

Håkon Austlid Taskén

# Scalable Breast Cancer Simulations

A Stochastic, Hybrid Cellular Automaton Model  
for Precision Medicine

Master's thesis in Applied Physics and Mathematics

Supervisor: Jon Andreas Støvneng

Co-supervisor: Arnaldo Frigessi

December 2022



Håkon Austlid Taskén

# Scalable Breast Cancer Simulations

A Stochastic, Hybrid Cellular Automaton Model for  
Precision Medicine

Master's thesis in Applied Physics and Mathematics  
Supervisor: Jon Andreas Støvneng  
Co-supervisor: Arnaldo Frigessi  
December 2022

Norwegian University of Science and Technology  
Faculty of Natural Sciences  
Department of Physics



Norwegian University of  
Science and Technology





# Preface

This thesis is submitted as the final part and conclusion of my master's degree in Applied Physics and Mathematics at the Norwegian University of Science and Technology (NTNU). The thesis centers around a peer-reviewed scientific paper published in January of 2022, where I am shared first author. The paper is processed by an introduction that provides background information and motivation for the work.

## Acknowledgements

First and foremost I would like to extend a huge "thank you" to Arnaldo Frigessi for the support given over several years. It is especially appreciated in hard times. Without him this thesis would not have been possible. I would also like to express my gratitude to my internal supervisor Jon Andreas Støvneng for facilitating this thesis. I know I join a very long list of student at NTNU who is appreciative of the extra effort you have made for them. My co-authors does of course deserve a thank you, and specially Xiaoran (Richard) Lai and Alvaro Köhn-Luque for many fruitful discussions. Finally, my family deserves a very special thank you for the support over the past years.

# Table of Contents

|   |           |
|---|-----------|
| <b>Preface</b>  | <b>i</b>  |
| <b>1 Introduction</b>   | <b>1</b>  |
| 1.1 Background and Motivation . . . . .   | 1         |
| 1.2 Personalized Treatment . . . . .  | 2         |
| 1.3 Modeling and Simulating Treatment . . . . .   | 4         |
| 1.4 A Breast Cancer Simulation Needing Improvements . . . . .   | 5         |
| <b>Bibliografi</b>  | <b>7</b>  |
| <b>2 Paper</b>  | <b>11</b> |
| Sammendraget fra artikklen . . . . .  | 11        |
| Paper: A scalable solver for a stochastic, hybrid cellular automaton model<br>of personalized breast cancer therapy . . . . . | 12        |

# 1 | Introduction

## 1.1 Background and Motivation

Cancer is the second most common cause of death in the world [1, 2], with almost 10.0 million people dying from the disease in 2020 [3]. That equates to roughly one in six deaths being caused by cancer globally. Cancer incidences are furthermore estimated to increase by 47% within 2040, as the world population increases and ages [4]. The number of new cancer cases each year will then go up from 19.3 million in 2020 to 28.4 million in 2040 [4].

With such enormous ramifications, there is no surprise that a huge effort has been put into improving the treatment of cancer. U.S. National Cancer Institute alone has spent over 150 billion dollar on cancer research since its establishment in 1937 [5]. This effort has led to substantial improvements in cancer screening, diagnosis and treatment. One strong indication of the difference these improvements have made can be found in one of the oldest national cancer registries, The Cancer Registry of Norway, which shows that the 5-year relative survival in Norway has steadily increased from 34% in 1965 to 77% in 2021 [6].<sup>1</sup> This trend is also representative for other developed countries such as The United States, England, Wales and Australia, although the curves differs somewhat for the available year [9–11].

Still, a 77% 5-year relative survival means that we fail to provide just under a quarter of all cancer patients, that would not have died from other causes, with effective enough treatment to be alive five years after their diagnosis. Cancer treatment is simply ineffective for a lot of patients. With all the resources that have been put into curing cancer, the question of why we have not made even greater improvements have been asked many times. Cancer have indeed turned out to be very difficult to cure, and understanding why is useful for progressing its treatment. Let us start by exploring some of the answer to this question by looking at what cancer actually is, its definition.

Cancer is a malignant neoplasm. A neoplasm is a collection of cells in the body

---

<sup>1</sup>Note that this number aggregate all types of cancer and does not account for the changing prevalence of risk factors nor how earlier diagnosis could improve 5-year survival compared to longer term survival [7, 8]. It is thus only an indication of the effect of these improvements and not a direct display of it.

with changes in their genome which causes abnormal cell growth. If a neoplasm has the ability to intrude into neighbouring tissue or spread to other parts of the body (metastasizing), it is called malignant. This malignant property is what makes cancer lethal. Meanwhile, the fact that cancer can result from any DNA changes that causes abnormal cell growth in any of the body's cells, turns out to make it incredibly challenging to treat in general as the sheer variety of cancers is enormous. This has gotten increasing attention over the past decades as more and more of this diversity has been uncovered. Cancer is not a single disease, but a group of more than 200 diseases or types of cancer—corresponding to the different types of cells in the body it initially affects [12]. As the cost of DNA sequencing have plummeted [13], we have discovered more and more of the huge genetic variety within each of these types of cancer through projects such as The Cancer Genome Atlas [14]. This work has also unveiled a substantial genetic diversity within individual tumors (intra-tumor genetic heterogeneity), both spatially and temporally as the cancer evolves, in addition to the diversity between tumors. That the cancers consists of the body's own cells also constitutes a challenge, as any treatment must deal with the problem of distinguishing between cancer and normal cells—getting rid of one without harming the other too much. It also contributes to more diversity of interaction with the body's signaling pathways, and makes the problem considerable more complex.

## 1.2 Personalized Treatment

With the tremendous variety of cancers, a very natural way of improving its treatment is to find a better selection of therapies for the particular variety of cancer the patient has. This has indeed been a persistent trend in cancer therapy—patients being divided into more and more groups by improved diagnostics and given complex combinations of advanced treatments that have proven more effective for that particular group.

As a fairly representative example of how individualized cancer treatment has become, we can look at treatment in Norway for breast cancer, which is the most common cancer type in the world [4]. Norway is one of several countries that have introduced national standardised cancer care pathways, and since 2015 patients in Norway have followed national action programs for diagnosis and treatment of 28 different types of cancer [15, 16]. These programs provides an excellent overview of the current recommended treatments. The therapy for breast cancer can be divided into local treatments, consisting of surgery and/or radiation, and systemic treatments, consisting of drug therapies such as chemo-, hormonal and targeted therapies. The program assigns the patient into groups with associated sets of

recommended treatments based on TNM-classification, which consists of tumor size, presence and extent of regional lymph node metastases and presence of distant metastases, along with histological grading, hormone receptor status (HR), human epidermal growth factor receptor 2 status (HER2), proliferation index (Ki67) and gene panel tests [17]. The patients given therapeutic treatment are first divided into initially inoperable and operable. The operable are further divided into 8 groups for local treatment and 26 groups for adjuvant systematic treatment. Furthermore, the initially inoperable are divided into 4 groups for neoadjuvant systemic treatment and 16 groups for the local treatment potentially given afterwards. The treatment for some of the groups are no treatment. The 5-year relative survival in Norway of breast cancer has reached 92% with the use of this treatment scheme [6], which likely is among the very highest in the world [18–20].

This exemplifies the need for individualization in the treatment of cancer. Such division of patients into groups with different therapies is the start of what is called personalized or precision medicine, which works towards tailoring the treatment to the individual characteristics of each patient. Precision medicine is an approach that can be used for any disease. There is, however, reason to believe that precision medicine will provide greater improvements in effectiveness of treatment in cancer than in most other diseases, because, as discussed above, cancer have enormous variations and providing a large degree of differentiated treatment based on the patients characteristics has already proven very successful.

Providing each cancer patient with their own optimal combination of treatments is a lofty goal. For the systematic treatment with drugs, there is a combinatorial growth of possible combinations of different drugs, dosages and timing of administration for each set of biomarkers the patient may have. The interaction of different combinations of drug treatments can further be very hard to predict in some circumstances [21, 22]. So, to come close to reaching this goal, there is a real need to develop a way to help the clinician in charge to select the best treatment for their individual patients. It would be beyond reach for even a team of experts to find this optimum on their own with the incredible complex interactions between multiple drugs, tumors and the human body, in addition to the tremendous variety of cancers.

This need for improved guidance on selecting treatment as the personalization increases is also present in just the same way when it comes to *in vivo* testing. This includes the gold standard of randomized clinical trials in a fairly straightforward way—comparing a few treatments to find the one that is best on average for each of the groups of patients with some common combination of biomarkers. There is a limit to how many treatment regimens can be tested in this way, because the group sizes will become too small to give significant insight. Personalized *in vitro* models, such as cells from a patient’s biopsy grown in a culture that mimics the *in*

*vivo* environment, could provide insight into which treatment that specific patient would respond best to out of a much larger set of treatments [23]. While this would undoubtedly give very valuable data and give great improvement, the combinatorial growth of treatments with possible significant differences gives way too many possibilities for even *in vitro* testing to conquer.

