# Likelihood-free inference for hybrid cellular automaton models for personalized simulation of breast cancer treatment

Xiaoran Lai[1]*, Henri Pesonen [1], Alvaro Köhn-Luque [1], Samuel Kaski [2], Jukka Corander[1], and Arnoldo Frigessi [1]

1 Oslo Centre for Biostatistics and Epidemiology, Faculty of Medicine, University of Oslo, Oslo, Norway
2 Helsinki Institute for Information Technology HIIT, Department of Computer Science, Aalto University, Helsinki, Finland
*xiaoran.lai@medisin.uio.no

## INTRODUCTION

• Mathematical modelling and simulation can be used to provide clinically relevant predictions of response to cancer therapy. However, the complexity and heterogeneity of cancer make it difficult to create and parameterise mechanistic models that allow for precise patient-specific predictions and treatment protocol optimisation.

• More detailed and realistic multi-scale cancer models can integrate multi-type clinical data[1], but are difficult to personalize both because the necessary measurements are sometimes not accessible and because analytical inference based on the likelihood on such complex models is impossible.

• We focus on a complex multi-scale hybrid cellular automaton model of breast cancer treated by a combination of chemotherapy and anti-angiogenic agents[1]. Several dependent processes, such as cell dynamics, the blood vessel birth and death process, and the processes governing concentrations of oxygen, VEGF and drugs in the tissue are only partially observable.

• We investigate if some of those key parameters that have a big impact on the treatment outcome can be estimated from a series of measurements of cell density in the tumour tissue using Bayesian optimization for likelihood-free inference (BOLFI)[2], as well as how often the measurements need to be taken.

• This work outlines the measures that should be done in clinical practise to ensure that enough patient data is accessible to run reliable tailored therapy simulations.

## METHODS

The system is partly initialised and tuned based on a biopsy from a real patient, but we use only simulated data to demonstrate the inference and prediction procedure.

The model consists of a hybrid cellular automaton model that couples stochastic and discrete model formalisms with deterministic and continuous components accounting for biological processes at different spatio-temporal scales (Fig. 1)

we focus on the inference of three key parameters determining the outcome of each patient to the drug treatment: chemo-sensitivity of cancer cells, α, minimal cell cycle length of cancer cells $T_c$ and birth probability of vessels based on the evolution of state $x_t$.

$$x_t = f_t(\alpha, T_c, p_{birth}, x_{t-1}, v_t),\ t = 0, 2, \ldots, T$$

where $f_t(.)$ is the nonlinear transition model at t, and $v_t$ is the stochastic component of the simulator.

$x_t$ consists of cells, vessels, and extracellular concentration of oxygen and VEGF within the simulation grid. Simulated data were generated using a set of fixed parameter values and assumed to be collected every three days, t = 0, 3, 6, 9.

Prior distributions for the unknown parameters were modelled as uniform distribution. Summary statistics were cancer cell and vessel densities, and VEGF distribution. The results are approximations to the posterior distribution of the parameters $p(\alpha, T_c, p_{birth} | y_{t0}, \ldots y_{t3})$

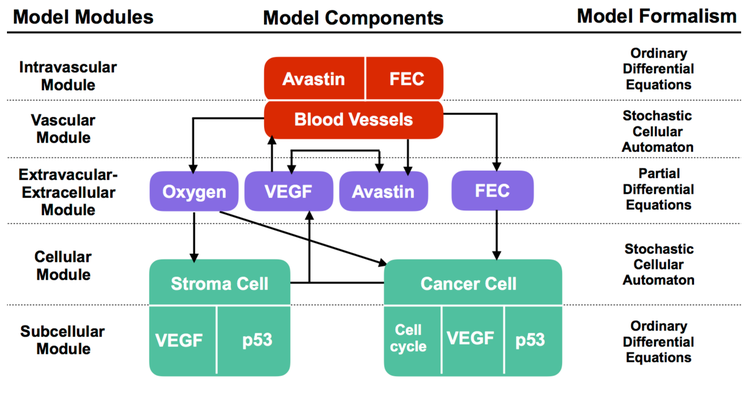
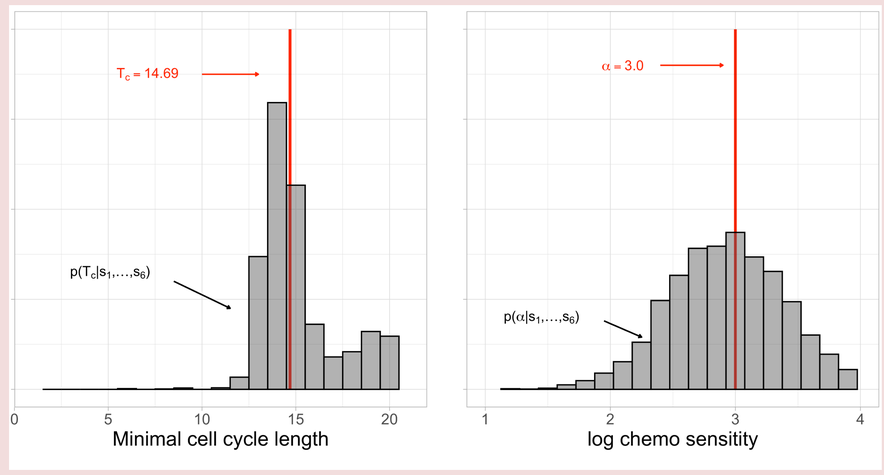



Fig. 1 Model summary of the simulator. The diagram shows the main components of each module and the interactions among them. The right column shows the different model formalisms used for each of the model modules.

