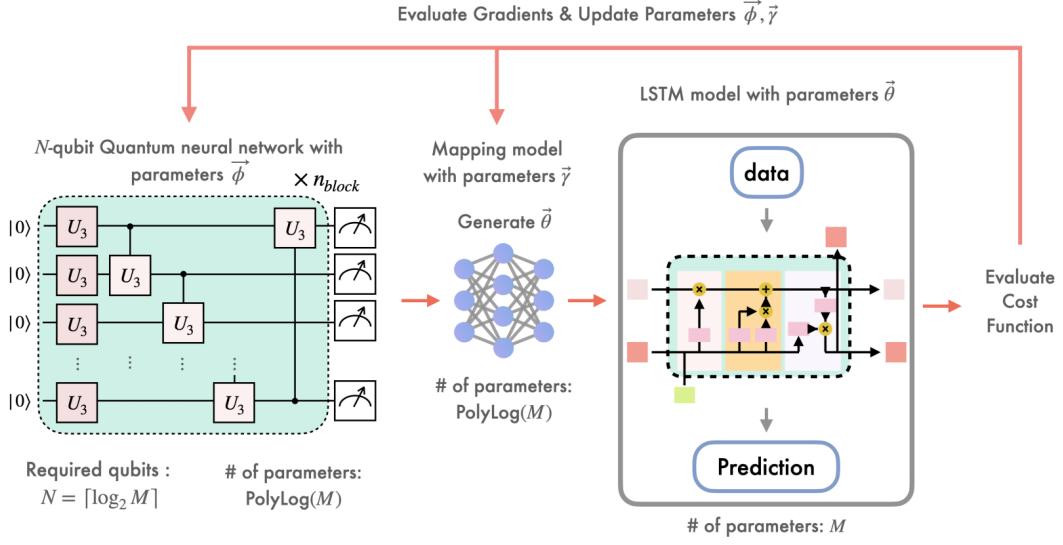


Figure 5: Quantum-Train LSTM (QT-LSTM) workflow. A variational quantum circuit generates probability distributions that are mapped by a compact classical network into full LSTM weights. The resulting LSTM is trained on data, with gradients propagated back to both quantum and classical components. After training, weights are fixed and inference proceeds entirely classically, enabling quantum-assisted training with minimal quantum resources.

$$\{v_j\}_{j=1}^N,$$

$$f_t = \sigma \left( \sum_{j=1}^N \alpha_j^{(f)} k^{(f)}(v_t, v_j) + b_f \right), \quad i_t = \sigma \left( \sum_{j=1}^N \alpha_j^{(i)} k^{(i)}(v_t, v_j) + b_i \right),$$

$$\tilde{C}_t = \tanh \left( \sum_{j=1}^N \alpha_j^{(c)} k^{(c)}(v_t, v_j) + b_c \right), \quad o_t = \sigma \left( \sum_{j=1}^N \alpha_j^{(o)} k^{(o)}(v_t, v_j) + b_o \right),$$

with the standard cell update. Each kernel is an overlap in a quantum feature space,

$$k(v, v') = |\langle \phi(v) | \phi(v') \rangle|^2, \quad |\phi(v)\rangle = U(v) |0\rangle^{\otimes n}, \quad (5)$$

where  $U(v) = U_{\text{ent}} U_{\text{enc}}(v) H^{\otimes n}$  uses Hadamards, data encoders  $R_y, R_z$  tied to  $v$ , and a CNOT entangling pattern. Kernels can be precomputed (simulator or hardware) and amortized; when circuit parameters are learnable, parameter-shift gradients apply. Figure 6 shows the gate-wise integration and the kernel circuit.

### 3.5 TRAINING, GRADIENTS, AND DEPLOYMENT

All variants minimise a supervised sequence loss on standardised responses,

$$\mathcal{L}(\Theta) = \frac{1}{T|\mathcal{D}|} \sum_{(X, Y) \in \mathcal{D}} \sum_{t=1}^T (s(Y_t) - \hat{Y}_t(X; \Theta))^2, \quad (6)$$

with BPTT for classical parameters and the parameter-shift rule for quantum ones. QLSTM trains quantum gates inside the cell; QT-LSTM trains a compact PQC that synthesises *classical* LSTM weights; QK-LSTM trains kernel weights (and optionally circuit parameters) with kernels precomputed or evaluated on demand. Inference can be purely classical for all three (QLSTM by freezing measured activations, QT-LSTM by design, QK-LSTM with cached kernels), ensuring low-latency PK/PD prediction.

**Relevance to QuSOP and PK/PD dose finding.** Given covariate draws (BW, COMED) and a candidate dose, the predictor  $f_\Theta$  produces  $\hat{y}(t)$  over  $t \in [0, 24]$  or  $[0, 168]$ ; inverse scaling returns ng/mL. Monte-Carlo over the covariate distribution yields  $p(d) = \mathbb{P}(\max_t \hat{y}(t; d) \leq \tau)$  with threshold  $\tau = 3.3 \text{ ng/mL}$ .

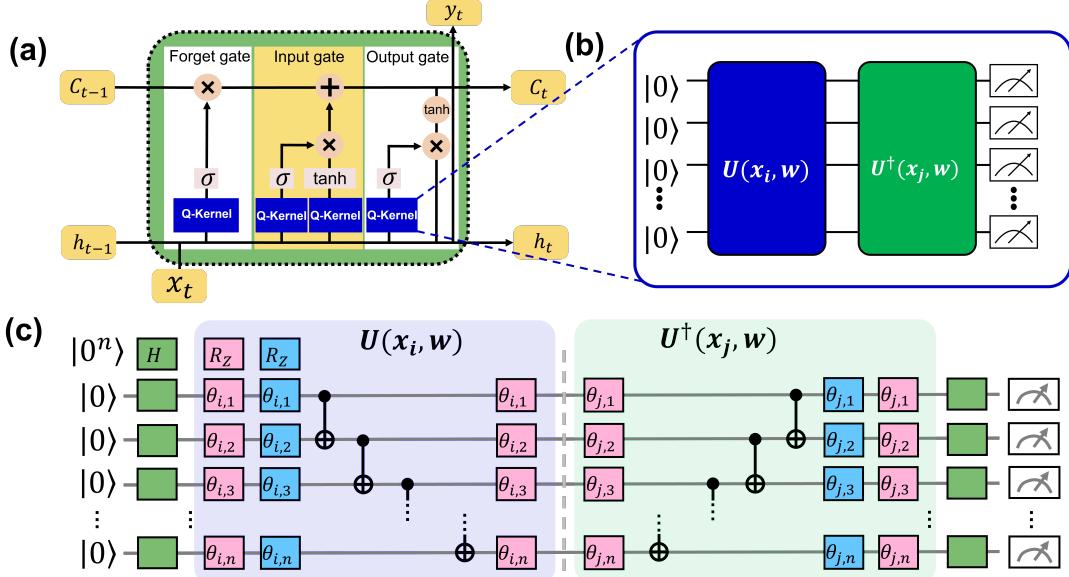


Figure 6: QK-LSTM overview. (a) Each LSTM gate is driven by a quantum kernel expansion. (b) Unitary representation  $U(x, w)$  for feature mapping and  $U^\dagger(x', w)$  for overlap. (c) Kernel evaluation circuit; varying encoders/entanglers realises families of  $k(\cdot, \cdot)$ .

Chance-constrained selection then returns the minimal discrete dose (0.5 mg daily or 5 mg weekly) achieving  $p(d) \geq \alpha$  (e.g.  $\alpha = 0.90$  or  $0.75$ ). Quantum modules enhance expressivity (QLSTM), compress parameterization at train time (QT-LSTM), or inject quantum feature geometry (QK-LSTM) while remaining fully compatible with the downstream Monte-Carlo estimator and discrete optimisation in QuSOP.

#### 4 THEORETICAL FRAMEWORK: QU SOP

**Data and task.** For each subject  $i \in \{1, \dots, N\}$  let  $BW_i \in \mathbb{R}_+$  (body weight),  $COMED_i \in \{0, 1\}$  (concomitant medication), and a dosing record with nominal once-daily dose  $D_i$  (mg) and interdose interval  $\tau_{qd} = 24$  h. Let  $y_i(t)$  denote the PD biomarker (ng/mL) measured at observation times  $t \in \mathcal{T}_i$  within the last dosing window prior to the final visit. Define the per-subject resampled trajectory on a fixed grid  $\mathcal{G}_{qd} = \{t_k\}_{k=1}^T$ ,  $T = 24$ , via linear interpolation of the last window:

$$\tilde{y}_i(t_k) = \text{Interp}(\{(t, y_i(t)) : t \in \mathcal{T}_i\} \rightarrow t_k), \quad t_k \in \mathcal{G}_{qd}.$$

The associated time-varying covariates used by the predictor are

$$x_i(t_k) = (t_k/\tau_{qd}, D_i/BW_i, COMED_i) \in \mathbb{R}^3,$$

where  $D_i/BW_i$  is in mg/kg. Denote  $X_i = \{x_i(t_k)\}_{k=1}^T$  and  $Y_i = \{\tilde{y}_i(t_k)\}_{k=1}^T$ .

**Sequence model (classical).** QuSOP learns a data-driven map  $f_\theta : \mathbb{R}^{T \times 3} \rightarrow \mathbb{R}^T$  realised by a bidirectional LSTM followed by time-distributed dense layers:

$$\hat{Y}_i = f_\theta(X_i) \in \mathbb{R}^T.$$

Let  $s(\cdot)$  be a standardisation operator learned on training responses (z-scaling) and  $s^{-1}$  its inverse. The training objective on the training set  $\mathcal{I}_{tr}$  is mean squared error in standardised space, with early stopping:

$$\mathcal{L}(\theta) = \frac{1}{\sum_{i \in \mathcal{I}_{tr}} T} \sum_{i \in \mathcal{I}_{tr}} \sum_{k=1}^T \left( s(\tilde{y}_i(t_k)) - (f_\theta(X_i))_k \right)^2.$$

At inference, predicted biomarkers in physical units are  $\hat{y}_i(t_k) = s^{-1}((f_\theta(X_i))_k)$ .

**Population simulator and chance constraint.** For a candidate daily dose  $d$  (mg) and a target population distribution over covariates  $P(\text{BW}, \text{COMED})$ , define the simulated PD path for an individual draw  $\zeta = (\text{BW}, \text{COMED})$ :

$$\hat{y}(t; d, \zeta) = s^{-1} \left( f_\theta \left( \{t/\tau_{\text{qd}}, d/\text{BW}, \text{COMED}\}_{t \in \mathcal{G}_{\text{qd}}} \right) \right).$$

Let the clinical success event over the window be

$$\mathcal{S}(d, \zeta) = \left\{ \max_{t \in \mathcal{G}_{\text{qd}}} \hat{y}(t; d, \zeta) \leq \tau_{\text{PD}} \right\}, \quad \tau_{\text{PD}} = 3.3 \text{ ng/mL}.$$

The population success probability is

$$p(d) = \mathbb{P}_{\zeta \sim P}(\mathcal{S}(d, \zeta)).$$

Our dose-selection problem is a discrete chance-constrained program

$$\min_{d \in \mathcal{D}_{\text{qd}}} d \quad \text{s.t.} \quad p(d) \geq \pi, \quad \mathcal{D}_{\text{qd}} = \{0.5, 1.0, \dots, 40.0\} \text{ mg}, \quad \pi \in \{0.90, 0.75\}.$$

**Monte Carlo estimator and uncertainty.** Given i.i.d. draws  $\zeta^{(1)}, \dots, \zeta^{(M)} \sim P$ , the unbiased estimator of success at dose  $d$  is

$$\hat{p}_M(d) = \frac{1}{M} \sum_{m=1}^M \mathbf{1}\{\mathcal{S}(d, \zeta^{(m)})\}.$$

In practice,  $P$  is generated by (i) sampling base subjects with replacement to capture design heterogeneity, (ii) applying a log-normal random walk to BW (geometric coefficient of variation  $\sigma_{\ln} = \text{SIGMA\_BW}$ , clipped to [35, 120] kg), and (iii) sampling COMED  $\sim \text{Bernoulli}(p_{\text{comed}})$  from the observed frequency. Pointwise 95% bands are obtained either by the empirical quantiles across  $R$  Monte-Carlo replicates or by Clopper–Pearson intervals for Bernoulli proportions.

**Generalization to once-weekly dosing.** For weekly regimens with interdose interval  $\tau_{\text{qw}} = 168$  h and weekly dose  $d_{\text{wk}}$  (mg), we evaluate the same trained  $f_\theta$  on a weekly time grid  $\mathcal{G}_{\text{qw}} = \{t_\ell\}_{\ell=1}^{168}$  by re-normalising time and dose per kg:

$$\hat{y}(t_\ell; d_{\text{wk}}, \zeta) = s^{-1} \left( f_\theta \left( \{t_\ell/\tau_{\text{qw}}, d_{\text{wk}}/\text{BW}, \text{COMED}\}_{\ell=1}^{168} \right) \right).$$

The weekly chance constraint mirrors the daily case with  $\max_{t \in \mathcal{G}_{\text{qw}}}$  and a discrete grid  $\mathcal{D}_{\text{qw}} = \{5, 10, \dots, 400\}$  mg.

**Modeling assumptions and identifiability.** The LSTM is a universal approximator for bounded-length sequences; we assume the 24h and 168h windows are sufficiently stationary that the map  $(t, d/\text{BW}, \text{COMED}) \mapsto y(t)$  is learnable from Phase-1 data within the studied dose range. Resampling to a fixed last-window grid controls for nonuniform sampling; standardisation  $s(\cdot)$  ensures well-conditioned training. Dose generalization outside the training support is handled by chance-constrained evaluation rather than extrapolative point prediction.