## 1.3 Modeling and Simulating Treatment

With both *in vivo* and *in vitro* testing arriving short of the desired goal of providing each patient with their optimal treatment on their own, it is natural to explore how *in silico* testing (i.e., computer simulations) could help to make progress. By making a mathematical model of cancer that can be handled by a computer and providing it with sufficient input data, such *in silico* methods can make predictions of how the cancer will develop under a treatment regimen. This is already done with great success for complex systems in other domains including weather forecasting or molecular modeling in drug design [24, 25]. With such predictions for many different treatments, the computer can hone in on the optimal treatment for a specific patient and their input values. In essence a digital representation of the patient's cancer is created, a digital twin of the cancer, on which experiments can be performed to discover what will happen. The goal of precision medicine can in other words be greatly aided, if one can develop a mathematical model that makes accurate predictions of the outcome of a treatment based on the biomarkers of an individual patient.

While *in silico* experiments will always have to be verified *in vivo*, it does offer some key advantages. As discussed above, a major limitation of *in vivo* and *in vitro* experiments when it comes to personalized medicine is the number of experiments that can be performed, but one can likely perform several orders of magnitude more experiments *in silico* depending on the model used without running into the same financial burden of *in vivo* experiments. The perfect control one has over experiments *in silico* is also a major advantage, and one does not have the ethical concerns associated with *in vivo* experiments.

While there has been very limited use of mathematical cancer models in clinical use, there has been a sharp incline in the development of them on the research level over the past three decades—leading to a rich selection of models [26]. There are two prominent approaches of data-driven models in the form of machine learning and simulations of mechanistical models based on qualitative knowledge, as well as some that combine the two. Each has their merits, but machine learning requires a huge amount of data to have a chance at producing good prediction for the large variety of patients there are. This is data we do not have access to, and the rest of

this thesis will therefore focus on simulation of mechanistic models.

The many mechanistic cancer models that exist can be classified by the representation they use to describe cancer evolution in the body and they range from continuous models governed by partial differential equations to discrete models such as cellular automaton, petri net, cellular Potts models or agent-based modelling [26]. There are also hybrid models that combine them by using a discrete model for the individual cells and a continuous model for other parts of the model such as transport of drugs from blood vessels to cancer cells. The models can also be classified by the length scales that are included in the model—from how genes and subcellular processes affect a single cell to how cancer and vascular density changes throughout the tumor. Models that incorporate both small and large length scales are referred to as multiscale models.

## 1.4 A Breast Cancer Simulation Needing Improvements

In a recent paper [27] our research group presented a complex hybrid, multiscale model of breast cancer that can be initiated by the measurements and samples taken as standard in the standardized pathway for breast cancer patients in Norway. The model does in a nutshell consist of discrete cancer cells, initialised from an annotated biopsy, that proliferate or die according to a stochastic cellular automaton's response to the surrounding environment and sub-cellular processes modelled for each cell. The environment surrounding each cancer cell consists of a number of substances such as chemotherapeutic and anti-angiogenic agents diffusing continuously between discrete blood vessels and the cells. While the model is specific to breast cancer, it is fairly straight forward to adapt it to any solid cancer by replacing the pharmacokinetic and pharmacodynamic equations for the relevant chemotherapies of the relevant cancer type.

The paper demonstrates the potential of this approach by reproducing the outcome of four patients, while calibrating some parameters that could not be measured in the patients nor be found in the literature, and correctly predicting a fifth patient's response to their treatment using these parameters. While these results are encouraging, there are also significant limitations. A major one of these limitations is the small simulation size of  $200 \times 300 \mu\text{m}$  dictated by prohibitively large runtimes. At this size the simulated piece of the cancer can not by any means be expected to be representative for the whole tumor, as it is not uncommon for breast cancers to be two orders of magnitude larger than this in each direction and cancer is as we have discussed highly heterogeneous in nature [28, 29].

In chapter 2 our solution to remedy most of this limitation is presented in the form of a research paper. We were able to simulate a clinically relevant cancer size of a full biopsy, roughly 500 times larger than previously achieved. This was done by utilising a novel way of implementing the stochastic evolution of the discrete cancer cells and blood vessels, in a scalable manner, on top of known methods for solving the continuous parts of the model.



# Bibliography

- [1] H. Wang et al. “Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015”. In: *The Lancet* 388.10053 (2016), pp. 1459–1544.
- [2] C. Wild et al. *World Cancer Report: Cancer Research for Cancer Prevention*. Lyon, France: International Agency for Research on Cancer, 2020. Available at: <http://publications.iarc.fr/586>.
- [3] J. Ferlay et al. “Cancer statistics for the year 2020: An overview”. In: *International Journal of Cancer* 149.4 (2021), pp. 778–789.
- [4] H. Sung et al. “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries”. In: *CA: A Cancer Journal for Clinicians* 71.3 (2021), pp. 209–249.
- [5] National Cancer Institute. *NCI Professional Judgment Budget, President’s Budget and Appropriations*. May 2022. Available at: <https://www.cancer.gov/about-nci/budget/fact-book/historical-trends/bypass-appropriations> (visited on 09/25/2022).
- [6] Cancer Registry of Norway. *Cancer in Norway 2021 - Cancer incidence, mortality, survival and prevalence in Norway*. Oslo: Cancer Registry of Norway, 2022. Available at: [https://www.kreftregisteret.no/globalassets/cancer-in-norway/2021/cin\\_report.pdf](https://www.kreftregisteret.no/globalassets/cancer-in-norway/2021/cin_report.pdf).
- [7] H. G. Welch, L. M. Schwartz, and S. Woloshin. “Are Increasing 5-Year Survival Rates Evidence of Success Against Cancer?” In: *JAMA* 283.22 (June 2000), pp. 2975–2978. Available at: <https://doi.org/10.1001/jama.283.22.2975>.
- [8] F. R. Lichtenberg. “Are Increasing 5-Year Survival Rates Evidence of Success Against Cancer? A Reexamination Using Data from the U.S. and Australia”. In: *Forum for Health Economics & Policy* 13.2 (2010). Available at: <https://doi.org/10.2202/1558-9544.1202>.
- [9] A. Noone et al. *SEER Cancer Statistics Review, 1975–2015*. Bethesda, MD: National Cancer Institute, Apr. 2018. Available at: [https://seer.cancer.gov/csr/1975\\_2015/](https://seer.cancer.gov/csr/1975_2015/).

## BIBLIOGRAPHY

---

- [10] M. Quaresma, M. P. Coleman, and B. Rachet. “40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971–2011: a population-based study”. In: *The Lancet* 385.9974 (2015), pp. 1206–1218.
- [11] A. I. of Health and Welfare. “Cancer survival and prevalence in Australia: Period estimates from 1982 to 2010”. In: *Asia-Pacific Journal of Clinical Oncology* 9.1 (Mar. 2013), pp. 29–39.
- [12] C. L. Sawyers et al. “AACR cancer progress report 2013”. In: *Clinical Cancer Research* 19.20\_Supplement (2013), S1–S98.
- [13] N. H. G. R. Institute. *The cost of sequencing a human genome*. Nov. 2021. Available at: <https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost> (visited on 09/27/2022).
- [14] J. N. Weinstein et al. “The cancer genome atlas pan-cancer analysis project”. In: *Nature Genetics* 45.10 (2013), pp. 1113–1120.
- [15] Y. Nilssen et al. “Decreasing waiting time for treatment before and during implementation of cancer patient pathways in Norway”. In: *Cancer Epidemiology* 61 (2019), pp. 59–69. Available at: <https://www.sciencedirect.com/science/article/pii/S1877782119300384>.
- [16] M. Silva. *Cancer patient pathways in Norway*. Jan. 2021. Available at: <https://www.kreftregisteret.no/en/Research/Projects/cancer-patient-pathways-in-norway/> (visited on 10/08/2022).
- [17] Helsedirektoratet. *Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av pasienter med brystkreft*. Mar. 2022. Available at: [https://www.helsedirektoratet.no/produkter/\\_/attachment/inlin e/54aa081b-9a76-409b-850f-abf657174663:ab33a788ac75510aca73ae3e a1c2805dee434810/Nasjonalt%20handlingsprogram%20for%20pasienter%20med%20brystkreft.pdf](https://www.helsedirektoratet.no/produkter/_/attachment/inlin e/54aa081b-9a76-409b-850f-abf657174663:ab33a788ac75510aca73ae3e a1c2805dee434810/Nasjonalt%20handlingsprogram%20for%20pasienter%20med%20brystkreft.pdf) (visited on 10/07/2022).
- [18] C. Allemani et al. “Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries”. In: *The Lancet* 391.10125 (2018), pp. 1023–1075.
- [19] R. L. Siegel et al. “Cancer statistics, 2022”. In: *CA: A Cancer Journal for Clinicians* 72.1 (2022), pp. 7–33.
- [20] A. I. of Health and Welfare. *Cancer data in Australia, Cancer survival data visualisation*. Oct. 2022. Available at: <https://www.aihw.gov.au/reports/cancer/cancer-data-in-australia/contents/cancer-survival-data-visualisation> (visited on 10/10/2022).