## RESULTS

### Simulation scenario 1

Fig 2. Approximated posterior distributions of the posterior marginals. The red vertical lines on the first two subplots indicate the true simulation values which were used for simulating the cancer cell growth trajectory



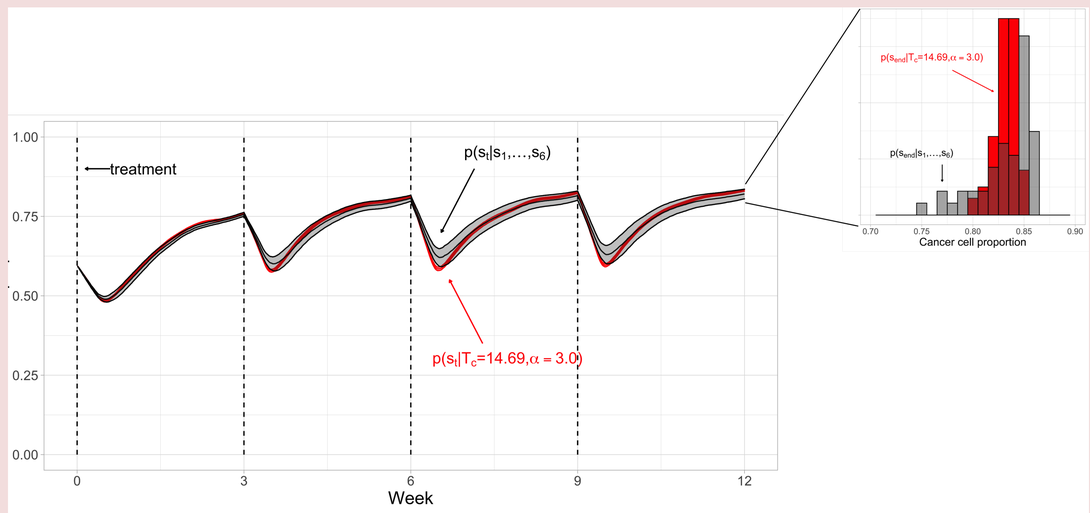




Fig 3. Simulated trajectories of the evolution of cancer cell proportions in the simulation grid given the true parameter values, and the forecast of the trajectory as simulated given the posterior distribution of the parameter values. The simulation end point is highlighted as the histogram on the right

### Simulation scenario 2

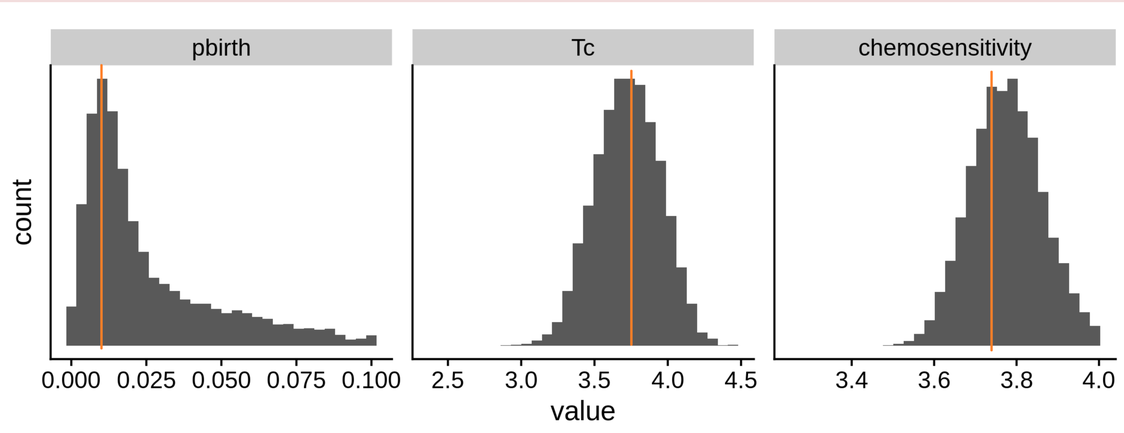



Fig 4. Approximated posterior distributions of the posterior marginals. The red vertical lines indicate true simulation values which were used for simulating the cancer cell growth trajectory

## CONCLUSION

• Our results were based on simulated data as real-patient data are only available at the start and the end of a therapy.
• From the simulated complete 12-week therapy history of two distinct patients, we can extract data at different interval to simulate various data collection intensities.
• For each collection intensity, we can test if we were able to estimate the personal parameters of the patient well enough and make good prediction of outcomes after therapy.

## REFERENCE

1. Lai, X, et al. "Toward personalized computer simulation of breast cancer treatment: A multiscale pharmacokinetic and pharmacodynamic model informed by multitype patient data." Cancer research 79.16 (2019): 4293-4304
2. Gutmann, M. U., & Corander, J. (2016). Bayesian optimization for likelihood-free inference of simulator-based statistical models. Journal of Machine Learning Research.
3. Pesonen, H., Simola, U., Köhn-Luque, A., Vuollekoski, H., Lai, X., Frigessi, A., ... & Corander, J. (2021). ABC of the Future. arXiv preprint arXiv:2112.12841.
4. Lintusaari, J., Vuollekoski, H., Kangasraasio, A., Skytén, K., Jarvenpaa, M., Marttinen, P., ... & Kaski, S. (2018). Elfi: Engine for likelihood-free inference. Journal of Machine Learning Research, 19(16), 1-7.