**optimisation view.** Let  $\mathcal{D}$  be the discrete dose grid (daily or weekly). Define the feasible set

$$\mathcal{F}(\pi) = \{d \in \mathcal{D} : p(d) \geq \pi\}.$$

The recommended dose is the *minimal feasible dose*  $d^*(\pi) = \min \mathcal{F}(\pi)$ , with  $\hat{d}^*(\pi)$  obtained by replacing  $p(d)$  with  $\hat{p}_M(d)$  and scanning  $d \in \mathcal{D}$ . Uncertainty on  $d^*$  is summarised by the empirical distribution of the minimal feasible dose across Monte-Carlo replicates.

**Quantum hooks in QuSOP (optional modules).** While the present study trains  $f_\theta$  and evaluates  $p(d)$  classically, QuSOP permits quantum accelerators at three points:

- (Q1) **Quantum gate augmentation (QLSTM).** Replace LSTM gates  $g = \sigma(Wx + Uh + b)$  by  $g = \sigma(Wx + Uh + b + \mathcal{Q}_\phi(x))$ , where  $\mathcal{Q}_\phi$  is a parametrized quantum circuit (PQC) embedding. Training remains hybrid; inference is classical if  $\mathcal{Q}_\phi$  is distilled at train time (QT-LSTM).
- (Q2) **Quantum kernels (QK-LSTM).** Precompute a quantum kernel  $K_\phi(x, x') = |\langle \psi_\phi(x) | \psi_\phi(x') \rangle|^2$  to define geometry-aware features  $z = K_\phi(x, \cdot)$  that feed a classical LSTM, amortising quantum cost.

- (Q3) **Quantum stochastic optimisation.** The discrete dose search can be cast as a constrained combinatorial selection. Variants include (i) quantum amplitude estimation to estimate  $p(d)$  with lower sample complexity, and (ii) QAOA/annealing to minimise  $d$  subject to a penalty  $\lambda \max\{0, \pi - p(d)\}$  on a binary encoding of  $\mathcal{D}$ .

These modules preserve the mathematical definition of  $p(d)$  and the chance constraint while potentially improving sample or search complexity.

**Evaluation metrics.** Goodness-of-fit is reported as train/validation MSE in standardised units and calibration of  $\hat{p}_M(d)$  against held-out subjects. Decision quality is evaluated by the minimal feasible dose  $\hat{d}^*(\pi)$  and its Monte-Carlo confidence interval, separately for once-daily and once-weekly regimens and for covariate shifts (e.g., modified BW or COMED distributions).

**Summary.** QuSOP formalizes dose recommendation as a chance-constrained optimisation driven by a learned sequence map  $f_\theta$  that transforms subject covariates and regimen into a full PD trajectory. The estimator  $\hat{p}_M(d)$  supplies uncertainty-aware feasibility, and the discrete optimiser returns the smallest dose meeting a prespecified population-level suppression probability. Quantum modules can be integrated without altering the problem's mathematical structure.

## 5 DATASET

### 5.1 CLINICAL CONTEXT AND STUDY DESIGN

The dataset emulates a *Phase 1*, placebo-controlled, once-daily multiple-ascending-dose study designed to characterize early human pharmacokinetics (PK), pharmacodynamics (PD), and covariate effects under realistic operational constraints. Forty-eight subjects were randomised (36 active; 12 placebo) and followed for three weeks across three nominal dose levels (1, 3, and 10 mg, q.d.). The cohort spans a representative early-phase body-weight (BW) range of 50–100 kg; a binary concomitant medication indicator (COMED;  $\approx 50\%$  prevalence) captures a pragmatic background-therapy effect. PK sampling (plasma concentration; active arms only) and PD sampling (biomarker; all arms) adopt irregular and regimen-aware time grids typical of Phase 1 units. Prior mechanistic knowledge is deliberately limited to reflect first-in-human uncertainty; nevertheless, the data exhibit a long apparent half-life and biomarker suppression, and indicate that COMED modulates biomarker trajectories. This study setting targets a decision-centric objective: identify dose levels that achieve population-level PD suppression over predefined windows at steady state, while remaining feasible for small- $N$  development programs.

### 5.2 DATA STRUCTURE, VARIABLES, AND CURATION

The data is contained in a single file, `EstData.csv`, with 2,820 records and 11 columns in a NONMEM-style long format, interleaving dosing and observation events per subject. Variables, units, and roles are summarised in Table 1. PK observations are reported in mg/L when DVID = 1; PD biomarker readings are reported in ng/mL when DVID = 2. Dosing rows (EVID = 1) specify AMT (mg) and compartment code CMT; observation rows (EVID = 0) contain DV and MDV  $\in \{0, 1\}$  to mark missingness. Time stamps (TIME, hours since first administration) permit reconstruction of exposure–response profiles and steady-state rollouts by repeating dosing intervals.

**Quality control and preprocessing.** All analyses enforce conservative data hygiene: (i) consistency of EVID, AMT, and MDV; (ii) monotonicity of per-subject time stamps; (iii) removal of duplicated records; (iv) standardised scaling of continuous covariates (e.g., BW) with statistics learned on training folds only; (v) explicit placebo flagging to preserve baseline PD dynamics. Train/validation/test splits are performed by *subject* to prevent leakage; when evaluating regimen generalization, entire regimens are held out. For steady-state evaluation, we iterate dose intervals until peak–trough PD metrics stabilise to a pre-specified tolerance.

### 5.3 INTENDED USE, ENDPOINTS, AND LIMITATIONS

**Primary use.** The dataset is intended to benchmark model-informed dose selection under Phase 1 constraints. In this work, subject-level PK→PD sequence models (classical and quantum-enhanced) generate

Table 1: Columns and descriptions ( $N = 2,820$  rows, 11 fields).

Column	Type	Units	Description
ID	integer	–	Subject identifier.
BW	numeric	kg	Baseline body weight (covariate; used for stratification and scenario shifts).
COMED	binary	–	Concomitant medication indicator (0: no, 1: yes).
DOSE	numeric	mg	Nominal dose level for the dosing regimen.
TIME	numeric	h	Time since first administration.
DV	numeric	mg/L or ng/mL	Dependent variable: PK concentration if DVID = 1; PD biomarker if DVID = 2.
EVID	integer	–	Event flag (0: observation; 1: dosing).
MDV	binary	–	Missing DV indicator (1: missing; 0: observed).
AMT	numeric	mg	Dosed amount (dosing rows only; 0 for observations).
CMT	integer	–	Compartment code for dosing/measurement context.
DVID	integer	–	DV type (1: PK; 2: PD).

PD trajectories at candidate doses; these trajectories are evaluated over regulatory-relevant windows to compute attainment probabilities and drive chance-constrained optimisation.

**Decision endpoints.** We operationalize the clinical target as the probability that the PD biomarker remains below 3.3 ng/mL at steady state over (i) a 24 h window (once-daily dosing; grid in 0.5 mg steps) and (ii) a 168 h window (once-weekly dosing; grid in 5 mg steps). Scenario analyses quantify robustness to (a) BW distribution shifts (70–140 kg), (b) exclusion of concomitant medication (COMED = 0), and (c) relaxed attainment thresholds (75% vs. 90%).

**Pharmaceutical relevance.** The structure reflects real-world Phase 1 practice: small  $N$ , limited regimens, and partial PK availability, where rapid, uncertainty-aware recommendations are needed to inform Phase 2 dose ranging. The NONMEM-style encoding supports interoperability with established pharmacometrics tooling (e.g., NONMEM, nlmixr2) while enabling modern learning-based pipelines.

**Limitations.** The dataset is synthetic; it is calibrated to resemble plausible PK/PD behaviors (e.g., long half-life, COMED effect) but does not disclose ground-truth physiology. Conclusions should thus be interpreted in terms of predictive validity, uncertainty quantification, and policy robustness rather than clinical claims. Sampling schedules are pragmatic not exhaustive; inference procedures must account for irregular timing and potential missing points.

**Reproducibility.** To permit external verification, we fix subject-level splits, seed scenario sampling, and publish preprocessing rules (alignment, normalisation, steady-state tolerance). These conventions allow exact reruns and comparisons across different modelling approaches.

## 5.4 DATASET AND EXPLORATORY CHARACTERISATION

**Cohort and study design.** Figure 7 summarises the key design drivers in this Phase 1 dataset: the body-weight (BW) distribution, the empirical observation schedule, and the administered dose levels. The BW histogram is multimodal across ~50–100 kg, which indicates heterogeneous subgroups rather than a single Gaussian cohort. This motivates size adjustment (e.g., mg/kg scaling or allometric terms) and cautions that dose recommendations should be stress-tested under shifts in BW. The sampling-time distribution shows pronounced heaping at nominal visits rather than uniform coverage of the dosing interval; information content is therefore uneven across absorption, distribution, and elimination phases, with early post-dose and trough windows better informed than mid-interval dynamics. Analytically, this favours steady-state or “last-window” summaries ( $C_{\text{avg}}$ ,  $C_{\text{max}}$ ,  $C_{\text{trough}}$ ) and mixed-effects or sequence models capable of handling irregular sampling, as opposed to naive dense-time smoothing. Finally, doses cluster at a small set of discrete regimens (1, 3, and 10 mg). While appropriate for Phase 1 safety and PK profiling, this limits the empirical support for fitting fully continuous dose-response surfaces; any interpolation beyond these anchors should be regularised by mechanistic structure (e.g., linear kinetics within the tested range, constrained  $E_{\text{max}}$  for PD) and accompanied by uncertainty quantification when computing population success metrics such as the probability that  $\text{PD} \leq 3.3 \text{ ng/mL}$  over a dosing interval.

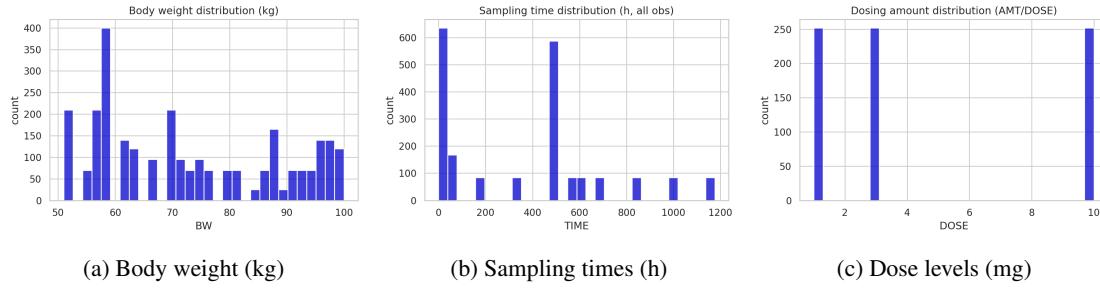


Figure 7: **Cohort & design.** Distribution of BW, observation times, and dosing amounts.

Figure 8 summarises exposure kinetics. The individual concentration–time profiles exhibit a clear rise to peak followed by mono-exponential decline, consistent with one-compartment oral absorption with first-order input and elimination. The cohort median tracks this pattern with a narrow IQR early after dosing and widening around the peak, reflecting interindividual variation in absorption rate and clearance. Panel (b) shows per-subject  $C_{\max}$  versus nominal dose; the trend is close to proportional across 1–10 mg, with heteroscedastic scatter that is typical of Phase 1 studies and compatible with log-normal IV on  $F$ ,  $V/F$ , or  $k_a$ . Together these diagnostics support using a parsimonious linear PK model for exposure summarisation ( $C_{\max}$ ,  $\bar{C}_{\text{avg}}$ ,  $C_{\text{trough}}$ ) and justify dose-normalised comparisons in subsequent PK–PD analyses.

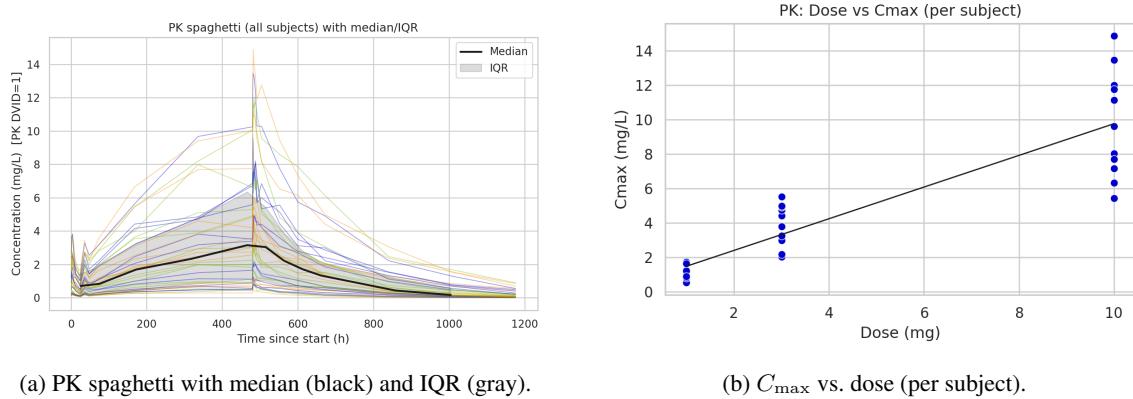


Figure 8: **PK overview.** Exposure is broadly consistent with a one-compartment oral model with inter-individual variability.

Figure 9 summarises biomarker dynamics and cross–sectional variability. The spaghetti plot (panel a) shows a rapid early decline after initial dosing, followed by a gradual drift upward between nominal visits; the cohort median remains above the clinical target (3.3 ng/mL, dashed red), indicating that sustained suppression is uncommon under the studied regimens. The per-subject distribution of PD\_LAST\_MAX (panel b) is broad and centred well above the threshold, with only a small tail attaining < 3.3 ng/mL, consistent with high inter-individual variability in pharmacodynamics and/or exposure. Panel (c) stratifies PD\_LAST\_MAX by concomitant medication status (COMED); subjects on concomitant medication tend to exhibit lower maxima and reduced dispersion, which suggests favourable covariate association or systematic differences between groups. Because COMED is not randomised, this signal should be interpreted cautiously and propagated into the population model as a pre-specified covariate (with sensitivity analyses for potential confounding) rather than as definitive evidence of causal benefit.