- [21] C. Bode. “The nasty surprise of a complex drug–drug interaction”. In: *Drug Discovery Today* 15.9 (2010), pp. 391–395. Available at: <https://www.sciencedirect.com/science/article/pii/S1359644610000760>.
- [22] B. Percha and R. B. Altman. “Informatics confronts drug–drug interactions”. In: *Trends in Pharmacological Sciences* 34.3 (2013), pp. 178–184. Available at: <https://www.sciencedirect.com/science/article/pii/S0165614713000072>.
- [23] C. Pauli et al. “Personalized In Vitro and In Vivo Cancer Models to Guide Precision Medicine”. In: *Cancer Discovery* 7.5 (2017), pp. 462–477.
- [24] T. Haiden et al. *Evaluation of ECMWF forecasts, including the 2021 upgrade*. eng. Sept. 2021. Available at: <https://www.ecmwf.int/node/20142>.
- [25] G. B. Barcellos et al. “Molecular Modeling as a Tool for Drug Discovery”. In: *Current Drug Targets* 9.12 (2008), pp. 1084–1091. Available at: <https://www.ingentaconnect.com/content/ben/cdt/2008/00000009/00000012/art00009>.
- [26] S. Bekisz and L. Geris. “Cancer modeling: From mechanistic to data-driven approaches, and from fundamental insights to clinical applications”. In: *Journal of Computational Science* 46 (2020), p. 101198.
- [27] X. Lai et al. “Toward personalized computer simulation of breast cancer treatment: A multiscale pharmacokinetic and pharmacodynamic model informed by multitype patient data”. In: *Cancer Research* 79.16 (2019), pp. 4293–4304.
- [28] S. Narod et al. “Are two-centimeter breast cancers large or small?” In: *Current Oncology* 20.4 (2013), pp. 205–211.
- [29] C. L. Carter, C. Allen, and D. E. Henson. “Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases”. In: *Cancer* 63.1 (1989), pp. 181–187.

## BIBLIOGRAPHY

---



## 2 | Paper

### Sammendraget fra artikkelen

Matematisk modellering og simulering er en lovende tilnærming til persontilpasset kreftmedisin. Å sette sammen diskrete celle-baserte modeller med kontinuerlige modeller til en hybrid, cellulær automata (CA) er en kraftfull framgangsmåte for å etterlikne biologisk kompleksitet og beskrive den dynamiske utvekslingen av informasjon over forskjellige størrelsesordner. Kreftens kompleksitet, heterogenitet og multiskala-natur utgjør derimot, en betydelig beregningsmessig utfordring. Når man ser på klinisk relevante utsnitt av svulsten blir det beregningsmessig svært kostbart. Selvom effektive parallelliseringsteknikker finnes for kontinuerlige modeller, så krever deres sammenkobling med diskrete modeller, CA spesielt, avanserte og forseggjorte løsninger. Med utgangspunkt i FEniCS, en populær og kraftig platform for å løse partielle differensiallikninger med vitenskapelige beregninger, utviklet vi en parallell algoritme som kobler sammen stokastisk CA med differensiallikninger (<https://bitbucket.org/HTasken/cansim>). Algoritmen minimerer kommunikasjonen mellom prosessene som deler nabovertidene i CA, samtidig som den også tillater reproduserbarhet under stokastiske oppdateringer. Vi demonstrerte potensialet for vår løsning på en kompleks hybrid cellulær automaton modell for brystkreft behandlet med cellegiftskombinasjoner. På en enkeltkjerne prosessor, oppnådde vi nær linjær skalering med økende problemstørrelse, mens svak parallell skalering viste moderat vekst i løsningstiden relativt til størrelsesøkningen av problemet. Til slutt brukte vi algoritmen på et problem som er 500 ganger større enn tidligere arbeid, som tillater oss å kjøre simuleringer av persontilpasset terapi basert på heterogen celledetthet og perfusjon av svulsten estimert fra magnetresonanstomografi data på en enestående skala.

**Paper: A scalable solver for a stochastic, hybrid cellular automaton model of personalized breast cancer therapy**

# A scalable solver for a stochastic, hybrid cellular automaton model of personalized breast cancer therapy

Xiaoran Lai<sup>1</sup>  | Håkon A. Taskén<sup>1</sup> | Torgeir Mo<sup>2</sup> | Simon W. Funke<sup>3</sup>  | Arnoldo Frigessi<sup>1,4</sup> | Marie E. Rognes<sup>3</sup> | Alvaro Köhn-Luque<sup>1</sup>

<sup>1</sup>Oslo Centre for Biostatistics and Epidemiology, Faculty of Medicine, University of Oslo, Oslo, Norway

<sup>2</sup>Institute for Cancer Research, Oslo University Hospital, Oslo, Norway

<sup>3</sup>Simula Research Laboratory, Lysaker, Norway

<sup>4</sup>Oslo Centre for Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway

## Correspondence

Alvaro Köhn-Luque, Oslo Centre for Biostatistics and Epidemiology, Faculty of Medicine, University of Oslo, Norway.  
Email: alvaro.kohn-luque@medisin.uio.no

## Funding information

European Research Council; Horizon 2020 Framework Programme; Norges Forskningsråd; Universitetet i Oslo

## Abstract

Mathematical modeling and simulation is a promising approach to personalized cancer medicine. Yet, the complexity, heterogeneity and multi-scale nature of cancer pose significant computational challenges. Coupling discrete cell-based models with continuous models using hybrid cellular automata (CA) is a powerful approach for mimicking biological complexity and describing the dynamical exchange of information across different scales. However, when clinically relevant cancer portions are taken into account, such models become computationally very expensive. While efficient parallelization techniques for continuous models exist, their coupling with discrete models, particularly CA, necessitates more elaborate solutions. Building upon FEniCS, a popular and powerful scientific computing platform for solving partial differential equations, we developed parallel algorithms to link stochastic CA with differential equations (<https://bitbucket.org/HTasken/cansim>). The algorithms minimize the communication between processes that share CA neighborhood values while also allowing for reproducibility during stochastic updates. We demonstrated the potential of our solution on a complex hybrid cellular automaton model of breast cancer treated with combination chemotherapy. On a single-core processor, we obtained nearly linear scaling with an increasing problem size, whereas weak parallel scaling showed moderate growth in solving time relative to increase in problem size. Finally, we applied the algorithm to a problem that is 500 times larger than previous work, allowing us to run personalized therapy simulations based on heterogeneous cell density and tumor perfusion conditions estimated from magnetic resonance imaging data on an unprecedented scale.

## KEYWORDS

breast cancer, cancer modeling, domain decomposition, FEniCS, multi-scale modeling, parallel computing, personalized cancer therapy

Xiaoran Lai and Håkon A. Taskén contributed equally to this paper.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *International Journal for Numerical Methods in Biomedical Engineering* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Mathematical modeling and computer simulations, informed by patient-specific clinical data, can be used to make personalized predictions of response to cancer therapy.<sup>1</sup> The methodology can in principle be used to design patient-specific treatment plans and constitutes a promising approach to personalized cancer medicine. One of the main goals is the development of quantitative and computational tools that can effectively and efficiently simulate the outcome of multiple therapeutical strategies in individual patients. However, the complexity, heterogeneity and multi-scale nature of cancer present severe computational challenges to that goal. Specifically, treated cancer tissue entails multiple interacting processes occurring at different spatio-temporal scales (for instance, intracellular signaling pathways, drug pharmacokinetics or single cell fates such as division or death). One approach to model the interactions is the hybrid cellular automata (HCA) framework coupling discrete cellular automata (CA) and continuous model components accounting for the different phenomena and scales.<sup>2,3,4,5</sup>

To simulate HCA models, computational algorithms have to balance the numerical schemes used to efficiently solve the different discrete and continuous formalism together with the sharing of partial states between them.<sup>6,7</sup> Computational time can be easily misspent in situations in which solvers are on hold while others need to reach stationary states. Furthermore, because different solvers can have different numerical meshes in space and time, resolving states sharing is not always a straightforward task. The large number of cells and the heterogeneity present within the tissue add to the complexity of simulating cancer tissue. Tumors, for example, frequently have areas with different cell densities and perfusion characteristics, cells and vessels may have different characteristics, and so on. As such, heterogeneity is known to influence treatment outcomes, large scale simulations that can capture it are needed. Furthermore, a simulation algorithm is only useful in clinical settings if the computational time is compatible with clinical decision-making. As a result, parallel computing is expected to play an important role.

FEniCS is a finite element computing platform for solving partial differential equations (PDEs) that has been fundamentally designed for parallel processing.<sup>8</sup> While FEniCS has already been extended to couple PDEs with ordinary differential equations (ODEs) associated with mesh nodes, the coupling with stochastic CA models required to solve HCA, is not straightforward. Specifically, mesh partitioning methods used by FEniCS are unsuitable for applying CA rules based on values in different CA meshes. We solved this issue by developing an algorithm that creates a map of what information needs to be exchanged between processes during each CA update. In addition, we implement a priority system to resolve potential conflicts caused by multiple processes attempting to set a CA node value at the same time. This, combined with an efficient and lightweight random number generation implementation, enables us to reproduce stochastic model simulations, produce unbiased results, and claim statistical significance.

