Supplementary Information

Title: InSilico-K digital twin platform for optimizing neoadjuvant regimens and surgical de-escalation in lung cancer

Author: Junichi Shimada

S1. Dimensional Analysis and Empirical Scaling Coefficients

The computational pipeline of the InSilico-K engine operates on a multi-scale biophysical framework derived from the Starling equation governing capillary fluid shifts and Darcy's law for interstitial flow. To ensure mathematical stability and physical plausibility, all empirical coefficients (k1, k2, k3, k4, and k_telmisartan) were subjected to strict dimensional analysis.

Table S1: Master Array of Variables and Calibrated Coefficients

Parameter	Symbol	Unit	Value	Derivation Source / Biophysical Justification
Tumor Volume	V	cm3	4.2	Median resectable target volume (Ref 17: Okami et al., 2019).
Mean Arterial Pressure	MAP	mmHg	105.0	Standard systemic perfusion baseline.
Serum Albumin	Alb	g/dL	3.8	Normal host intravascular oncotic baseline.
Volume-Pressure Scaler	k1	mmHg/cm3	0.85	Solid stress expansion bounds governing cellular density.
Filtration Scaler	k2	Dimensionless	0.15	Systemic-to-capillary pressure transmission ratio (Starling forces).
Oncotic Buffer	k3	mmHg/(g/dL)	4.20	Colloid osmotic

Parameter	Symbol	Unit	Value	Derivation Source / Biophysical Justification
Scaler				pressure conversion factor.
Decompression Rate	k_telmisartan	day-1	0.18	Biological half-life of vascular decompression (Ref 6: Chauhan et al., 2013).
Diffusion Sensitivity	k4	mmHg-1	0.08	Empirical tuning to match macromolecular penetration limits under high IFP.

Dimensional Integrity in Equation One (Baseline IFP)

The target parameter Baseline_IFP is measured in absolute pressure (mmHg). The dimensional integrity of the macro-environmental input function is validated as follows:

(k1 * V) --> (mmHg/cm3) * cm3 = mmHg

(k2 * MAP) --> Dimensionless * mmHg = mmHg

(k3 * Alb) --> (mmHg/(g/dL)) * g/dL = mmHg

The linear summation of these vectors rigorously satisfies the output dimension of mmHg, physically representing the balance between solid stress, hydrostatic filtration, and intravascular oncotic absorption.

S2. Spatial Grid Matrix (B_j,k) Initialization

To calculate the spatial distribution and cumulative efficacy of the therapeutic agents (Equation Five), the simulated tumor microenvironment is discretized into a 3D matrix, denoted as B_j,k.

- Grid Resolution: The matrix utilizes isotropic voxels with a spatial resolution of 1 mm3 per cell, providing an optimal balance between micro-scale biological fidelity and macro-scale tensor computation efficiency.
- Initial Carrying Capacity: The baseline cell density (carrying capacity) at day 0 is initialized with a radial gradient. The solid core is assigned maximum mechanical density, while the peripheral margins simulate lower-density cellular diffusion boundaries.

- Dynamic Update: The matrix is recalculated iteratively over the 28-day window to track the spatio-temporal reduction in carrying capacity as the local Effective_Occupancy score accumulates, corresponding to the trajectories represented in Figure 2.

S3. Computational Infrastructure and Deployment Protocol

To guarantee the absolute reproducibility of the high-performance multi-threaded tensor calculations required for clinical digital twins, the InSilico-K platform was deployed on a localized, strictly containerized architecture. This protocol eschews variable cloud environments in favor of a robust, deterministic three-layer physical-to-virtual deployment standard.

Hardware Specifications (Layer 1)

The primary execution host (designated locally as "Aurora7920") requires a high-end workstation configuration tailored for intensive computational oncology workloads:

- Processor: Dual-core Intel Xeon scalable architecture.
- Memory: 384 GB of physical ECC RAM.
- GPU Acceleration: NVIDIA RTX PRO 5000 providing 48 GB of dedicated VRAM.
- Persistent Storage: A 44 TB high-capacity physical storage array, formatted as NTFS (GPT partition). All project scripts, AlphaFold 3 coordinates, and transient state outputs are strictly mapped natively to an isolated drive (Z: drive) to prevent permission degradation and ensure data persistence across operating system boundaries.

Transport and Interface Layer (Layer 2)

The communication bridge is managed via Windows Subsystem for Linux 2 (WSL2), running an Ubuntu 22.04 LTS distribution. The physical host drive is securely mounted at /mnt/z/ within the Linux environment.

Containerized Execution Protocol (Layer 3)

The execution engine is encapsulated within a secure Docker container. To eliminate dependency on unstable external package servers and guarantee exact local compilation, the build protocol enforces the following localized rules:

1. Repository Optimization (Mandatory Localized Mirror):

Standard Ubuntu archive endpoints must be systematically replaced with a highly stable domestic mirror (e.g., the Japan Advanced Institute of Science and Technology [JAIST] mirror) to ensure rapid, secure, and reproducible package retrieval.

Plaintext

```
RUN sed -i 's@archive.ubuntu.com@ftp.jaist.ac.jp@g' /etc/apt/sources.list ¥
    && sed -i 's@security.ubuntu.com@ftp.jaist.ac.jp@g' /etc/apt/sources.list ¥
    && apt-get update
```

2. Framework Stack and Hardware Compilation:

- Base Image: Official NVIDIA CUDA 12.8 runtime environment.

- Python Stack: Python 3.10.
- Tensor Framework: PyTorch 2.2+ must be compiled natively for the sm_120 architecture to maximize the parallel processing efficiency of the target RTX PRO 5000 tensor cores.
- Array Operations: NumPy is strictly pinned to v1.26 (maintaining sub-2.0 versions) to prevent multi-dimensional array method conflicts during real-time metric generation.

3. Graphical Interface and Localization:

The interactive Streamlit dashboard is rendered securely via X11 display communication across multi-monitor host setups. The fonts-noto-cjk package is embedded within the container image to ensure localized text rendering without host-system dependency.

Plaintext

```
xhost +local:docker
sudo docker run --runtime nvidia -it --rm --network host ¥
  -v /mnt/z/InSilico_System:/workspace ¥
  -v /tmp/.X11-unix:/tmp/.X11-unix ¥
  -e DISPLAY=:0 ¥
  insilico-k-engine:latest
```

4. State Persistence and Container Management:

To protect the integrity of the pure Linux environment, all session configurations and matrix state updates are preserved through a strict preservation policy. Following any destructive testing or environmental shutdown, the active container must be preserved via the docker commit command, and the FROM instruction within the baseline Dockerfile updated to ensure zero loss of computational progress.