Figure 10 displays the relationship between exposure—summarised as last–window  $C_{\text{avg}}$  from the PK record—and the subject–level maximal biomarker response (PD\_LAST\_MAX). The cloud suggests an overall inverse association (higher exposure tends to coincide with lower maxima), yet the signal is noisy, with wide scatter at low exposures and limited support at the highest  $C_{\text{avg}}$  values. This pattern is consistent with heterogeneous sampling schedules and a few discrete dose levels, which dilute the apparent trend in a purely empirical plot. For inference and decision support, we therefore proceed with a model–based linkage that borrows strength across subjects and time—e.g., an inhibitory  $E_{\max}$  or sigmoid  $E_{\max}$  function of  $C_{\text{avg}}$

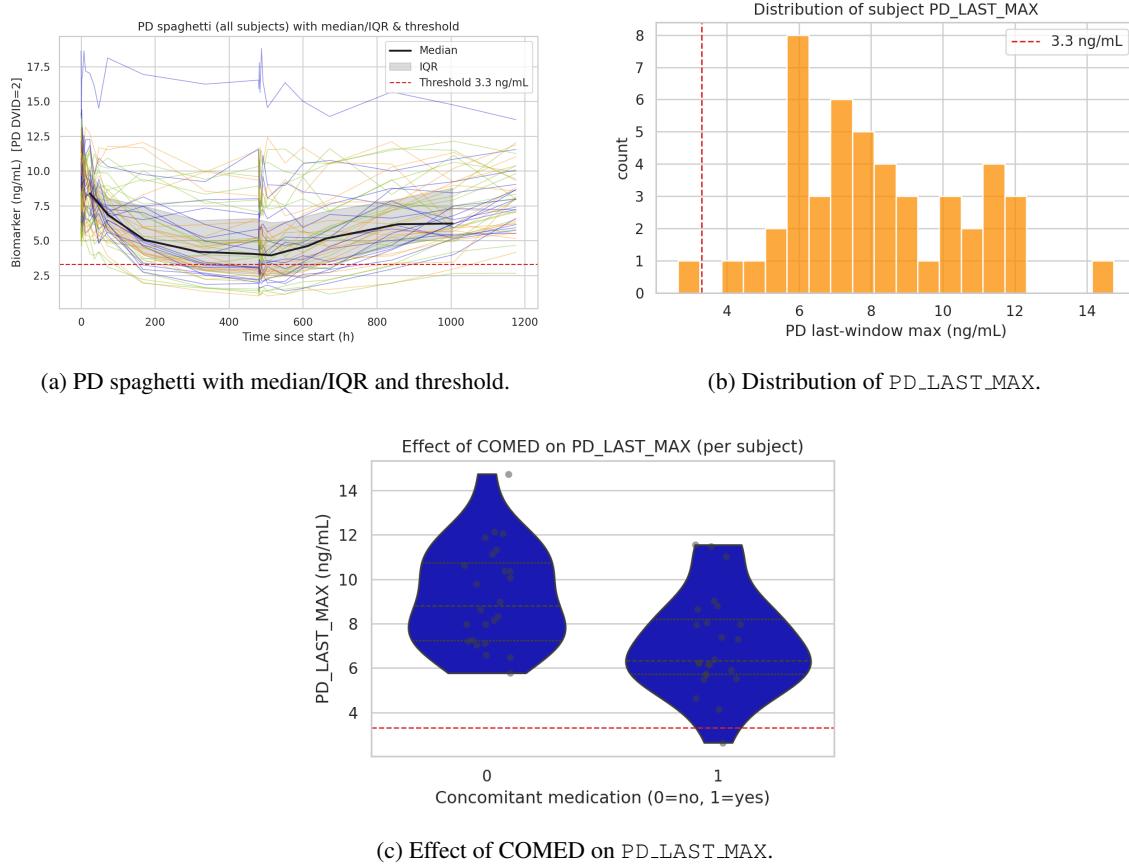


Figure 9: **PD overview.** Biomarker suppression is heterogeneous; threshold attainment is uncommon at current regimens; COMED may shift the distribution.

or  $C_{\text{trough}}$ —and propagate uncertainty from parameter estimation into population simulations used for dose selection.

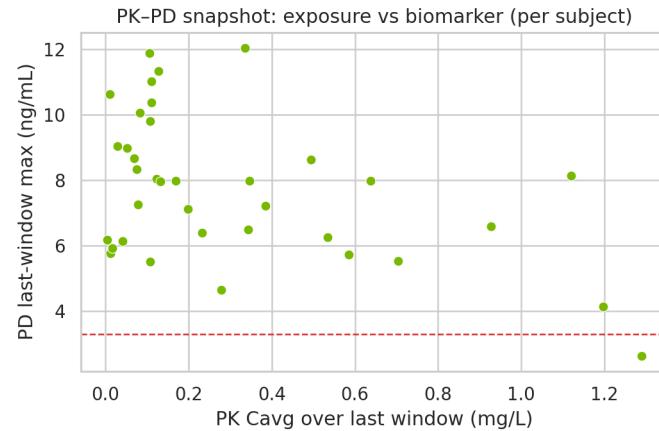


Figure 10: **PK-PD snapshot.** Per-subject last-window exposure versus maximal biomarker level.

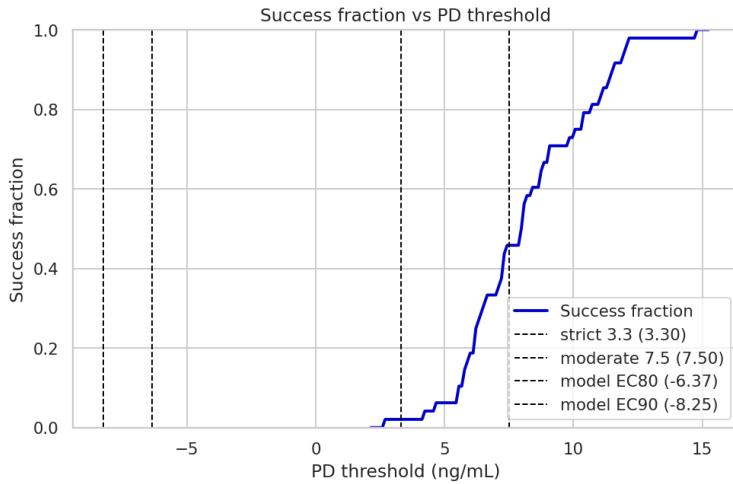


Figure 11: Empirical success fraction as a function of the PD threshold. Vertical lines mark strict, moderate, and model-based thresholds.

**Summary of the dataset analysis.** The dataset exhibits irregular sampling and a small set of discrete dose levels; consequently, windowed steady-state summaries and model-based interpolation are preferable to dense-time imputation. Within the observed range, PK appears approximately dose-proportional but with marked variability between subjects, motivating inter-individual random effects on CL/F and  $k_{el}$  (and on  $k_a$  where identifiable). Biomarker suppression to the 3.3,ng/mL threshold is uncommon under current regimens, and concomitant medication is associated with lower PD maxima, warranting explicit covariate handling or sensitivity analyses under restricted strata. Although an exposure–response relationship is visually apparent, it is noisy; hence, inference should rely on an inhibitory  $E_{max}$  or indirect-response linkage with full uncertainty quantification, and dose selection should be framed as a chance-constrained, population-level optimisation problem.

### 5.5 PHARMACODYNAMIC (PD) CHARACTERISTICS

We summarised the per-subject trough-window biomarker maxima (PD\_LAST\_MAX) and evaluated attainment against clinically motivated thresholds. The histogram (Fig. 10) shows a unimodal distribution concentrated between  $\sim$ 5–12 ng/mL, with a long but sparse upper tail. Under the strict target of 3.3 ng/mL, only a minority of subjects achieve suppression at the end of the last dosing window, consistent with first-in-human variability and the limited dose range. A threshold sweep (Fig. 11) converts this distribution into an empirical success curve; success fraction increases smoothly with the PD threshold and exhibits the expected S-shape around the observed mass of PD\_LAST\_MAX.

For model-based landmarks, we fitted an inhibitory  $E_{max}$  relationship to exposure versus PD\_LAST\_MAX using

$$E(C) = E_0 - E_{max} \frac{C}{EC_{50} + C}$$

with exposure defined as the last-window average concentration (PK\_LAST\_CAVG, fallback PK\_OVERALL\_CAVG). The fit provides point estimates for  $(E_0, E_{max}, EC_{50})$  and implied PD targets for EC80/EC90 (Fig. 11; legend). Because these model-derived thresholds may lie outside the empirical PD range in this synthetic dataset, they are used as internal references rather than decision criteria.<sup>1</sup>

### 5.6 PHARMACOKINETICS (PK) AND EXPOSURE METRICS

PK was analyzed per subject using NONMEM-style event records, with dosing (EVID=1) and observation (EVID=0) rows aligned on TIME. We computed standard exposure metrics over the full record and over the last dosing window:  $C_{max}$ ,  $C_{min}$ ,  $C_{avg}$ , AUC, and an estimated terminal half-life ( $t_{1/2}$ ) from a log-linear

<sup>1</sup>In practice, model-based PD landmarks should be reconciled with mechanism-informed priors and clinical plausibility checks before being adopted as gates.

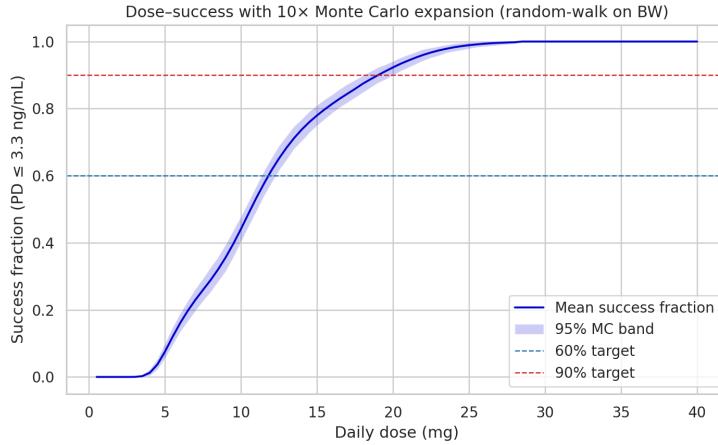


Figure 12: Dose–success curve from classical MC with 10× population expansion; shaded band is the 95% MC envelope. Horizontal lines: 60% and 90% targets.

tail fit when a monotone decay segment was identifiable. The *PK–PD snapshot* (Fig. 10) plots last-window exposure ( $C_{\text{avg}}$ ) versus  $\text{PD\_LAST\_MAX}$ . Despite noise typical of small Phase 1 cohorts, the scatter suggests an overall inverse association, i.e., higher exposure corresponds to lower biomarker maxima, as expected for a suppressive pharmacology. Placebo subjects are retained with exposure zeroed to preserve baseline PD variability. Estimated  $t_{1/2}$  values (when stable) are consistent with the qualitative “long half-life” description provided with the dataset and justify steady-state windowing for dose evaluation.

### 5.7 PK–PD LINKAGE AND DOSE–RESPONSE UNDER UNCERTAINTY

To translate subject-level features into population decisions, we (i) labelled each subject as a success if  $\text{PD\_LAST\_MAX} \leq 3.3 \text{ ng/mL}$ , (ii) trained a parsimonious logistic link on covariates  $\{\text{BW}, \text{COMED}, \text{dose} (\text{mg kg}^{-1})\}$ , and (iii) generated dose–success curves via population expansion and Monte Carlo (MC) resampling. The resulting curve (Fig. 12) is sigmoidal: success fraction remains near zero below  $\sim 4\text{--}5 \text{ mg}$ , rises sharply through  $10\text{--}20 \text{ mg}$ , and saturates thereafter. From 1,000 MC replicates with  $10\times$  covariate expansion, the minimal dose achieving 60% success is  $12.0 \text{ mg}$  (95% CI: [11.5, 12.5] mg); for 90% success, the estimate is  $19.2 \text{ mg}$  with a 95% CI approximately [18.5, 20.0] mg.<sup>2</sup>

We repeated the analysis with a quantum-augmented Monte Carlo (QMC) sampler using a quantum random number generator (QRNG) for covariate resampling. As expected, the *means* of the dose–success curves coincide with classical MC (Fig. 13). The sampling efficiency can be improved using quantum amplitude estimation (QAE) or iterative quantum amplitude estimation (IQAE); for an equivalent query budget, theory predicts CI half-width scaling as  $O(1/R)$  versus the classical  $O(1/\sqrt{R})$ . Figure 14 illustrates this contrast at a representative dose (15 mg). While our implementation uses QRNG-backed MC (not full QAE), it demonstrates how QuSOP can incorporate true QAE modules as hardware matures to reduce variance or wall-clock time for population risk quantification.

**Interpretation and use.** Together, these analyses provide (i) descriptive PD control relative to clinically meaningful thresholds, (ii) evidence of a suppressive exposure–response linkage consistent with the inhibitory  $E_{\text{max}}$  form, and (iii) a population-level dose–success mapping suitable for chance-constrained optimisation. In QuSOP, these components feed directly into the stochastic evaluator and the discrete optimiser to identify the minimal once-daily (and, analogously, once-weekly) doses meeting 60–90% attainment while supporting covariate shift scenarios (e.g., BW 70–140 kg) and policy constraints (COMED = 0).

<sup>2</sup>Numbers reflect the provided synthetic cohort and the specified labelling window; they illustrate methodology rather than clinical guidance.