To demonstrate the potential of our approach and algorithms, we consider an updated multi-scale HCA model, previously used for simulating personalized breast cancer therapy.<sup>9,10</sup> In previous work,<sup>9</sup> model simulations were used to explain treatment outcomes at the level of individuals and suggest more successful regimes for tumors that did not respond to therapy. These findings were encouraging, but they were limited to small sections of the tumor with only a few hundred biological cells and were simulated on a single CPU. While the new model is parameterized using the same methods as described,<sup>9</sup> this work is dedicated to describing the unique numerical schemes implementing the stochastic hybrid discrete-continuum modeling. The new parallel solver allows to simulate more clinically relevant pieces of tumors and to capture the tumor heterogeneity as observed in magnetic resonance imaging (MRI) data. In addition to testing the solver scalability of single core performance, we performed a weak scaling study on a cluster of up to 80 cores as problem size increased. Finally, we demonstrate that it is now possible to run simulations of 2D tumor sections 500 times larger than previously.<sup>9,10</sup>

## 2 | METHODS

### 2.1 | Magnetic resonance imaging

Variation in cellular and vascular density across breast tumor tissue of one patient was assessed with MRI. The patient underwent MRI examinations before the start of treatment, and after 1 and 12 weeks of neoadjuvant treatment. Examinations were performed on an ESPREE 1.5 T MR scanner (Siemens, Erlangen, Germany) equipped with a phased-array bilateral breast coil (CP breast coil, Siemens, Erlangen, Germany). The MRI protocol was a state-of-the-art MRI protocol<sup>11</sup> with T2-weighted, diffusion-weighted (DW), and dynamic contrast-enhanced (DCE) MRI. DCE-images were acquired using a radial, spoiled gradient echo with k-space weighted image contrast (KWIC), using spectral adiabatic inversion



recovery (SPAIR) for fat-suppression (TE = 2.59 ms, TR = 5.46 ms, flip angle = 15°, field of view = 320 mm × 320 mm, in-plane resolution = 1 mm × 1 mm, slice thickness = 15 mm). After eight pre-contrast series, the contrast agent (Gadovist, Bayer Pharma, Germany) was administered at a dosage of 0.8 mmol kg<sup>-1</sup>, at a rate of 3 ml s<sup>-1</sup>, followed by a 20 ml saline flush. Subsequently, 32 post contrast image-series were acquired at a frequency of 1/13 s<sup>-1</sup>.

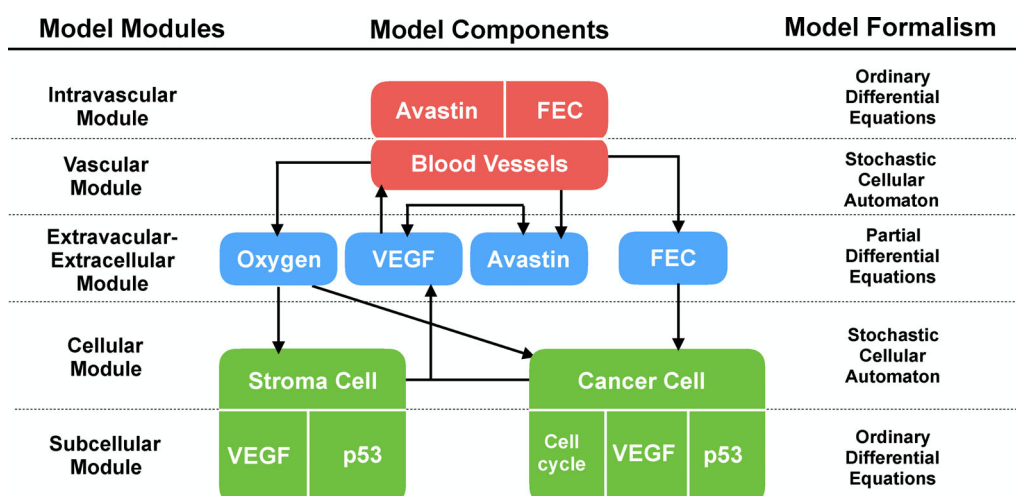
DCE images were analyzed using an extended Tofts two-compartment pharmacokinetic model, yielding 1 mm × 1 mm-resolution maps of in-vivo perfusion parameters in the tumor. Perfusion parameters calculated includes the permeability-surface area product,  $K_{trans}$  of the vasculature, volume fraction of the vascular space,  $v_p$  and volume fraction of the extravascular, extracellular space,  $v_e$ . Additionally, the cellular density was estimated as  $v_c = 1 - (v_p + v_e)$ . Image analysis was performed using nICE (Nordic NeuroLab, Bergen, Norway). Patient specific arterial input function was sampled from a region of the right atrium of the heart, visible in the DCE images.

## 2.2 | Multi-scale mathematical model

We begin by summarizing the multi-scale mathematical model of breast cancer and its treatment by a combination of administrated drugs. A more in-depth biological and clinical discussion of the model can be found in our previous work.<sup>9</sup> Using hybrid CA,<sup>2,3</sup> the model accounts for the response of a 2D cross section of tumor tissue to a combination of chemotherapeutic and anti-angiogenic agents. Thus, a tumor section is represented by a finite regular square lattice  $L$ , consisting of a set of nodes labeled by their positions  $\mathbf{x} \in L$ ,  $\mathbf{x} = (i\Delta x, j\Delta y)$ ,  $i = 0, \dots, n-1, j = 0, \dots, m-1$ ,  $\Delta x$  being the distance between nearest nodes. Biological cells and cross-sectional cuts of functional blood vessels are modeled as individual agents occupying a single lattice node. Microenvironmental factors in the tissue section, such as oxygen, are modeled as continuous variables over the domain  $D = [0, (n-1)\Delta x] \times [0, (m-1)\Delta y] \subset \mathbb{R}^2$ . Intracellular and intravascular processes are modeled as continuous variables associated to each cell and blood vessel, respectively. To account for cell and blood vessel dynamics and the molecular factors that control them, we build five interlinked model modules: the cellular, subcellular, vascular, intravascular and extravascular-extracellular modules. Figure 1 shows a diagram with the main components of each module, the interactions between them and the model formalism used in each case. A more detailed description of each module is provided below.

### 2.2.1 | Stochastic cellular automaton for the cellular module

The presence of cancer and other cells, referred here as stroma, on the lattice sites at time  $t$ , is described by a function  $\mathcal{L}$  that takes three possible values:  $\mathcal{L}(\mathbf{x}, t) = 0$ , if the site  $\mathbf{x}$  is empty;  $\mathcal{L}(\mathbf{x}, t) = 1$ , if  $\mathbf{x}$  is occupied by a cancer cell; or



**FIGURE 1** Modular structure of the hybrid cellular automaton model for vascular tumor growth and combination therapy. The directions of arrows indicate influence between different components of each module. The last column shows the mathematical formalism used to describe each of the modules

$\mathcal{L}(\mathbf{x}, t) = 2$ , if  $\mathbf{x}$  is occupied by a stroma cell. At most one cell can occupy one site of the grid. It is assumed that cancer cells can divide or die but cell movements are neglected. Cancer cell division is controlled by an internal cell cycle clock described in the subcellular module below. When a cell at position  $\mathbf{x}_{\text{cell}}$  is committed to divide, it can be killed by the chemotherapy agent  $i = 1, 2, 3$  with a probability given by:

$$I\left(\frac{G_i(\mathbf{x}_{\text{cell}}, t)}{\max(G_{i,0})}; 1, b_i\right), \quad (1)$$

where  $I(x; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x y^{a-1} (1-y)^{b-1} dy$ . The value  $\max(G_{i,0})$  refers to the maximal concentration of chemotherapy  $i$  in the blood and  $b_i$  is a parameter describing the sensitivity toward chemotherapy  $i$ . If the cell at position  $\mathbf{x}$  dies at time  $t$ , it is removed from the computational grid by setting  $L(\mathbf{x}, t) = 0$ . If chemotherapy does not kill the proliferating cell, a daughter cell is placed at an empty (Moore) neighbor location with the highest oxygen pressure. If, however, no free space is available in the neighborhood, the cell cycle of the parent cell is reset to zero and no daughter cell is produced.

Stroma cells in our cellular automaton do not proliferate or die. Thus, their role in the model is restricted to competition for space and oxygen with cancer cells and to production of the vascular endothelial growth factor (VEGF) under hypoxia (low oxygen tension).

### 2.2.2 | Stochastic cellular automaton for the vascular module

We consider only perfused functional blood vessels and assume that all are cylindrical and perpendicular to the modeled tumor section. The presence of cross-sectional vessel cuts in the lattice at time  $t$  is given by the function  $\mathcal{G}$ , with  $\mathcal{G}(\mathbf{x}, t) = 1$  representing the presence of a functional vessel and  $\mathcal{G}(\mathbf{x}, t) = 0$  its absence. Vessel dynamics is modeled by a birth-death process, with the probability of creating and removing vessels depending on the extracellular spatial distribution of the angiogenic factor VEGF. Specifically, we assume there is a range of VEGF concentrations where vessels are viable but outside that concentration range vessels are more likely to regress and disappear. Therefore, at every fixed time step  $\Delta v$ , we define the probability of creating and removing a vessel at  $\mathbf{x} \in L$  as:

$$\begin{aligned} P(\mathcal{G}(\mathbf{x}, t + \Delta v) = 0 | \mathcal{G}(\mathbf{x}, t) = 1) &= p_{\text{death}} \text{ if } V_{\mathbf{x}} < V_{\text{low}} \text{ or } V_{\mathbf{x}} > V_{\text{high}} \\ P(\mathcal{G}(\mathbf{x}, t + \Delta v) = 1 | \mathcal{G}(\mathbf{x}, t) = 0) &= p_{\text{birth}}(\mathbf{x}, t) \text{ if } V_{\text{low}} \leq V_{\mathbf{x}} \leq V_{\text{high}} \end{aligned} \quad (2)$$

where  $V_{\text{low}}$  and  $V_{\text{high}}$  are lower and upper thresholds of VEGF concentration where functional vessels are viable, and  $V_{\mathbf{x}} = V(\mathbf{x}, t)$  denotes the extracellular concentration of VEGF at location  $\mathbf{x}$  at time  $t$ , which is described below. Extending our original model,<sup>9</sup> we now assume that the probability of birth  $p_{\text{birth}}(\mathbf{x}, t)$  is proportional to the VEGF concentration  $V(\mathbf{x}, t)$ , instead of being constant. Following Owen et al.,<sup>12</sup> we also assume that at each time-step  $\Delta v$ , the expected value of new created vessels is given by  $n_{\text{new}}$ :

$$E[n_{\text{new}}] = Pr_{\text{sprout}} 2\pi r_0 h_c \Delta v \int \frac{r_{\mathbf{x}}}{r_0} \frac{V_{\mathbf{x}}}{V_{\text{sprout}} + V_{\mathbf{x}}} \sum_{\mathbf{x} \in \mathbf{x}_v} \delta(\mathbf{x} - \mathbf{x}, t) d\mathbf{x}, \quad (3)$$

where  $Pr_{\text{sprout}}$  is the maximum probability that an endothelial sprout emerges from a surface of a vessel and forms a new vessel,  $X_v = \{\mathbf{x} : \mathcal{G}(\mathbf{x}, t) = 1\}$  is the set of all blood vessel locations at time  $t$ .  $r_{\mathbf{x}} = r(\mathbf{x}, t)$  is the vessel radius,  $r_0$  is the average radius of initial vessels,  $h_c$  is the height of the simulated tissue layer and  $V_{\text{sprout}}$  is the VEGF concentration at which the probability is half-maximal.

### 2.2.3 | Systems of ordinary differential equations for the intravascular module

The time-dependent concentrations of four drugs in the blood, that is, Avastin and a cocktail of three chemotherapies (Fluorouracil, Epirubicin, and Cyclophosphamide) together known as FEC100, are modeled by their respective

pharmacokinetic equations, dosage and drug administration schedule. It is assumed that all vessels share the same drug concentration at a given time point. For Avastin, a two-compartment model is used:

$$\frac{dA_1}{dt} = -\frac{q}{v_1}A_1 - \frac{cl}{v_1}A_1 + \frac{q}{v_2}A_2, \quad (4a)$$

$$\frac{dA_2}{dt} = -\frac{q}{v_2}A_2 - \frac{cl}{v_1}A_1, \quad (4b)$$

where  $A_1 = A_1(t)$  and  $A_2 = A_2(t)$  are the concentrations of Avastin at time  $t$  in the plasma and peripheral compartments, respectively, and  $q$ ,  $cl$ ,  $v_1$ ,  $v_2$ , are kinetic and compartment volume parameters. For Fluorouracil a single compartment model is used:

$$\frac{dG_{1,0}}{dt} = -\frac{v_{max}}{k_m + G_{1,0}}G_{1,0}, \quad t \neq t_{D_n} \quad (5)$$

and

$$G_{1,0}(t_{d_n}) = \frac{d_n}{V}, \quad n = 1 \dots N \quad (6)$$

where  $G_{1,0} = G_{1,0}(t)$  is the plasma concentration at time  $t$ ,  $d_n$  is the  $n$ th dose, and  $v_{max}$ ,  $k_m$  and  $V$  are kinetic parameters. For Epirubicin, a three-compartment model is used:

$$\frac{dG_{2,0}}{dt} = -\frac{q_2}{w_1}G_{2,1} - \frac{q_3}{w_1}G_{2,2} - \frac{cl_2}{w_1}G_{2,0} + \frac{q_3}{w_3}G_{2,2} + \frac{q_2}{w_2}G_{2,1}, \quad (7a)$$

$$\frac{dG_{2,1}}{dt} = -\frac{q_2}{w_2}G_{2,1} + \frac{q_2}{w_1}G_{2,0}, \quad (7b)$$

$$\frac{dG_{2,2}}{dt} = -\frac{q_3}{w_3}G_{2,2} + \frac{q_3}{w_1}G_{2,0}, \text{ and } G_{2,0}(t_{d_n}) = \frac{d_n}{V_1}, \quad n = 1 \dots N \quad (7c)$$

and

$$G_{2,0}(t_{d_n}) = \frac{d_n}{V_1}, \quad n = 1 \dots N \quad (7d)$$

where  $G_{2,0} = G_{2,0}(t)$  is the plasma concentration at time  $t$  and  $G_{2,1} = G_{2,1}(t)$  and  $G_{2,2} = G_{2,2}(t)$  are concentrations in two peripheral compartments at time  $t$ . The parameters  $q_1$ ,  $q_2$ ,  $q_3$ ,  $cl_2$  are kinetic rates and  $w_1$ ,  $w_2$ ,  $w_3$  are compartment volumes. For Cyclophosphamide, a single compartment model is again used:

$$\frac{dG_{3,0}}{dt} = -\frac{cl_3}{u}G_{3,0}, \quad (8)$$

and

$$G_{3,0}(t_{d_n}) = \frac{d_n}{V}, \quad n = 1 \dots N \quad (9)$$

where  $G_{3,0} = G_{3,0}(t)$  is the plasma concentration at time  $t$  and  $cl_3$ ,  $u$  are kinetic parameters. For  $G_{2,0}$  and  $G_{3,0}$ , closed-form solutions, described in elsewhere,<sup>13</sup> were used for the calculation of plasma chemotherapy concentration.

### 2.2.4 | Ordinary differential equations for the intracellular module

For each cell  $c$  at position  $\mathbf{x}_c$ , the intracellular amount of p53, denoted by  $P_c = P_c(t)$ , and the intracellular amount of VEGF, denoted by  $V_c = V_c(t)$ , are modeled by the following system of equations:

$$\frac{dP_c}{dt} = k_1 - k'_1 \frac{K_{x_c}}{K_{p53} + K_{x_c}} P_c, \quad (10)$$

$$\frac{dV_c}{dt} = k_2 + k''_2 \frac{P_c V_c}{J_5 + V_c} - k'_2 \frac{K_{x_c}}{K_V + K_{x_c}} V_c \quad (11)$$

where  $k_1, k'_1, K_{p53}, k_2, k'_2, k''_2, J_5$  and  $K_V$  are kinetic parameters. The function  $K_{x_c} = K(\mathbf{x}_c, t)$  represents the oxygen concentration at cell location  $x_c$  and is modeled by a PDE (described below).

Similarly, the progression through the cell cycle  $\phi = \phi(t)$  of each cell  $c$  at position  $\mathbf{x}_c$  is modeled by the equation:

$$\frac{d\phi}{dt} = \frac{K_{x_c}}{T(K_\phi + K_{x_c})}, \quad (12)$$

where  $T$  and  $K_\phi$  are scalar parameters. At start, all cancer cells are assigned a random  $\phi$  value between 0 and 1. The value of the cell cycle of a newly divided cell is set to  $\phi = 0$ , and it increases until the first cellular update after  $\phi \geq 1$ . At that point, the cancer cell is committed to divide and will either divide or die due to the effect of chemotherapy as described above.