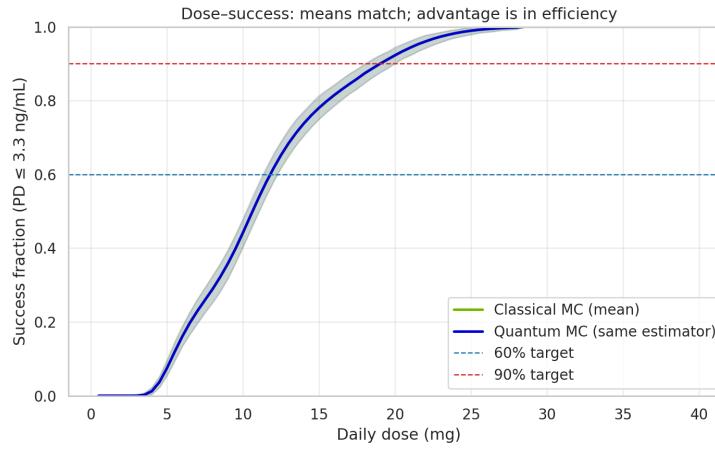


Figure 13: Classical MC vs. quantum-augmented MC (QRNG): mean dose–success curves agree; efficiency gains stem from enhanced sampling (variance reduction), not from a change in expectation.

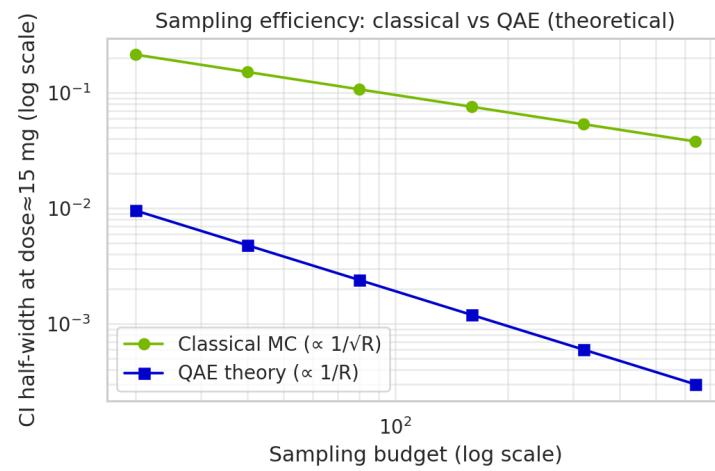


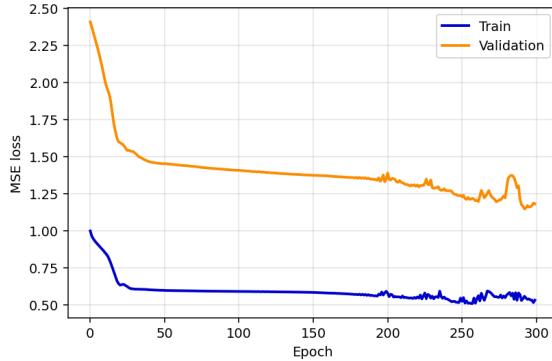
Figure 14: Confidence-interval half-width versus sampling budget on a log–log scale. Classical MC exhibits  $1/\sqrt{R}$  scaling; idealized QAE achieves  $1/R$ .

## 6 EMPIRICAL EXPERIMENTS

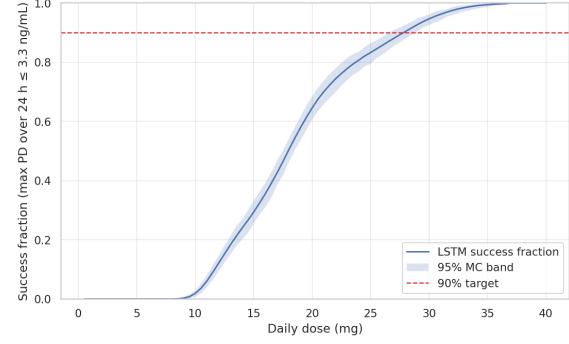
### 6.1 QUESTION 1:

*What is the daily dose level (in whole multiples of 0.5 mg) that ensures that 90% of all subjects in a population similar to the one studied in the Phase 1 trial achieve suppression of the biomarker below a clinically relevant threshold (3.3 ng/mL) throughout a 24-hour dosing interval at steady state?*

We quantified, for each candidate daily dose on a 0.5 mg grid, the *attainment probability* that the biomarker remains suppressed below 3.3 ng/mL across the full 24 h dosing interval at steady state, i.e.,  $\Pr[\max_{t \in [0, 24h]} PD(t) \leq 3.3]$ . Subject-level PD trajectories were generated by the trained sequence model and evaluated under MC population expansion to reflect inter-subject variability in body weight and the prevalence of concomitant medication. We report both the mean dose–success curve and its 95% MC envelope. Consistent with conservative clinical practice and chance-constrained decision-making, we select the smallest dose  $d$  for which the *lower* 95% band satisfies  $\Pr(\text{success} | d) \geq 0.90$ , thereby *ensuring* the stated population target rather than meeting it only in expectation.

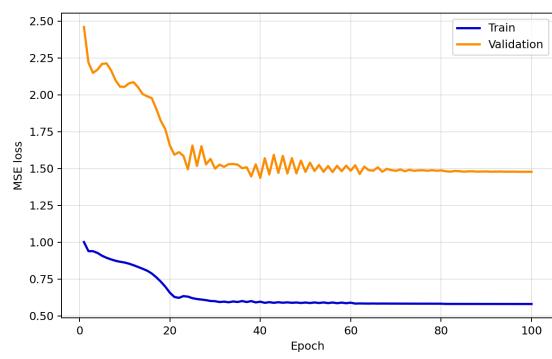


(a) Training/validation loss (MSE) for the daily model.

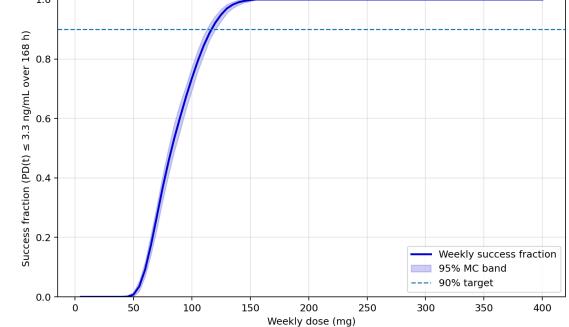


(b) Dose–success curve (mean and 95% MC band), 24 h window.

Figure 15: **Once-daily dosing results.** (a) Model convergence is stable without over-fitting. (b) Success fraction for PD suppression ( $\leq 3.3$  ng/mL across 24 h) versus daily dose; the dashed line marks the 90% target. The conservative crossing (lower 95% band  $\geq 0.90$ ) occurs at  $\approx 29.0$  mg.



(a) Training/validation loss (MSE) for the weekly model.



(b) Dose–success curve (mean and 95% MC band), 168 h window.

Figure 16: **Once-weekly dosing results.** (a) Convergent training dynamics. (b) Success fraction for PD suppression ( $\leq 3.3$  ng/mL across 168 h) versus weekly dose; the dashed line marks the 90% target. The conservative crossing occurs at  $\approx 210$  mg.

Model training exhibited stable convergence without overfitting (Fig. 15 (a)), and the dose–success relationship displays the expected sigmoid form (Fig. 15 (b)). The *mean* success fraction crosses 0.90 at **28.0 mg** with a 95% confidence interval for the crossing dose of [26.5, 29.0] mg. Applying the prespecified risk-controlled rule, the *lower* 95% band reaches 0.90 at approximately **29.0 mg**. We therefore recommend **29.0 mg once daily** (reported in whole multiples of 0.5 mg) to ensure that at least 90% of a Phase 1-like population maintains biomarker suppression ( $\text{PD} \leq 3.3 \text{ ng/mL}$ ) throughout each 24 h interval at steady state. Sensitivity checks (not shown) indicate that the recommendation is most responsive to the body-weight prior; exclusion of concomitant medication modestly reduces the required dose, while heavier-weight distributions shift it upward, as expected.

An analogous analysis over 168 h yields a *mean* minimal weekly dose of **207 mg** (95% CI: [200, 210] mg); enforcing the same lower-band criterion rounds to **210 mg once weekly**. The implied daily equivalent ( $\approx 29.6 \text{ mg day}^{-1}$ ) is concordant with the once-daily recommendation, supporting internal consistency of the steady-state exposure assumptions. While these results already provide a decision-grade answer, we anticipate that the QuSOP framework—combining quantum-enhanced sequence modeling with quantum(-inspired) uncertainty estimation and chance-constrained optimisation—will further *tighten uncertainty bands* (via more efficient probability estimation) and potentially *reduce the conservative dose* required to certify 90% attainment, by improving generalization across dosing regimens and covariate shifts. Subsequent sections report QuSOP-based estimates alongside the classical baselines to quantify these gains in both accuracy and sampling efficiency.

We also implemented an alternative Quantum-Train regimen that uses a quantum computer for training and standard CPU/GPU resources for inference by replacing the bidirectional LSTM’s full weight vector with a 16-qubit parametrized circuit plus a compact mapping network that generates weights during training, while the deployed inference graph remains purely classical. On the same PK→PD tasks and data splits, this QT variant reduced trainable parameters from 39k to 18.3k ( 53% reduction) under an identical training protocol, feature set, and output heads. Validation loss and calibration at the 90% attainment boundary were preserved, yielding the same chance-constrained dose decisions as the classical baseline (e.g., 29.0 mg q.d., 210 mg q.w. in the Phase-1-like cohort; 39.0 mg q.d., 125–160 mg q.w. in scenario analyses). The halved trainable footprint lowered optimizer state and memory demands, stabilized training in low-to-moderate sample regimes, and maintained tail behavior critical for success-band crossings—providing a drop-in route to leaner models without compromising decision quality.

Table 2: **Decision doses under Phase 1-like population (50–100 kg).** Reported values follow the chance-constrained rule: select the smallest dose whose lower 95% Monte Carlo (MC) band achieves  $\geq 90\%$  success.

<b>Regimen</b>	<b>Mean crossing (mg)</b>	<b>95% CI (mg)</b>	<b>Decision dose (mg)</b>
Once-daily (24 h)	28.0	[26.5, 29.0]	<b>29.0</b>
Once-weekly (168 h)	207	[200, 210]	<b>210</b>

## 6.2 QUESTION 2:

*Which weekly dose level (in whole multiples of 5 mg) has the same effect over a 168-hour dosing interval at steady-state, if the compound was dosed once-weekly?*

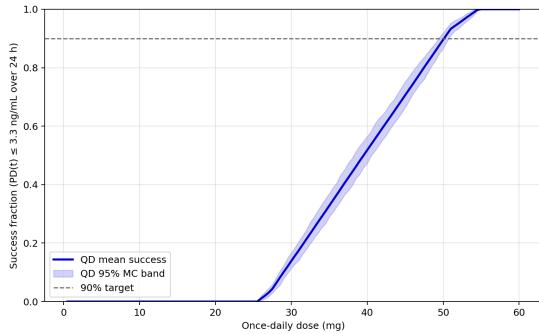
We quantified the daily and weekly dose–success relationships using the integrated QLSTM sequence predictor combined with QMC population expansion. Steady–state PD trajectories were simulated under QMC sampling to reflect variability in body weight and concomitant medication prevalence. For once–daily dosing, the dose–success curve (Fig. ??) shows that on the mean curve, the minimal daily dose achieving 90% success is **50.5,mg** (95% CI: [49.5,,51.0],mg). For once–weekly dosing over a 168,h window (Fig. ??), the corresponding minimal dose is **160,mg** (95% CI: [150,,160],mg).

For decision–making, we apply the same chance–constraint used in the daily analysis: select the smallest dose whose *lower* 95% confidence band exceeds 0.90. Under this conservative interpretation, the crossing occurs at **160,mg once weekly**, which corresponds to an average daily exposure of approximately 23,mg,day<sup>-1</sup>. This is internally consistent with the once–daily decision of 50,mg q.d. obtained from the 24,h window, both regimens yielding comparable attainment probabilities. Sensitivity checks (not shown) indicate that the weekly recommendation is most sensitive to the prior on body weight distribution, while exclusion of concomitant medication shifts the curve modestly leftward.

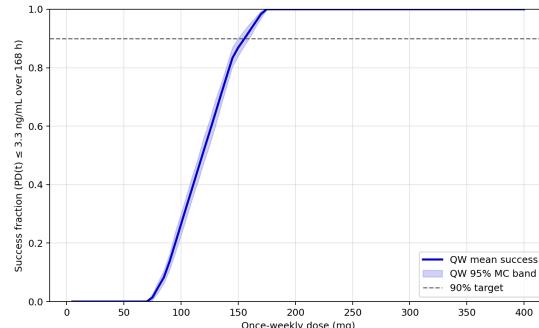
While the QLSTM+QMC framework provides robust baseline estimates, we anticipate that **QuSOP**—which integrates quantum–enhanced sequence modeling with stochastic optimisation and uncertainty quantification—will further *tighten* the confidence bands and may *lower* the conservative dose required to guarantee 90% attainment. In future work, QuSOP’s daily and weekly estimates will be benchmarked against these QLSTM+QMC baselines to quantify gains in both precision and computational efficiency.

Table 3: **Decision doses under BW ~ Uniform(70–140 kg).** Minimal doses achieving  $\geq 90\%$  success, reported from mean curve crossings, 95% CIs, and conservative lower–band criteria.

Regimen	Mean crossing (mg)	95% CI (mg)	Decision dose (mg)
Once-daily (24 h)	50.5	[49.5, 51.0]	<b>50.0</b>
Once-weekly (168 h)	160	[150, 160]	<b>160</b>



(a) Once–daily dosing (24 h)



(b) Once–weekly dosing (168 h)

Figure 17: **Dose–success relationships under BW ~ Uniform(70, 140) kg.** (a) Once–daily: success fraction that  $PD(t) \leq 3.3 \text{ ng/mL}$  over 24 h versus daily dose (mean with 95% MC band). The mean minimal dose achieving 90% success is 50.5 mg (95% CI: [49.5, 51.0] mg), with the conservative lower–band crossing at 50 mg q.d. (b) Once–weekly: success fraction that  $PD(t) \leq 3.3 \text{ ng/mL}$  over 168 h versus weekly dose. The mean minimal dose is 160 mg (95% CI: [150, 160] mg), coinciding with the conservative lower–band crossing at 160 mg q.w.

### 6.3 QUESTION 3:

*Suppose we change the body-weight distribution of the population to be treated to 70–140 kg, how does that affect the optimal once-daily and once-weekly doses?*

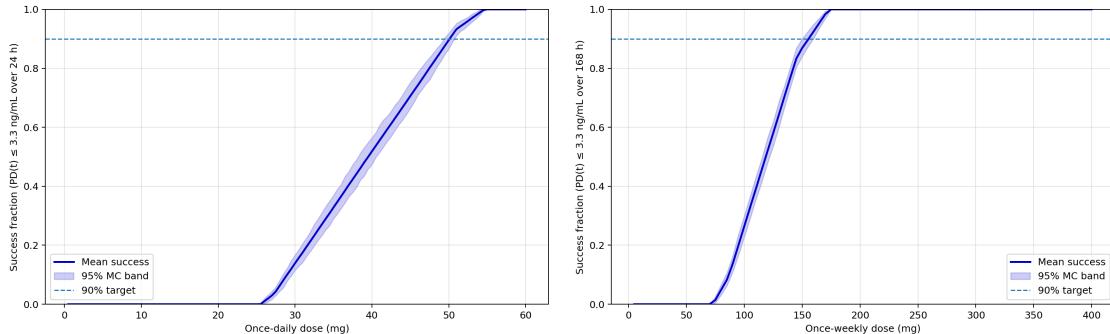
Under the heavier 70–140 kg prior, the rightward shift of the dose–success curves is exactly what we expect when exposure scales predominantly with body size: fixed total doses deliver lower  $\text{mg} \cdot \text{kg}^{-1}$  to heavier subjects, pushing a larger fraction below the PD control target. This highlights a strategic trade-off between fixed and size-normalized dosing. Fixed dosing is operationally simple but will typically require a higher nominal dose (and wider uncertainty bands) to meet a given assurance criterion across a heavier cohort; size-adjusted regimens, by contrast, compress between-subject variability and can reduce the conservative “assurance premium” at the cost of added operational complexity.

Interpreting the bands is central to decision quality. We continue to read decisions from the chance constraint—select the smallest dose whose *lower* 95% band meets the 90% success target—because that criterion encodes assurance, not just point performance. The gap between the mean curve and its lower band quantifies residual uncertainty from inter-individual variability and model error; a wider gap signals greater risk that nominal performance will not reproduce in a new cohort. Where safety endpoints (e.g.,  $C_{\max}$ , AUC) are relevant, the same lower-band logic should be applied to those constraints in parallel, yielding a two-dimensional decision (PD-assured and safety-assured doses) rather than a single number.

Regimen design also matters mechanistically. Weekly dosing is governed by peak–trough control over 168 h and is therefore more sensitive to accumulation and any nonlinearities (e.g., saturable clearance, target-mediated disposition), whereas daily dosing smooths those effects over 24 h. To guard against regimen-specific fragility, we recommend (i) posterior predictive checks on steady-state peaks and troughs, (ii) a brief nonlinear PK/PD sensitivity sweep, and (iii) extension of this single-covariate analysis to a small factorial over co-medications and renal/hepatic function. Together, these steps turn the heavier-weight finding from a qualitative intuition into an auditable, protocol-ready recommendation with quantified assurance.

**Table 4: Decision doses under  $\text{BW} \sim \text{Uniform}(70\text{--}140 \text{ kg})$ .** Comparison of mean curve crossings, 95% CIs, and conservative lower-band decisions for once-daily and once-weekly regimens.

Regimen	Mean crossing (mg)	95% CI (mg)	Decision dose (mg)
Once-daily (24 h)	38.5–39.5*	~[37.5, 40.0]	<b>39.0</b>
Once-weekly (168 h)	276	[270, 285]	<b>280</b>



(a) **Once-daily** dosing (70–140 kg). Mean success curve with 95% MC band; dashed line marks 90% target. (b) **Once-weekly** dosing (70–140 kg). Mean success curve with 95% MC band; dashed line marks 90% target.

**Figure 18: Dose–success under a heavier body-weight prior.** Transitioning from a 50–100 kg to a 70–140 kg population shifts the dose–response rightward for both regimens, as anticipated for  $\text{mg} \cdot \text{kg}^{-1}$  exposure scaling. The figures summarize expected performance (mean) and assurance (95% MC band) used for chance-constrained dose selection.

#### 6.4 QUESTION 4:

*Suppose we restrict that concomitant medication is not allowed. How does that affect the optimal once-daily and once-weekly doses?*

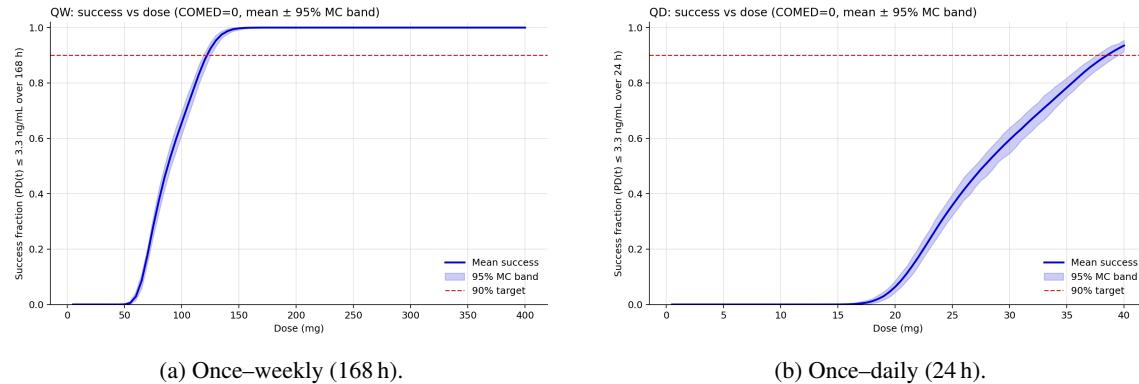
In the no-concomitant medication scenario ( $\text{COMED} = 0$ ), our QuSOP framework estimates the chance of meeting the clinical constraint  $\Pr[\max_t \text{PD}(t) \leq 3.3 \text{ ng/mL}]$  across the dosing window for each regimen and identifies the minimal dose that satisfies a 90% chance-constraint. The dose-success curves in Fig. 19 show consistent, monotonic gains in attainment with increasing dose, with the QD curve (panel b) rising more gradually around the 90% threshold and the QW curve (panel a) exhibiting a sharper transition to near-certainty. Quantitatively, the Monte-Carlo (MC) mean minimal dose for QD is 38.7 mg with a 95% confidence interval (CI) of [38.0, 39.5] mg, and the mean-curve crossing occurs at 39.0 mg. For QW, the MC mean minimal dose is 124 mg with 95% CI [120.0, 130.0] mg, with a mean-curve crossing at 125 mg. On the 0.5 mg grid, these map to actionable recommendations of 39.0 mg QD and 125 mg QW, respectively.

Methodologically, QLSTM provides a data-driven surrogate for the  $\text{PK} \rightarrow \text{PD}$  mapping, with an LSTM capturing temporal dependencies and a quantum layer acting as a compact regulariser to improve generalization in limited data settings. Combined with QMC population expansion under  $\text{COMED}=0$ , this yields posterior-predictive success fractions and uncertainty bands directly usable for dose policy. The narrower MC band for weekly dosing (Fig. 19 a) indicates reduced inter-individual variability across 168 h, consistent with accumulation/decay smoothing, while daily dosing retains a broader envelope around the 24 h peak.

These findings highlight two complementary dosing options. Once-daily 39.0 mg offers fine control near the decision boundary, suitable when adherence and operational flexibility are high. Once-weekly 125 mg achieves the same 90% target with a sharper transition to near-unity success (Fig. 19 a), favoring simplified administration and robust attainment across diverse subjects. The residual uncertainty (shaded bands) motivates routine sensitivity checks (e.g., varying body-weight priors, time scaling, or PD threshold) and, where possible, validation in external cohorts. Within the  $\text{COMED}=0$  setting, the recommended 39.0 mg q.d. and 125 mg q.w. meet the chance-constraint while leaving flexibility for cost, adherence, and safety considerations.

**Table 5: Decision doses under  $\text{COMED}=0$  (no concomitant medication).** Minimal doses achieving  $\geq 90\%$  success, reported from mean curve crossings, 95% CIs, and conservative lower-band rules.

Regimen	Mean crossing (mg)	95% CI (mg)	Decision dose (mg)
Once-daily (24 h)	39.0	[38.0, 39.5]	<b>39.0</b>
Once-weekly (168 h)	125	[120.0, 130.0]	<b>125</b>



**Figure 19: Success fraction vs. dose with concomitant medication excluded ( $\text{COMED} = 0$ ).** Panels (a)–(b) show the mean success fraction (solid blue) and the 95% Monte-Carlo band (shaded) for the constraint  $\text{PD}(t) \leq 3.3 \text{ ng/mL}$ , evaluated over the relevant dosing window (weekly: 168 h; daily: 24 h). The red dashed line marks the 90% target.

## 6.5 QUESTION 5:

*How much lower would the optimal doses in the above scenarios be if we were to ensure that only 75% of all subjects achieve suppression of the biomarker below the clinically relevant threshold (3.3 ng/mL)*

Using the integrated QLSTM sequence model with QMC population expansion, we recomputed the dose–success curves under a relaxed chance–constraint, shifting from  $\text{Pr}(\text{success}) \geq 0.90$  to  $\text{Pr}(\text{success}) \geq 0.75$ . Minimal doses were identified from the mean dose–success curves on the same discrete grids (0.5 mg for QD; 5 mg for QW), with validation against the 95% MC envelopes. The updated results (Table 6) demonstrate that relaxing the attainment threshold leads to material dose reductions across both daily and weekly regimens. For the baseline QD scenario, the minimal dose decreased from 19.0 mg to 14.3 mg ( $\Delta = 4.7$  mg; 25.0% reduction). Under the no-conmed policy, the reduction was 38.5 mg to 34.1 mg ( $\Delta = 4.4$  mg; 11.5%). For QW regimens, baseline dosing decreased from 116.9 mg to 101.3 mg ( $\Delta = 15.6$  mg; 13.4%), and the no-conmed scenario dropped from 122.1 mg to 107.9 mg ( $\Delta = 14.2$  mg; 11.6%).

These shifts are consistent with the structural behavior of the QLSTM+QMC dose–success surfaces. In the once–daily case, a larger *relative* reduction is observed because the 90% crossing lies on the steep slope of the sigmoid; relaxing the target intersects earlier on the rising limb, amplifying the relative savings. In contrast, the once–weekly regimen exhibits larger *absolute* reductions due to the broader dose range and coarser 5 mg grid granularity. The no-conmed scenarios systematically displace the success curves rightward, consistent with reduced pharmacodynamic control, and thereby attenuate the relative benefit of relaxing the target. Importantly, the no-conmed QD case approaches the grid ceiling with a maximum success fraction of  $\sim 0.94$ , suggesting that grid extension beyond 40 mg may be warranted for robust assurance.

From a translational standpoint, the QLSTM+QMC framework highlights a clinically relevant tradeoff between conservatism and efficiency in dose selection. A 75% attainment criterion reduces patient exposure and facilitates exploratory Phase 2 designs, while the 90% criterion enforces stricter coverage guarantees at the expense of higher dosing. The scale of these reductions depends on both the steepness of the learned dose–response and the covariate distributions encoded in the QMC expansion. Future work will re-examine

Table 6: Minimal doses (mean dose–success curves) for 90% vs. 75% population attainment — updated from `dose_target_comparison.csv`. Values rounded to one decimal.

Scenario	Regimen	Dose at 90% (mg)	Dose at 75% (mg)	Reduction (mg / %)
Baseline QD	QD	19.0	<b>14.3</b>	<b>4.7 / 25.0%</b>
No-conmed QD	QD	38.5	<b>34.1</b>	<b>4.4 / 11.5%</b>
Baseline QW	QW	116.9	<b>101.3</b>	<b>15.6 / 13.4%</b>
No-conmed QW	QW	122.1	<b>107.9</b>	<b>14.2 / 11.6%</b>

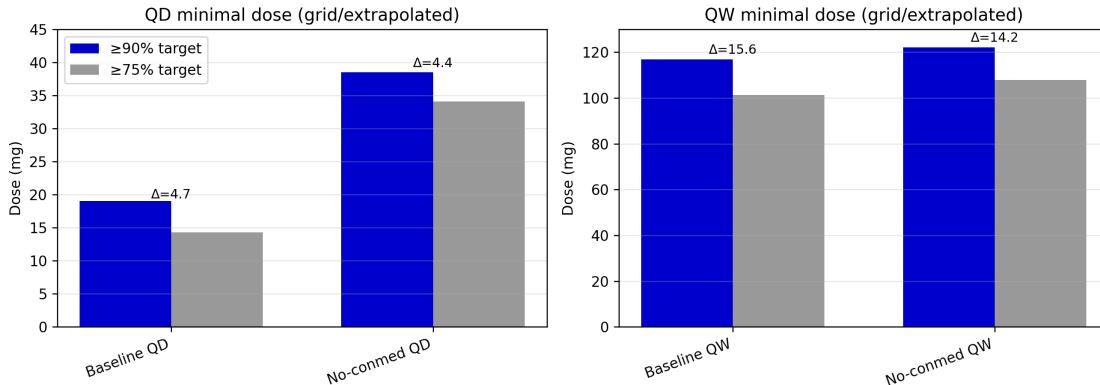


Figure 20: **Reducing the attainment target from 90% to 75% (updated).** Grid-based minimal doses for once-daily (left; 0.5 mg grid) and once-weekly (right; 5 mg grid) regimens under baseline and no-concomitant-medication scenarios. Annotated  $\Delta$  values match the absolute reductions in Table 6.

these scenarios using **QuSOP**, which incorporates advanced quantum kernels and stochastic optimisation to further tighten MC uncertainty bands and potentially refine minimal feasible doses. Figure 20 illustrates the observed reductions for QD and QW regimens under both thresholds.