### 2.2.5 | Reaction–diffusion equations for the extravascular-extracellular module

Each of the four drugs (Avastin  $A$ , Fluorouracil  $G_{1,0}$ , Epirubicin  $G_{2,0}$ , and Cyclophosphamide  $G_{3,0}$ ) is modeled as moving from the vessels (concentrations denoted by  $A_1, G_{1,0}, G_{2,0}, G_{3,0}$  respectively) to the extravascular-extravascular space (concentrations denoted by  $A, G_1, G_2, G_3$ , respectively). These rates depend on the difference in concentrations between the vessel  $U_v = U_v(t)$  and the extravascular-extravascular space (EES),  $U_x = U(\mathbf{x}, t)$ , and the drugs' vessel surface permeabilities  $P_U$ , and are modeled as:

$$R_U = \int_{V_0} (U_v - U_x) P_U S_x d\mathbf{x} \quad (13)$$

where the integration domain  $V_0$  is the total volume of the cross-section, and  $S_x = S(\mathbf{x})$  is the vessel surface-per-volume ratio for a vessel at  $\mathbf{x}$ .

$$\begin{aligned} R_U &= \int_{V_0} (U_v - U_x) P_U \frac{2\pi r_x h}{\Delta x^3} \sum_{\mathbf{x}_v \in \mathbf{X}_v} \delta(\mathbf{x} - \mathbf{x}_v, t) d\mathbf{x} \\ &= \frac{P_U 2\pi r_0 h}{\Delta x^3} \int_{V_0} (U_v - U_x) \frac{r_x}{r_0} \sum_{\mathbf{x}_v \in \mathbf{X}_v} \delta(\mathbf{x} - \mathbf{x}_v, t) d\mathbf{x}, \end{aligned} \quad (14)$$

where  $r_0$  is the average radius of the initial vessels,  $r_x = r(\mathbf{x}, t)$  is the radius of the vessel at  $\mathbf{x}$ , and  $h$  is the vessel length (which is the height of tissue section and equal to the cell size  $\Delta x$ ). Function  $\delta$  is the Dirac delta function and for convenience, we introduce a short-hand for the sum of Dirac delta functions over a set of points  $X_a$ :

$$\mathcal{N}_a = \mathcal{N}_a(\mathbf{x}, t) = \sum_{\mathbf{y} \in X_a} \delta(\mathbf{x} - \mathbf{y}, t) \quad (15)$$

This forms part of the reaction terms in Equations (17) to (21) below.

For chemotherapy and Avastin, we assumed that  $P_U = c_U P_{\text{Gado}}$ .  $P_{\text{Gado}}$  is the permeability of gadolinium, the contrast agent used in MRI, and  $c_U \in \mathbb{R}$  is a scaling constant. Given that the initial transfer rate  $K_{\text{trans}}^{\text{init}} = P_{\text{Gado}} S^{\text{init}} = P_{\text{Gado}} N_v^{\text{init}} 2\pi r_0 h / (S_{\text{MRI}} h)$ , is the Tofts model permeability-surface area product of gadolinium,<sup>14</sup> where  $N_v^{\text{init}}$  is the initial number of vessels and  $S_{\text{MRI}}$  is the permeability-surface area product of a voxel in MRI, we can derive that  $P_U = c_U K_{\text{trans}}^{\text{init}} S_{\text{MRI}} / (N_v^{\text{init}} 2\pi r_0)$ . It can then be substituted in Equation (14) for Avastin and the chemotherapies to give:

$$R_U = c_U K_{\text{trans}}^{\text{init}} \frac{S_{\text{MRI}}}{\Delta x^2 N_v^{\text{init}}} \int_{V_0} \frac{r_x}{r_0} (U_v - U_x) \mathcal{N}_v d\mathbf{x}, \quad (16)$$

### Oxygen

Oxygen pressure  $K = K(\mathbf{x}, t)$  in the EES is modeled by the following reaction–diffusion equation for all  $\mathbf{x} \in D$  and  $t > 0$ :

$$s_K \frac{\partial K}{\partial t} = D_K \nabla^2 K - \frac{\phi_K K}{K_1 + K} \mathcal{N}_c + \frac{P_K 2\pi r_0}{\Delta x^2} (K_0 - K) \mathcal{N}_v, \quad (17)$$

where  $D_K$ ,  $\phi_K$ ,  $K_1$  and  $K_0$  are scalar parameters,  $X_c = \{\mathbf{x} : \mathcal{L}(\mathbf{x}, t) \neq 0\}$  is the set of all cell locations.  $s_K \in \{0, 1\}$  is constant. When  $s_K = 1$ , this becomes to the time-dependent reaction–diffusion PDE, while  $s_K = 0$  corresponds to the steady-state equation. The latter is used to initialize the oxygen pressure given initial locations of cells and vessels. The first term in the right-hand-side of Equation (17) accounts for oxygen diffusion, the second term accounts for the consumption by cells and the third term represents point sources accounting for the flow of oxygen from the blood vessels.

### Chemotherapies

The concentration of each chemotherapy  $G_i = G_i(\mathbf{x}, t)$ , with  $i = 1, 2, 3$  in the EES is modeled by the following reaction–diffusion equation for  $\mathbf{x} \in D$  and  $t > 0$ :

$$s_C \frac{\partial G_i}{\partial t} = D_{G_i} \nabla^2 G_i - \psi_{G_i} G_i + c_G K_{\text{trans}}^{\text{init}} \frac{S_{\text{MRI}}}{\Delta x^2 N_v^{\text{init}}} \frac{r_x}{r_0} (G_{i,0} - G_i) \mathcal{N}_v, \quad (18)$$

where  $D_{G_i}$  is the diffusion constant of  $G_i$ ,  $\psi_{G_i}$  is the linear decay rate of  $G_i$ .  $G_{i,0}$  is the concentration of chemotherapy  $i$  in the blood. The first term in the right-hand-side of Equation (18) accounts for diffusion of the concentration, the second term accounts for drug decay, the third term represent point sources accounting for the flow chemotherapy from the vessels.

### VEGF-Avastin complex

Avastin is a VEGF-inhibitor that binds to the VEGF and produces an inactive VEGF-Avastin complex, thereby reducing the availability of active VEGF. The interactions between the Avastin, VEGF and VEGF-Avastin complex concentrations  $A = A(\mathbf{x}, t)$ ,  $V = V(\mathbf{x}, t)$  and  $C = C(\mathbf{x}, t)$  are given by the following system of reaction–diffusion equations for  $\mathbf{x} \in D, t > 0$ :

$$\frac{\partial V}{\partial t} = D_V \nabla^2 V + a V_c \mathcal{N}_c - k_a A V + k_d C - \psi_V V \quad (19)$$

$$\frac{\partial A}{\partial t} = D_A \nabla^2 A - k_a A V + k_d C - \psi_A A + c_A K_{\text{trans}}^{\text{init}} \frac{S_{\text{MRI}}}{\Delta x^2 N_v^{\text{init}}} \frac{r_x}{r_0} (A_1 - A) \mathcal{N}_v \quad (20)$$

$$\frac{\partial C}{\partial t} = D_C \nabla^2 C + k_a A V - k_d C - \psi_C C, \quad (21)$$

where  $D_V, D_A, D_C, a, b, k_a, k_d, \psi_V, \psi_A, \psi_C$  and  $c_A$  are scalar parameters and  $V_c$  is the time-dependent intracellular concentration of VEGF described in Equation (11) and  $A_1$  is the time-dependent concentration of Avastin intravascularly. The first term in the right-hand-side of Equation (19) accounts for diffusion of the VEGF concentration, the second term accounts for the release of VEGF by cells, the third and fourth terms account for VEGF binding/unbinding to the VEGF inhibitor Avastin and the last term accounts for natural VEGF decay. In Equation (20), the first term in the right-hand-side accounts for diffusion of the Avastin concentration, the second and third terms account for Avastin binding/unbinding to VEGF, the fourth term accounts for drug decay and the last term represent point sources accounting for the flow of Avastin from the vessels. Finally, the first term in the right-hand-side of Equation (21) accounts for diffusion of the VEGF-Avastin complex, the second and third terms account for Avastin binding/unbinding to VEGF and the last term represents natural decay of the complex.

### 2.3 | Patient-specific model initialization and parameterization

All simulations run in this study were personalized by data from a specific breast cancer patient (Patient 3) from a clinical trial.<sup>9</sup> Model initialization and parameterization are, unless differently specified, as in that publication, where clinical, histological, MRI and molecular data were used to estimate model parameters and initial values. Patient 3 is a complex patient and was chosen for exhibiting very heterogeneous perfusion conditions as observed by MRI. More precisely, MRI showed a tumor core with very low perfusion and viable cells and a tumor edge with much higher perfusion values. All parameter used for the simulations are listed in Table 1.

### 2.4 | Numerical methods

The main algorithm for the numerical solution of the full model equations is presented in Algorithm 1 below. Further specification of the numerical techniques for solving the separate subproblems are given in the text below. The numerical solver was implemented using the open source FEniCS Project finite element library.<sup>8</sup> The code is available at <https://bitbucket.org/HTasken/cansim>.

#### 2.4.1 | Numerical solution of the PDE systems

The time-dependent, nonlinear systems of PDEs describing the evolution of the oxygen and chemotherapy concentrations and the VEGF-Avastin complex (Equations (17) to (21)) are solved using the finite element method in space and finite difference method in time. The non-linear problems are solved using the Newton–Raphson method and all linear systems are solved using iterative Krylov methods designed to scale to large-scale simulations.