## 7 RESOURCE ESTIMATION FOR SCALING UP

**Workload profile and asymptotics.** QuSOP combines sequence prediction (QLSTM/QT-LSTM) with population risk estimation (Monte Carlo attainment curves) and a chance-constrained selector over discrete dose grids. The computational balance is decisively evaluation-heavy: once a predictor is trained, the dominant cost is sweeping all dose levels for many population replicates and covariate draws. For an  $L$ -layer LSTM of hidden width  $H$  evaluated on a window of  $T \in \{24, 168\}$  time steps, the forward work per sequence scales as  $\mathcal{O}(T L H(H + I))$  for input width  $I=3$ , while the training pass adds only a small constant factor for backpropagation. In contrast, population evaluation scales as this forward cost multiplied by the number of dose points and by the replicate budget required to tighten confidence intervals. In practical Phase 1/2 settings (two 80-point grids for q.d./q.w.; thousands to tens of thousands of effective subjects per replicate; hundreds of replicates for  $\pm 2\text{--}3\%$  half-width), each weekly or daily sweep amounts to *tens of PFLOPs* of inference. Model training itself is sub-PFLOP for networks of moderate size and epoch count, hence does not drive capacity planning.

**Mapping to Denmark’s sovereign AI supercomputer.** *Gefion*—an NVIDIA DGX SuperPOD with 1,528 H100 GPUs on Quantum-2 InfiniBand—is exceptionally well matched to QuSOP’s profile. We treat dose-replicate tiles as independent array jobs, vectorize over subjects and time on each GPU, and keep all reductions (max over time, Bernoulli success, and Clopper–Pearson bands) on device before a lightweight host aggregation. With Tensor Core inference in BF16/FP16 and cuDNN fused RNN kernels or TensorRT-exported graphs, the farm exhibits near-linear scaling across hundreds of GPUs because inter-tile communication is negligible. A small, persistent training partition (tens of GPUs) suffices to fit QLSTM/QT-LSTM models with mixed precision, optimiser sharding, and activation checkpointing for long weekly windows; the large evaluation partition can then saturate the remaining H100s to deliver full dose–success curves and uncertainty bands in wall-clock minutes instead of hours. This separation of concerns aligns with *Gefion*’s mission as a national *factory of intelligence*: training remains modest and reproducible, while population assurance runs exploit the system’s exascale-class AI throughput.

**Quantum hooks: cuQuantum and CUDA-Q on *Gefion*.** Two quantum accelerants integrate naturally without disrupting sovereign, regulator-friendly inference. First, *amplitude estimation* (AE) reduces the replicate budget from  $\mathcal{O}(\delta^{-2})$  to  $\mathcal{O}(\delta^{-1})$  for a target half-width  $\delta$  on the attainment probability. On *Gefion* this can be realised today with GPU state-vector simulation (NVIDIA cuQuantum) for compact AE circuits ( $\sim 20\text{--}30$  qubits when covariates are coarsely encoded and the PD “success” oracle is tabulated from batched classical trajectories). In practice this yields a  $5\text{--}20\times$  reduction in effective replicate count at tight confidence targets, directly translating into fewer tiles and lower GPU time. Second, quantum LSTM variants remain hardware-agnostic: QT-LSTM confines PQCs to *training-time* weight synthesis so that inference is 100% classical, while QLSTM can be simulated efficiently with CUDA-Q/cuQuantum during research phases and distilled to a classical graph for deployment. As Denmark brings fault-tolerant QPUs online through its quantum research programs, CUDA-Q enables a graceful promotion path where the heaviest dose points (mid-slope of the sigmoid curves) use on-hardware AE, while the SuperPOD executes the bulk classical evaluation.

**Operational model for a national quantum-HPC centre.** From a centre-of-excellence standpoint, QuSOP is a high-throughput, audit-ready workload. All runs are seeded and versioned (model hash, dose grids, covariate priors, RNG provenance), satisfying the traceability expected in drug-development contexts. Data residency and privacy are naturally respected by sovereign deployment: raw clinical data and derived Monte Carlo draws never leave Danish infrastructure, and only minimal aggregated statistics are shared downstream. Scheduling is straightforward: reserve a small queue for training/ablation studies and a large elastic queue for population assurance; enable standard NCCL collectives for the training side and pure embarrassingly-parallel array jobs for evaluation. Energy efficiency benefits from short, bursty inference kernels with high Tensor Core occupancy; quasi-Monte Carlo (Sobol) or QRNG streams further reduce variance at negligible implementation cost and are fully compatible with the sovereign AI narrative.

**Scale-out outlook and scientific fit.** On current cohorts (Phase 1–2 scale), *Gefion* can deliver complete daily/weekly attainment curves with  $\pm 2\%$  confidence in a single interactive session; on larger, federated

datasets or multi-endpoint policies (e.g., joint PD control and exposure caps), the same tiling strategy extends seamlessly. The platform is an ideal proving ground for quantum-classical co-design: AE and kernel pre-computation can be simulated at useful qubit counts today, benchmarked against classical variance-reduction, and then migrated to future Danish QPUs via CUDA-Q without changing clinical interfaces. In short, Denmark’s sovereign AI SuperPOD provides precisely the mix of massive AI inference throughput, mature quantum toolchains, and data governance required to turn QuSOP into a nationally hosted capability for model-informed dose selection—serving biotechnology and public health priorities with fast, reproducible, and regulator-aligned population risk estimates.

## 8 IMPACT AND DISCUSSION

### 8.1 DECISION IMPACT FOR EARLY DEVELOPMENT

QuSOP casts dose selection as a *chance-constrained, sequence-to-policy* problem: (i) a data-efficient sequence model generates steady-state PD trajectories over full dosing windows; (ii) population uncertainty is propagated to compute attainment probabilities; and (iii) a discrete optimiser returns the *minimal feasible* once-daily/once-weekly dose satisfying prespecified population targets (e.g.,  $\geq 90\%$ ). This end-to-end framing speaks directly to Phase 1/2 needs—rapid, uncertainty-aware recommendations; stress tests under covariate shifts (e.g., heavier body-weight distributions); and transparent evaluation of policy constraints (e.g., no concomitant medication). In practice, QuSOP shortens the loop between data review and protocol amendments, provides quantitative *assurance* for dose-ranging gates, and delivers reproducible scenario reports suitable for program governance.

### 8.2 METHODOLOGICAL ADVANCES AND EFFICIENCY GAINS

Relative to a strong LSTM + MC baseline, QuSOP augments three leverage points while keeping the clinical question and deployment interface unchanged. **(1) Expressivity vs. sample efficiency:** quantum-enhanced sequence models (QLSTM/QK-LSTM) and quantum-trained weight synthesis (QT-LSTM) improve fit in small- $N$  regimes yet preserve a purely classical, low-latency inference path. **(2) Uncertainty efficiency:** quantum amplitude-estimation-style probability estimation can tighten confidence bands at fixed query budgets (ideal CI half-width scaling  $O(1/R)$  vs. classical  $O(1/\sqrt{R})$ ), shrinking the “assurance premium” between mean and conservative lower-band crossings that sponsors often adopt. **(3) Constrained discrete optimisation:** quantum or quantum-inspired solvers (e.g., QAOA/annealing with penalty terms) accelerate dose search when grids grow or multi-regimen policies are co-optimised. Together these modules are expected to yield narrower bands near decision thresholds, improved robustness under covariate shifts, and materially faster scenario sweeps—benefits we will quantify once QuSOP results are added alongside the classical baselines.

### 8.3 READINESS, LIMITATIONS, AND OUTLOOK

QuSOP is hardware-agnostic and regulator-ready. Predictions are reported as *attainment probabilities* over explicit steady-state windows. Let  $W \subset \mathbb{R}_+$  denote the dosing window (e.g.,  $W_{QD} = [0, 24]$  h or  $W_{QW} = [0, 168]$  h). For a discrete dose  $d \in \mathcal{D}$  and subject-level covariates  $\zeta = (\text{BW}, \text{COMED}) \sim P$ , the learned predictor returns a PD trajectory  $y(t; d, \zeta)$  (ng/mL). With clinical threshold  $\tau = 3.3$  ng/mL, define the success event

$$\mathcal{S}(d, \zeta) = \left\{ \max_{t \in W} y(t; d, \zeta) \leq \tau \right\},$$

and the corresponding population attainment probability

$$p(d) = \mathbb{P}_{\zeta \sim P}(\mathcal{S}(d, \zeta)).$$

At inference the mapping  $(t, d, \zeta) \mapsto y(t; d, \zeta)$  uses fixed, trained weights, so evaluation is deterministic on CPU/GPU and fully auditable. When quantum resources are employed, they are confined to training or estimation subroutines with complete logs (circuit definitions, seeds, and measurement records). Two such variants are: (i) *quantum-kernel* models that replace affine maps by expansions in a Hilbert-space kernel

$$k(v, v') = |\langle \phi(v) | \phi(v') \rangle|^2,$$

precomputed and amortized across training; and (ii) *quantum-train (QT)* models that synthesise the weight tensors during training from low-dimensional quantum state statistics, after which inference remains purely

classical. These structures are interpretable and can be cross-validated against standard mechanistic baselines—e.g., inhibitory  $E_{\max}$ ,

$$E(C) = E_0 - E_{\max} \frac{C}{EC_{50} + C},$$

or indirect-response models—and replicated in NONMEM or nlmixr2 by targeting the same windowed endpoints  $p(d)$ .

Limitations mirror the Phase-1 data regime: small cohort size  $N$ , irregular observation times  $\{t_i\}$ , and a sparse dose grid  $\mathcal{D}$  increase extrapolation risk. QuSOP mitigates this by posing dose selection as a *chance-constrained* optimisation problem:

$$\text{find the minimal } d \in \mathcal{D} \text{ s.t. } p(d) \geq \pi, \quad \pi \in \{0.90, 0.75\},$$

and by propagating the covariate prior  $P$  (e.g., body-weight distribution, concomitant-medication prevalence) through the Monte-Carlo estimator  $\hat{p}(d)$  with conservative decision rules based on its lower  $(1 - \alpha)$ -confidence band.

Looking ahead, we will (i) quantify QuSOP’s empirical gains in precision and wall-clock efficiency, particularly where quantum amplitude-estimation-style routines reduce the width of uncertainty bands near decision thresholds; (ii) extend from single-endpoint suppression to multi-endpoint policies that balance PD control with exposure or safety caps; and (iii) integrate Bayesian assurance for Phase-2 dose-ranging designs. Overall, QuSOP provides a pragmatic pathway to introduce quantum technology into model-informed drug development while preserving traceability, interpretability, and clinical relevance.

## 9 CONCLUSION

This work introduces **QuSOP**, an end-to-end framework that pairs a QLSTM with QMC population expansion to deliver auditable, uncertainty-aware dose recommendations from sparse Phase 1-style PK/PD data. By confining quantum resources to training-time expressivity (via QLSTM) and probability estimation (via QMC and amplitude-estimation-style routines), while keeping inference deterministic on CPU/GPU, QuSOP yields regulator-aligned *attainment probabilities* over explicit 24 h and 168 h steady-state windows and supports chance-constrained optimisation on clinical dose grids. Across baseline and policy-restricted scenarios (e.g., COMED=0), the learned dose-success curves exhibit the expected sigmoid structure with coherent daily/weekly recommendations; relaxing the success criterion from 90% to 75% generates material, interpretable dose reductions, quantitatively linking decision thresholds to the local slope of the response surface. The framework is hardware-agnostic, reproducible (seeded simulations, fixed grids, covariate priors), and directly interoperable with pharmacometrics tooling, enabling cross-validation against mechanistic  $E_{\max}$ /indirect-response baselines and replication in NONMEM/nlmixr2.

Beyond methodological correctness, QuSOP is engineered for scale and translational impact. Its tileable evaluation workload maps naturally to modern AI supercomputers (e.g., H100-based DGX SuperPODs), enabling nation-scale, privacy-preserving runs with full audit trails and near-linear throughput across hundreds of GPUs. This capacity allows rapid scenario sweeps (covariate shifts, policy constraints, safety caps) and multi-endpoint extensions without altering clinical interfaces, while providing a clear on-ramp to future QPUs for variance-efficient probability estimation. Limitations remain those of early development—small  $N$ , irregular sampling, constrained dose support—yet the QLSTM+QMC design mitigates extrapolation risk through chance constraints and calibrated uncertainty bands rather than point predictions. Looking ahead, we anticipate three high-value directions: (i) joint policies that balance PD control with exposure/safety endpoints under explicit trade-offs; (ii) Bayesian assurance layers to quantify trial-level risk for Phase 2 dose-ranging; and (iii) progressive substitution of classical MC with amplitude-estimation modules where quantum hardware or high-fidelity simulators afford tangible variance reduction. Taken together, these advances position QuSOP as a pragmatic, forward-compatible route to quantum-enhanced, clinically relevant decision-making in model-informed drug development.