The computational domain  $D$ , representing a tissue slice, is a rectangular region:  $D = [0, (n-1)\Delta x] \times [0, (m-1)\Delta y] \subset \mathbb{R}^2$ . We call  $\mathcal{T}$  the base regular square mesh with  $n \times m$  vertices associated with  $D$ . The vertices are the potential locations for the vessels and biological cells. To avoid having the accuracy of the PDE solutions limited by the size of the biological cells, a finer mesh is used for the finite element discretization of the PDEs. Specifically, we used a uniform refinement  $\mathcal{T}_H$  of the base mesh with  $2 \cdot 2^2(n-1) \times 2^2(m-1)$  triangular elements. The same refined mesh was used for the discrete functions, that is, the delta function representation of the vessels and biological cells. See section 2.5 for more detail. The choice for the number of refinement was determined by a mesh independence study, in which the relative error of the solution was stable after twice refinement.

#### *Solving the chemotherapy concentration equations*

We first consider the numerical solution of the system of chemotherapy equations, 18 for  $i = 1, 2, 3$ . Each equation is time-dependent but linear, and the equations are independent of each other. We first discretize each PDE by the implicit second-order Crank–Nicolson scheme in time. At each time  $t^k$  for  $k = 1, \dots, K$ , given the concentrations at the previous time  $G_i^{k-1}$ , we solve for the concentrations  $G_i^k$  using the finite element method with continuous piecewise linear finite elements relative to the mesh  $\mathcal{T}_H$ . The resulting linear systems of equations are symmetric and positive

TABLE 1 Overview of all parameters

| Parameter    | Description  | Value                  | Units                                 |
|--------------|--|------------------------|---------------------------------------|
| $k_\phi$     | Oxygen concentration at half-maximal cycle speed       | 1.4                    | mmHg                                  |
| $k'_1$       | Degradation rate of p53 by oxygen                      | 0.01                   | min <sup>-1</sup>                     |
| $K_{TP53}$   | Oxygen concentration for half-maximal TP53 degradation | 0.01                   | mmHg                                  |
| $k_2$        | Synthesis rate of VEGF                                 | 0.002                  | min <sup>-1</sup>                     |
| $k'_2$       | Reaction rate of p53 with VEGF                         | 0.01                   | min <sup>-1</sup>                     |
| $J_5$        | sVEGF concentration for half-maximal sVEGF production  | 0.04                   | µg ml <sup>-1</sup>                   |
| $K_{VEGF}$   | Oxygen concentration for half-maximal VEGF degradation | 0.01                   | mmHg                                  |
| $D_k$        | Oxygen diffusion coefficient                           | $1.05 \times 10^5$     | µm <sup>2</sup> min <sup>-1</sup>     |
| $r_k$        | Oxygen supply rate                                     | $1.88 \times 10^4$     | min <sup>-1</sup>                     |
| $K_0$        | Oxygen concentration in the blood                      | 20                     | mmHg                                  |
| $\phi_k$     | Oxygen consumption rate                                | 900                    | min <sup>-1</sup>                     |
| $K_1$        | Oxygen concentration for half-maximal consumption      | 2.5                    | mmHg                                  |
| $D_v$        | VEGF diffusion coefficient                             | $3.52 \times 10^3$     | µm <sup>2</sup> min <sup>-1</sup>     |
| $a$          | VEGF secretion slope                                   | $6.66 \times 10^{-6}$  | min <sup>-1</sup>                     |
| $b$          | VEGF secretion intercept                               | $-1.10 \times 10^{-6}$ | µg ml <sup>-1</sup> min <sup>-1</sup> |
| $k_a$        | VEGF association rate to Avastin                       | $7.4 \times 10^{-1}$   | µg <sup>-1</sup> ml min <sup>-1</sup> |
| $k_d$        | VEGF dissociation rate from Avastin                    | $1.76 \times 10^{-3}$  | min <sup>-1</sup>                     |
| $\psi_v$     | VEGF decay rate  | $1.0 \times 10^{-2}$   | min <sup>-1</sup>                     |
| $D_A$        | Avastin diffusion coefficient                          | $2.4 \times 10^3$      | µm <sup>2</sup> min <sup>-1</sup>     |
| $\psi_C$     | Complex decay rate                                     | $1.0 \times 10^{-2}$   | min <sup>-1</sup>                     |
| $D_{G^i}$    | Chemotherapies diffusion coefficient                   | $9.6 \times 10^3$      | µm <sup>2</sup> min <sup>-1</sup>     |
| $\psi_{G^i}$ | Chemotherapies decay rate                              | $1.0 \times 10^{-2}$   | min <sup>-1</sup>                     |
| $v_1$        | Avastin plasma compartment volume                      | $2.66 \times 10^3$     | ml                                    |
| $v_2$        | Avastin peripheral compartment volume                  | $2.76 \times 10^3$     | ml                                    |
| $q$          | Avastin intercompartmental clearance                   | 0.412                  | ml/min                                |
| $cl$         | Avastin elimination clearance                          | 0.144                  | ml/min                                |
| $v_{max}$    | Fluororacil maximal degradation rate                   | 1.75                   | µg/ml/min                             |
| $k_m$        | Fluororacil half-maximal concentration                 | 27                     | µg/ml <sup>-1</sup>                   |
| $w_1$        | Epirubicin plasma compartment volume                   | $18 \times 10^3$       | ml                                    |
| $w_2$        | Epirubicin peripheral compartment volume               | $957 \times 10^3$      | ml                                    |
| $w_3$        | Epirubicin peripheral compartment volume               | $25 \times 10^3$       | ml                                    |
| $q_2$        | Epirubicin intercompartmental clearance                | $0.918 \times 10^3$    | µg/ml/min                             |
| $q_3$        | Epirubicin intercompartmental clearance                | $0.25 \times 10^3$     | µg/ml/min                             |
| $cl_2$       | Epirubicin elimination clearance                       | $0.983 \times 10^3$    | µg/ml/min                             |
| $u$          | Cyclophosphamide plasma compartment volume             | $2430 \times 10^3$     | ml                                    |
| $cl_3$       | Cyclophosphamide elimination clearance                 | $3.93 \times 10^3$     | µg/ml/min                             |
| $\Delta x$   | Space interval   | 10                     | µm                                    |
| $\Delta t$   | Time interval of cell cycle update                     | 30                     | min                                   |
| $\Delta v$   | Vessel update interval                                 | 720                    | min                                   |
| $Low_V$      | Lower VEGF angiogenic threshold                        | $10^{-6}$              | µg ml <sup>-1</sup>                   |
| $\alpha$     | FEC dose-response shape                                | 1                      | Dimensionless                         |
| $T_{min}$    | Minimum cell cycle duration                            | 3.74                   | Days                                  |
| $k_1$        | Basal p53 synthesis rate                               | 0.0004                 | min <sup>-1</sup>                     |
| $k'_2$       | Maximal p53 effect in VEGF production                  | -0.0002                | min <sup>-1</sup>                     |
| $d_{G^i}$    | Fluorouracil dose                                      | 600                    | mg m <sup>-2</sup>                    |

(Continues)



TABLE 1 (Continued)

| Parameter | Description           | Value | Units              |
|-----------|-----------------------|-------|--------------------|
| $d_{G^1}$ | Epirubicin dose       | 100   | mg m <sup>-2</sup> |
| $d_{G^3}$ | Cyclophosphamide dose | 600   | mg m <sup>-2</sup> |
| $d_A$     | Bevacizumab dose      | 0     | mg m <sup>-2</sup> |

### Algorithm 1 Numerical solution of the coupled PDE-ODE-CA model

- 1: Initialize computational meshes  $\mathcal{T}_H$  corresponding to the regular square lattice  $L$  (cf. 2.2).
- 2: Initialize the vascular lattice representation  $\mathcal{G}(\cdot, 0)$  and cellular lattice representation  $\mathcal{L}(\cdot, 0)$  based on available patient data.
- 3: Compute the initial oxygen concentration  $K(\cdot, 0)$  by solving (17) with  $s_K = 0$ .
- 4: Define the initial condition for the cell cycle  $\phi$  based on a uniform random distribution:

$$\phi_{\mathbf{x}}(0) \in \mathcal{U}(0, 1) \quad \forall \mathbf{x} \in L,$$

and compute the intracellular levels of  $P(t)$  and  $V_c(t)$  at initial time ( $t = 0$ ) by solving (10) and (11), respectively, using the initial oxygen concentration  $K(\cdot, 0)$ .