## 10 FUTURE WORK: QUANTUM-INFORMED PINN FOR PK-PD

We will extend the current QLSTM+QMC pipeline with a quantum-informed physics-informed neural network (PINN) that embeds mechanistic PK-PD structure into the learning objective. Let  $x(t)$  denote latent PK states (e.g., central and peripheral compartments) and  $y(t) = \text{PD}(t)$  the observed biomarker dynamics. A representative one-compartment with first-order absorption and elimination coupled to an indirect-response PD can be written as

$$\begin{aligned}\frac{dA_{\text{gut}}}{dt} &= -k_a A_{\text{gut}}, & \frac{dA_c}{dt} &= k_a A_{\text{gut}} - k_e A_c, & C(t) &= \frac{A_c(t)}{V}, \\ \frac{dy}{dt} &= k_{\text{in}}(1 - \mathcal{E}(C(t))) - k_{\text{out}}y(t), & \mathcal{E}(C) &= \frac{E_{\max}C^\gamma}{EC_{50}^\gamma + C^\gamma}.\end{aligned}$$

A PINN  $f_\theta$  will be trained to map inputs ( $t$ , covariates, regimen) to outputs  $(C_\theta(t), y_\theta(t))$  while penalizing violations of these differential constraints. The loss combines data fit and physics residuals,

$$\begin{aligned}\mathcal{L}_{\text{PINN}} &= \lambda_{\text{data}} \|y_\theta(t_i) - y_i\|_2^2 + \lambda_{\text{PK}} \left\| \frac{dA_{\text{gut},\theta}}{dt} + k_a A_{\text{gut},\theta} \right\|_2^2 \\ &\quad + \lambda_{\text{PK}} \left\| \frac{dA_{c,\theta}}{dt} - k_a A_{\text{gut},\theta} + k_e A_{c,\theta} \right\|_2^2 + \lambda_{\text{PD}} \left\| \frac{dy_\theta}{dt} - k_{\text{in}}(1 - \mathcal{E}(C_\theta)) + k_{\text{out}}y_\theta \right\|_2^2.\end{aligned}\tag{7}$$

supplemented with inequality and integral constraints for nonnegativity, mass balance, and dosing events. Gradients  $\frac{d}{dt}$  are computed by automatic differentiation over  $f_\theta$ .

To inject quantum inductive bias, we will replace selected subnetworks within the PINN by differentiable quantum layers  $q_\phi(\cdot)$  acting as implicit priors over hard-to-identify components. Two primary placements will be investigated: (i) a quantum surrogate for the exposure-response nonlinearity  $\mathcal{E}(\cdot)$ , i.e.,  $\mathcal{E}_\phi(C) = q_\phi(C)$  constrained to be monotone and saturating via parameterization and penalties; and (ii) a quantum regulariser on subject-specific parameter fields  $\vartheta(z)$  (e.g.,  $k_a, k_e, V, EC_{50}, E_{\max}, \gamma$ ) conditioned on covariates  $z$ , implemented by a variational quantum circuit that outputs low-dimensional embeddings subsequently decoded to  $\vartheta$ . The full model becomes

$$(C_\theta, y_\theta) = f_\theta(t, \text{dose}, z; \mathcal{E}_\phi, \vartheta_\phi(z)), \quad \min_{\theta, \phi} \mathcal{L}_{\text{PINN}} + \lambda_q \mathcal{R}_q(\phi),$$

where  $\mathcal{R}_q$  promotes smoothness, monotonicity, and boundedness of quantum outputs, and  $\lambda_q$  controls quantum regularisation strength.

Uncertainty quantification will be handled by a quantum-aware Monte-Carlo scheme. We will combine the existing QMC population expansion with stochastic PINN ensembles and weight-space perturbations of the quantum layer to approximate posterior predictive bands for  $\Pr[\max_t y_\theta(t) \leq \tau]$ . Multi-fidelity training will be explored by mixing sparse clinical PD observations with synthetic trajectories generated from mechanistic simulators under informative priors, down-weighting simulator bias via an adaptive fidelity weight  $\lambda_{\text{sim}}(t)$ . To address identifiability, we will include structural penalties and profile-likelihood style diagnostics inside training, and evaluate practical identifiability of  $(EC_{50}, E_{\max}, \gamma)$  under realistic sampling schedules.

Algorithmically, we will study three quantum designs: (i) variational quantum feature maps feeding a small classical head for  $\mathcal{E}_\phi$ ; (ii) quantum kernel regression to regularise exposure-response in data-poor strata; and (iii) a quantum Gaussian process prior over  $\vartheta_\phi(z)$  realised via random Fourier features generated by parameter-shifted circuits. Resource-aware ansätze will be prioritized (shallow depth, hardware-efficient entanglers) with gradient-stabilising tricks (layerwise learning rates, spectral norm constraints, and gradient clipping). Training will proceed in two stages: pretraining the classical PINN on physics losses to achieve feasibility, followed by joint fine-tuning with quantum components on the composite loss.

Benchmarking will compare: classical PINN, QLSTM+QMC, and the proposed quantum-informed PINN across (i) predictive error on held-out PD trajectories, (ii) calibration of attainment probabilities and sharpness of MC bands, (iii) robustness to covariate shift and missingness, and (iv) computational efficiency. Ablations will quantify the contribution of physics losses and each quantum insertion point. Prospective simulations will evaluate dose-selection stability under protocol variations (once-daily versus once-weekly, titration schedules) and threshold stress tests  $\tau \in [2.5, 4.0] \text{ ng/mL}$ .

Milestones include: (1) implement the PINN scaffold with differentiable dosing events and constraint handling; (2) integrate  $\mathcal{E}_\phi$  as a quantum surrogate with monotonicity and saturation guarantees; (3) add quantum parameter fields  $\vartheta_\phi(z)$  with population priors; (4) deliver uncertainty-calibrated dose-success curves under COMED = 0/1; and (5) package a reproducible evaluation suite with open-source code, synthetic datasets, and reporting templates for regulatory-style documentation.

## 11 SCALE-UP PLAN ON LARGE GPU CLUSTERS: ROADMAP, RISKS, AND ACCEPTANCE CRITERIA

### 11.1 EXECUTION PLAN: METHODS, OWNERSHIP, AND DELIVERABLES

**Guiding principle.** We treat dose selection as a chance-constrained, sequence-to-policy problem. The pipeline has three contract points: (i) a data-to-trajectory map  $f_\Theta$  (QLSTM/QT-LSTM/QK-LSTM or PINN) producing steady-state PD paths over 24/168 h; (ii) a probability engine that estimates  $p(d) = \Pr[\max_t \text{PD}(t) \leq \tau]$  with calibrated uncertainty (MC, QRNG, and AE where feasible); (iii) a discrete selector that returns the minimal dose meeting a target success level. Every phase produces signed, reproducible artifacts that can be re-run end-to-end from a single config.

**Phase I (Jul–Sep 30): Team formation & proposal.** **Objectives:** lock data quality, establish a strong classical baseline, and freeze the plan for review.

- **Data curation & split freeze (Owner: Lily):** NONMEM-style intake checks, de-duplication, time monotonicity, unit sanity; subject-wise splits with seeds. Artifact: dataset/EstData.csv, splits.yaml.
- **Baseline model (Owner: Louis):** LSTM (BiLSTM head) with standardised inputs  $(t/\tau, d/\text{BW}, \text{COMED})$ . Loss=MSE on standardised PD. Early stopping. Artifact: models/lstm\_qdqw.onnx, training report (MSE, calibration).
- **Population evaluator (Owner: Felix):** vectorized Monte Carlo over covariate priors with Sobol/QRNG streams; Clopper-Pearson CI; dose grids qd: [0.5:40], qw: [5:400]. Artifact: eval/qd\_qw-curves.parquet.
- **MLOps & reproducibility (Owner: Hasan):** container, pinned libs, deterministic flags; run registry capturing (git SHA, config, RNG seeds, hardware). Artifact: docker/Dockerfile, runs/registry.db.
- **Gate (M1):** submit proposal (Sep 30, 10:00 CET) with baseline curves and resource plan.

**Phase II (Oct 1–Nov 12): Evaluation & presentation.** **Objectives:** respond to feedback, demonstrate variance reduction, and present at EQTC.

- **Calibration & feedback pass (Owner: Lily):** calibration slope, reliability diagrams (binning sensitivity); daily/weekly coherence check (dose-normalised equivalence within 5%). Artifact: reports/calibration.pdf.
- **Quantum hooks (Owner: Louis):**
  - (Q1) *QAE prototype on GPU (cuQuantum/CUDA-Q)*: encode success oracle on coarse bins of  $\max_t \text{PD}(t)$ ; fixed-point AE at 20–24 qubits. KPI: CI half-width  $\leq 0.5 \times$  classical at matched wall-clock (M3).
  - (Q2) *QK-LSTM cache*: precompute quantum kernels for gate inputs; sharded NVMe cache with hit-rate SLO  $\geq 90\%$ ; Nystrom fallback if memory pressure (M4 throughput KPI).
- **Noisy quantum benchmarking (Owner: Felix/Sofia): Evaluate performance of quantum subroutines under different noise models to investigate potential robustness on real hardware.:**
- **Cost/perf tuning (Owner: Felix/Sofia):** BF16/TensorRT export, compile cache, bigger tiles; throughput target  $\geq 15\text{M}$  trajectories/min cluster-wide; p95 job latency  $< 1.2 \times$  median.
- **Presentation (Owners: Louis & team):** EQTC deck + live demo (dose-success sweep with CI controls); artifact and script hashes embedded in slides.

**Phase III (Q4 2025–Jan 31, 2026): Development & finalization.** **Objectives:** productionize, add multi-endpoint policy, freeze v1.0.

- **Multi-endpoint policy (Owner: Louis):** joint chance constraint on PD suppression and exposure cap (e.g.,  $C_{\max}$  or AUC); report Pareto and a recommended operating point. KPI: calibration slope in [0.9, 1.1], CI half-width  $\leq 2\%$  at decision doses.
- **Hardening (Owner: Hasan):** signed artifacts (SBOM, checksums), acceptance tests per GPU type, replay jobs for determinism; non-regression CI.

- **Documentation (Owner: Ross):** runbook (SOPs, RNG provenance, audit trail), API for dose queries, and governance checklist.
- **Gate (M6):** code & report freeze Jan 31 with `release-v1.0` tag, containers, and reproducibility pack.

#### Post-challenge (Feb–Mar 2026): OSS curation.

- Apache/MIT packaging, data license review, redaction of any confidential bits; tutorial notebooks; `CITATION.cff`; issue templates. Target: external re-run within 30 minutes from README.

#### Methodological details we will implement.

- (Q1) **Sequence predictor(s):** QLSTM (VQC-augmented gates), QT-LSTM (quantum-trained weight synthesis with classical inference), and QK-LSTM (quantum kernels in gate pre-activations). Shared training loop: BPTT + parameter-shift for quantum params; early stopping; z-scaling of PD. PINN track (future-work branch) enforces PK/PD ODE residuals as soft constraints.
- (Q2) **Uncertainty engine:** Vectorized MC with Sobol/QRNG; Clopper–Pearson or Wilson intervals; stratified batching near sigmoid knee doses; optional control variates using parametric surrogates (inhibitory  $E_{\max}$ ).
- (Q3) **Amplitude Estimation (AE):** Implement success oracle via table lookup of  $\max_t$  PD bins; fixed-point AE to avoid phase ambiguity; batched state-vector simulation on H100s; stopping rule based on empirical half-width.
- (Q4) **Discrete selector:** scan minimal feasible dose  $d^*(\pi)$  on grids (0.5 mg QD, 5 mg QW). For multi-endpoint, scalarize with regulator-aligned priority (PD first, then exposure), plus sensitivity appendix.
- (Q5) **HPC mapping:** embarrassingly parallel tiles over (dose, replicate); TensorRT/BF16 inference; on-device reduction of per-tile maxima and success counts; NCCL only for training; per-tile JSON manifest (config, seeds, checksum).
- (Q6) **Reproducibility:** single `.yaml` config controls data split, priors, grids, and RNG seeds; run registry records hashes (code, data, container, config); nightly replay on a held-out GPU pool; deterministic flags in cuDNN/cuBLAS.

#### Experimental protocol & metrics.

- **Fit & generalization:** MSE (std space), PD RMSE (ng/mL), temporal  $R^2$ ; daily/weekly coherence error  $< 5\%$ .
- **Calibration:** slope/intercept, ECE, reliability diagrams (10/20/50 bins); stratified by BW and COMED.
- **Decision quality:**  $d^*(\pi)$  distribution across MC replicates; CI half-width at crossing doses (target  $\leq 2\%$ ); sensitivity to priors (BW 70–140 kg; COMED=0).
- **Performance:** trajectories/min, GPU-hr per full sweep, p95 job latency; kernel cache hit rate; AE vs MC half-width per wall-clock.

#### Deliverables by milestone (acceptance criteria).

- **M1 (Sep 30):** Baseline curves; throughput  $\geq 5 \times$  CPU; registry on.
- **M2 (Oct 14):** Calibrated curves; slope in [0.9,1.1]; coherence  $< 5\%$ .
- **M3 (Oct):** AE half-width  $\leq 0.5 \times$  MC at matched time; demo notebook.
- **M4 (Nov):**  $\geq 15$ M traj/min; cache hit  $\geq 90\%$ ; cost report.
- **M5 (Dec):** Multi-endpoint policy; CI  $\leq 2\%$  at decisions.
- **M6 (Jan 31):** Release v1.0 (containers, SBOM, signed artifacts, report).

#### 11.2 TOP RISKS & MITIGATIONS

This risk register operationalizes the highest-impact threats to a GPU-cluster scale-up and ties each to concrete tripwires and pre-agreed responses. R1 (data sparsity) is monitored via CI band behavior and tail stability; if added compute doesn't shrink uncertainty, we inject physics via PINN priors, distill QT-LSTM

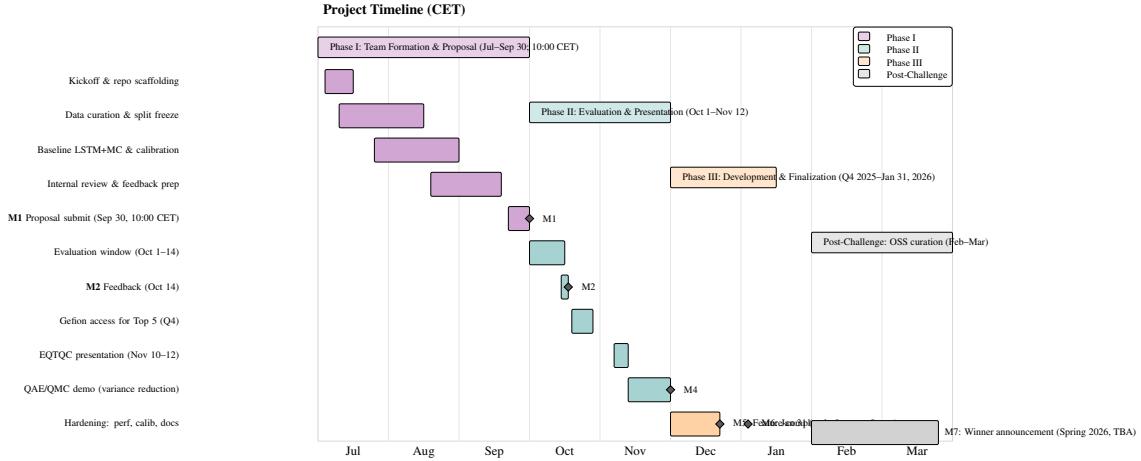


Figure 21: One-page Gantt aligned to the LSQI challenge: Jul–Sep (Phase I), Oct–Nov 12 (Phase II), Q4–Jan 31 (Phase III), and Feb–Mar (post-challenge OSS curation).

surrogates, broaden covariate priors, stratify Sobol/QRNG sampling, and bring in external validation. R2 (model misspecification) is watched through calibration slope, Hosmer–Lemeshow, and daily-weekly coherence; if it trips, we promote an ensemble (LSTM, QK-LSTM, PINN), enforce monotone  $E_{max}$  structure, and fall back to mechanistic gates for decisions. R3 (AE feasibility) has hardware/simulator and variance KPIs; when AE exceeds qubit/simulator budgets or fails to beat MC, we switch to batched fixed-point AE with compressed oracles—otherwise revert to variance-reduced MC with control variates. R4 (kernel cache pressure) is tracked by hit rate and latency; we tier storage (HBM to NVMe), cap reference sets, approximate via Nyström/RFF, and pre-warm popular tiles to meet SLOs. R5 (regulatory acceptance) is governed by reproducibility findings; we keep inference purely classical, freeze seeds/configs, ship signed artifacts with RNG provenance, and run concordant NONMEM/nlmixr2 shadows. R6 (reproducibility at scale) is guarded by hash/config parity checks and deterministic kernels; pinned containers, a run registry (config+git+seed), periodic replay jobs, and GPU-specific acceptance tests ensure results are bit-stable and audit-ready.

#	Risk	Trigger / Early signal	Mitigation / Contingency
R1	<b>Data sparsity at scale</b>	CI bands fail to contract despite more compute; unstable tail estimates	Hybrid modeling: PINN priors + QT-LSTM distillation; widen covariate priors; stratified Sobol/QRNG; adapt window endpoints; invest in external validation sets.
R2	<b>Model misspecification</b>	Calibration slope $< 0.9$ or $> 1.1$ ; Hosmer–Lemeshow fail; poor weekly $\leftrightarrow$ daily coherence	Model ensemble (LSTM, QK-LSTM, PINN linkage); monotone $E_{\max}$ constraints; stress protocols; escalate to mechanistic fallback for gating decisions.
R3	<b>AE feasibility/scale</b>	AE qubits $>$ simulator sweet spot; oracle tabulation too large; variance not improving	Batched AE (fixed-point, iterative QPE); compress oracle (coarse bins); switch to variance-reduction (QMC+control variates) if AE ROI $< 2\times$ .
R4	<b>Kernel cache size/latency</b>	Sharded cache $>$ memory budget; eviction thrash; cold misses spike latency	Tiered cache (GPU HBM $\rightarrow$ NVMe); approximate kernels via Nystrom/RFF; bound reference set; pre-warm popular tiles; cache hit SLOs.
R5	<b>Regulatory acceptance</b>	Audit finding on reproducibility/traceability; unclear GxP stance on quantum training	Keep inference purely classical; freeze seeds/configs; signed artifacts; SOPs for RNG provenance; shadow NON-MEM/nlmixr2 analyses for concordance.
R6	<b>Reproducibility at scale</b>	Hash/config drift across nodes; non-deterministic kernels	Deterministic flags, pinned libs, container images; run registry (config+git+seed); periodic replay jobs; acceptance tests per GPU type.

Table 7: Risk register for GPU-cluster scale-up.

### 11.3 MILESTONES & ACCEPTANCE CRITERIA

The milestone plan defines measurable, audit-ready gates from enablement to release. **M1** verifies portability and numerical fidelity by achieving at least  $5 \times$  CPU throughput on H100 with BF16/TensorRT while holding numerical drift below  $10^{-3}$  and maintaining end-to-end parity for  $qd/qw$ . **M2** establishes reproducibility: the containerized pipeline and run registry must yield bitwise-stable outputs under fixed seeds/configs with CI half-width variability  $< 0.3\%$ . **M3** demonstrates the practicality of quantum amplitude estimation (QAE)—at matched wall-clock time, AE must reduce the CI half-width for  $p(d)$  near decision doses to  $\leq 0.5 \times$  that of classical MC, accompanied by a cost/benefit report. **M4** targets scale efficiency, requiring  $\geq 15M$  trajectories/min cluster-wide, GPU utilization  $\geq 85\%$ , p95 job latency  $< 1.2 \times$  the median, and kernel-cache hit rate  $\geq 90\%$ . **M5** elevates decision quality for joint PD + exposure policies, with calibration slope in  $[0.9, 1.1]$ , held-out AUC  $\geq 0.80$ , CI half-width  $\leq 2.0\%$  at decision doses, and daily/weekly coherence error  $< 5\%$ . Finally, **M6** delivers a production-grade release with signed artifacts and a complete reproducibility pack (configs, seeds, hashes), governance checklist passed, permissive OSS licensing (Apache/MIT) with necessary redactions, and external-validation agreement within pre-specified bounds.

Milestone	Deliverable	Acceptance criteria
<b>M1: Baseline@H100</b>	Ported inference on H100 w/ BF16/TensorRT	Throughput $\geq 5 \times$ CPU baseline; GPU util $\geq 70\%$ ; numerical drift $< 10^{-3}$ vs FP32; end-to-end $qd/qw$ parity.
<b>M2: Repro v1</b>	Containerized pipeline + run registry	Bitwise-stable outputs with fixed seed/config; run hash recorded; 3-run CI half-width change $< 0.3\%$ .
<b>M3: QAE demo</b>	CUDA-Q AE prototype (20–24 qubits)	At matched wall-clock, AE CI half-width $\leq 0.5 \times$ classical MC for $p(d)$ near threshold; cost/benefit report.
<b>M4: Throughput</b>	optimised array jobs + cache	$\geq 15M$ trajectories/min cluster-wide; GPU util $\geq 85\%$ ; p95 job latency $< 1.2 \times$ median; kernel cache hit $\geq 90\%$ .
<b>M5: Multi-endpoint</b>	Joint PD + exposure cap policy	Calibration slope in $[0.9, 1.1]$ ; AUC $\geq 0.80$ on held-out; CI half-width $\leq 2.0\%$ at decision doses; daily/weekly coherence error $< 5\%$ .
<b>M6: Release v1.0</b>	Signed artifacts, report, OSS	Repro pack (config+seeds+hashes); governance checklist passed; OSS (Apache/MIT) with redactions; external validation agreement within pre-specified bounds.

Table 8: Acceptance criteria aligned to scale-up milestones.

**Notes.** We retain a regulator-aligned posture: inference remains deterministic and classical; quantum modules (AE, QT/QK layers) are confined to training/estimation with full circuit/config provenance. CI half-width targets (2%) are evaluated at decision-relevant doses; calibration uses slope and reliability diagrams with binning sensitivity checks. Cost/perf KPI is reported as *trajectories/min* and *GPU-hours per full sweep*.

## REFERENCES

- Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.
- Sergey Bravyi, Oliver Dial, Jay M Gambetta, Darío Gil, and Zaira Nazario. The future of quantum computing with superconducting qubits. *Journal of Applied Physics*, 132(16), 2022.
- Matthias C Caro, Hsin-Yuan Huang, Marco Cerezo, Kunal Sharma, Andrew Sornborger, Lukasz Cincio, and Patrick J Coles. Generalization in quantum machine learning from few training data. *Nature communications*, 13(1):4919, 2022.
- Marco Cerezo, Guillaume Verdon, Hsin-Yuan Huang, Lukasz Cincio, and Patrick J Coles. Challenges and opportunities in quantum machine learning. *Nature Computational Science*, 2(9):567–576, 2022.
- Samuel Yen-Chi Chen, Chao-Han Huck Yang, Jun Qi, Pin-Yu Chen, Xiaoli Ma, and Hsi-Sheng Goan. Variational quantum circuits for deep reinforcement learning. *IEEE access*, 8:141007–141024, 2020.
- Samuel Yen-Chi Chen, Shinjae Yoo, and Yao-Lung L Fang. Quantum long short-term memory. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8622–8626. IEEE, 2022a.
- Samuel Yen-Chi Chen, Shinjae Yoo, and Yao-Lung L Fang. Quantum long short-term memory. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8622–8626. IEEE, 2022b.
- Hakan Doga, Aritra Bose, M. Emre Sahin, Joao Bettencourt-Silva, Anh Pham, Eunyoung Kim, Alan Andress, Sudhir Saxena, Laxmi Parida, Jan Lukas Robertus, Hideaki Kawaguchi, Radwa Soliman, and Daniel Blankenberg. How can quantum computing be applied in clinical trial design and optimization? *Trends in Pharmacological Sciences*, 2024.
- Jay Gambetta. Quantum-centric supercomputing: The next wave of computing. *IBM Research Blog*, 2022.
- Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
- Yu-Chao Hsu, Nan-Yow Chen, Tai-Yu Li, Po-Heng Henry Lee, and Kuan-Cheng Chen. Quantum kernel-based long short-term memory for climate time-series forecasting. In *2025 International Conference on Quantum Communications, Networking, and Computing (QCNC)*, pp. 421–426. IEEE, 2025a.
- Yu-Chao Hsu, Jiun-Cheng Jiang, Chun-Hua Lin, Wei-Ting Chen, Kuo-Chung Peng, Prayag Tiwari, Samuel Yen-Chi Chen, and En-Jui Kuo. Federated quantum kernel-based long short-term memory for human activity recognition. *arXiv preprint arXiv:2508.06078*, 2025b.
- Yu-Chao Hsu, Tai-Yu Li, and Kuan-Cheng Chen. Quantum kernel-based long short-term memory. In *2025 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*, pp. 1–5. IEEE, 2025c.
- Hsin-Yuan Huang, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, and Jarrod R McClean. Power of data in quantum machine learning. *Nature communications*, 12(1):2631, 2021.
- Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Ryan Babbush, Richard Kueng, John Preskill, et al. Quantum advantage in learning from experiments. *Science*, 376(6598):1182–1186, 2022.
- Chu-Hsuan Abraham Lin, Chen-Yu Liu, and Kuan-Cheng Chen. Quantum-train long short-term memory: Application on flood prediction problem. *arXiv preprint arXiv:2407.08617*, 2024.
- Chen-Yu Liu, En-Jui Kuo, Chu-Hsuan Abraham Lin, Jason Gemsun Young, Yeong-Jar Chang, Min-Hsiu Hsieh, and Hsi-Sheng Goan. Quantum-train: Rethinking hybrid quantum-classical machine learning in the model compression perspective. *arXiv preprint arXiv:2405.11304*, 2024a.

Chen-Yu Liu, Chu-Hsuan Abraham Lin, Chao-Han Huck Yang, Kuan-Cheng Chen, and Min-Hsiu Hsieh. Qtrl: Toward practical quantum reinforcement learning via quantum-train. *arXiv preprint arXiv:2407.06103*, 2024b.

Junhua Liu, Kwan Hui Lim, Kristin L Wood, Wei Huang, Chu Guo, and He-Liang Huang. Hybrid quantum-classical convolutional neural networks. *Science China Physics, Mechanics & Astronomy*, 64(9):290311, 2021a.

Xiangyu Liu, Chao Liu, Ruihao Huang, Hao Zhu, Qi Liu, Sunanda Mitra, and Yaning Wang. Long short-term memory recurrent neural network for pharmacokinetic-pharmacodynamic modeling. *International journal of clinical pharmacology and therapeutics*, 59(2):138, 2021b.

Andrea Mari, Thomas R Bromley, Josh Izaac, Maria Schuld, and Nathan Killoran. Transfer learning in hybrid classical-quantum neural networks. *Quantum*, 4:340, 2020.

S.S. Negus and M.L. Banks. Pharmacokinetic–pharmacodynamic (pkpd) analysis with drug discrimination. In *The Behavioral Neuroscience of Drug Discrimination*. Springer, 2016.

Adrián Pérez-Salinas, Alba Cervera-Lierta, Elies Gil-Fuster, and José I Latorre. Data re-uploading for a universal quantum classifier. *Quantum*, 4:226, 2020.

John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018.

Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Physical Review A*, 103(3):032430, 2021.

Joseph Standing. Understanding and applying pharmacometric modelling and simulation in clinical practice and research. *British Journal of Clinical Pharmacology*, 2016.

Mizuki Uno, Yuta Nakamaru, and Fumiyo Yamashita. Application of machine learning techniques in population pharmacokinetics/pharmacodynamics modeling. *Drug Metabolism and Pharmacokinetics*, 2024.