- 5: Define the initial chemotherapy concentrations by setting  $G_j(\cdot, 0) = 0$  for  $j = 1, 2, 3$ . Define initial conditions for  $V$ ,  $A$  and  $C$  by setting  $V(\cdot, 0) = 0$ ,  $A(\cdot, 0) = 0$  and  $C(\cdot, 0) = 0$ .
- 6: Compute the intravascular chemotherapy concentrations  $G_{j,0}(t^i)$ ,  $j = 1, 2, 3$  and intravascular VEGF-inhibitor concentrations  $A_1(t^i)$  by solving (4)-(9) given doses  $d_k$ ,  $k = 1, 2, 3 \quad \forall t^i \leq T$
- 7: Set  $t^1 = \Delta t = 30$  min,  $n = 1$
- 8: **while**  $t^n \leq T$  **do**
- 9:     Compute  $K(\cdot, t^n)$  by solving (17) with  $s_K = 1$ .
- 10:     Compute  $V(\cdot, t^n)$ ,  $A(\cdot, t^n)$ ,  $C(\cdot, t^n)$  by solving eqs. (19) to (21).
- 11:     Compute  $G^j$  for  $j = 1, 2, 3$  by solving (18).
- 12:     Compute  $P(\cdot, t^n)$  and  $V_c(\cdot, t^n)$  by solving (10) and (11) and compute  $\phi(\cdot, t^n)$  by solving (12) using  $K(\cdot, t^n)$ .
- 13:     Compute  $\mathcal{L}(\cdot, t^n)$  and  $\mathcal{G}(\cdot, t^n)$  based on the cellular automaton rules described in section 2.2.1.
- 14:     Set  $t^n = t^{n-1} + \Delta t$ ,  $n = n + 1$ .
- 15: **end while**

definite, and were solved (with optimal complexity) using a conjugate gradient (CG) solver with algebraic multigrid (BoomerAMG<sup>15</sup>) preconditioning with a relative solver tolerance of  $10^{-8}$ .

#### Solving the oxygen concentration equation

We consider the time-dependent version of the nonlinear oxygen concentration Equation (17) (with  $s_K = 1$ ). As for the chemotherapy equations, we first discretize the PDE by the implicit second-order Crank–Nicolson scheme in time. The resulting nonlinear system of differential equations at each timestep  $t^k$  is discretized using continuous piecewise linear finite elements relative to the mesh  $\mathcal{T}_H$ . We solve the resulting nonlinear system using a Newton iteration, with a tolerance of  $10^{-8}$ . The inner loop linear systems are symmetric and positive definite, and were solved with a CG Krylov solver with Jacobi preconditioning, and a relative solver tolerance of  $10^{-8}$ .



We also used a linear approximation of the nonlinear system that could lead to significant CPU time reduction. In particular, the non-linear term at timestep  $t^k$  can be approximated by  $\frac{K(\mathbf{x}, t^k)}{K_1 + K(\mathbf{x}, t^{k-1})}$ , where  $K(\mathbf{x}, t^{k-1})$  denotes the solution of oxygen at time  $k-1$ . The resulting linearized system of differential equations at each timestep  $t^k$  is discretized using continuous piecewise linear finite elements relative to the mesh  $\mathcal{T}_H$ . The resulting linear system of equation is symmetric and positive definite, and were solved (with optimal complexity) using a CG solver with algebraic multigrid (BoomerAMG<sup>15</sup>) preconditioning with a relative solver tolerance of  $10^{-8}$ . We found that the difference between the non-linear and linear approximations to be negligible (with a difference of less than 0.03%).

#### *Solving the coupled system equations of VEGF/Avastin complex*

The system of PDEs was discretized by the implicit second-order Crank–Nicolson scheme in time. The resulting nonlinear system of differential equations at each timestep  $t^k$  is discretized using continuous piecewise linear finite elements relative to the base mesh  $\mathcal{T}_H$ . We solve the resulting nonlinear system using Newton's method, and the (non-symmetric) linear systems were solved with an iterative Krylov (GMRES) solver with Jacobi preconditioning. The stopping criteria for the Newton and Krylov solvers were set to  $10^{-8}$  for both.

The coupled system of VEGF/Avastin complex can be simplified when Avastin is not administered. In this case, the PDE system reduces to a linear PDE in  $V$  alone. The resulting linear equation is symmetric and positive definite, and were solved (with optimal complexity) using a CG solver with algebraic multigrid (BoomerAMG<sup>15</sup>) preconditioning with a relative solver tolerance of  $10^{-8}$ .

## 2.4.2 | Numerical treatment of ODEs

The ODEs for the intravascular module and cell cycle can be solved analytically<sup>13</sup> and evaluated at the necessary time points. For the coupled system of the intracellular module, solution to  $P(\mathbf{x}, t_k + dt)$  can be written in explicit form given solution  $P(\mathbf{x}, t_k)$  at  $t = t_k$ . For  $V_c(\mathbf{x}, t)$ , VODE solver with implicit Adams method from SciPy package was chosen for the non-stiff problem for each cell. This solver uses an adaptive time step, and the maximum time step was set to 1min..

## 2.5 | Numerical treatment of cellular automaton

The biological cells and vessels are in the equations above represented as delta functions, and these delta functions are here numerically approximated with continuous, piecewise linear functions defined relative to the refined mesh  $\mathcal{T}_H$ . Such an approximated delta function,  $\delta(\mathbf{x} - \mathbf{x}_1)$ , will take the form of a right square pyramid centered at  $\mathbf{x}_1$  with the value of 0 at all vertices, except at  $\mathbf{x}_1$  where it will take the value that makes the integral over the approximated delta function become 1 (as it should for a delta function).

For each cancer cell at  $\mathbf{x} = (x_i, x_j) \in \mathcal{L}$  with cell cycle  $\phi_x \geq 1$ , the cell is killed with probability drawn from a cumulative Beta distribution  $Beta(1, \beta), \beta > 0$ . If the cell is not killed, a new cell is placed at the lattice site  $\mathbf{x}' \in (x_p, x_q), p = i-1, \dots, i+1, q = j-1, \dots, j+1$  with the highest oxygen concentration  $K$ . No cell is placed if no space is available in the neighborhood, i.e. if all (Moore) neighbor sites are occupied by other cells. If more than one cell attempts to place a new cell at the same lattice site, the cell with the highest  $\phi$  wins. The new cell inherits its minimum cell cycle duration  $T_{\min}$ . Next,  $\phi_x$  and  $\phi_{x'}$  are set to  $(\phi_x \bmod 1)$  and 0 respectively. The intracellular VEGF and TP53 levels  $P, V_c$  are set to zero at both  $\mathbf{x}$  and  $\mathbf{x}'$  at time  $t$ . The sequence of updates for proliferation-ready cells is asynchronous such that the new state of a cell do not affect the calculation of states in neighboring cells.

## 2.6 | Parallelization of hybrid cellular automaton and finite element models

### 2.6.1 | Linking spatially parallel continuous and discrete modules

The computational expense of the whole algorithm is dominated by the operations within each time step. Each time step consists of “continuous update”—solving a sequence of PDEs and ODEs, and “discrete update”—computing states

of cellular automaton. Due to the inherent dependencies between the continuous and discrete components, parallelism in the time variable is more difficult to approach. Instead, since the solvers in the FEniCS system are spatially parallel, we consider a fork-join strategy. They are simple to integrate with the rest of the algorithm as long as the communication overhead is kept to a minimum.

Standard domain decomposition techniques, iterative linear solvers, and preconditioners were used to parallelize the PDE assembly and solvers. The parallelization of the ODEs was trivial: they are made up of spatially decoupled problems that are described on each biological cell and can thus be solved independently. The parallelization of the cellular automaton models necessitated further consideration because they are dependent on the oxygen levels in neighboring biological cells. Retrieving that information requires inter-process communication if that cell is at the boundary of the domain decomposition. Cells are classified into two categories based on their location within the sub-domain. Cells located inside a sub-domain are marked as internal, while those located at the boundary of two or more sub-domains, shared by two, three or four processors are marked as inter-facial. To achieve this, we added a setup routine (Algorithm 2) to the algorithm that allows parallel processes to determine which process owns their neighboring biological cells. These neighbor maps allowed the cellular automaton to retrieve neighboring biological cell values with minimal overhead.

### 2.6.2 | Choosing a scalable parallel random number generator

Since the model is stochastic, having a scalable parallel random number generator (RNG) is an important aspect of parallelization. To investigate and understand the simulation results, we must be able to reproduce the same scenarios and obtain the same confidence intervals each time the same stochastic experiment is performed; when debugging such parallel stochastic application, we also need to reproduce the same result to correct any anomalous behavior. Furthermore, for applications in predicting treatment effect, it is important to produce unbiased results and claim statistical significance. Implementation of pseudo RNG also reduces the number of simulations needed for claiming statistical significance in cross-comparison between different treatment regimens since the simulation would be the same until the point of divergence induced by different regimens.

Pseudo RNG in high-performance computing (HPC) applications, particularly in our HCA model, requires the following criteria. It cannot download, or store the hundreds of thousands of numbers needed to reproduce the experiments due to constraints of the server. The random number sequence assigned to each cell or vessel must not depend on the number of processors, that is, each cell or vessel should autonomously obtain either its own random sequence or its own sub-sequence of a global sequence. The pseudo RNG must also come with good statistical properties, approximating as close as possible a

#### Algorithm 2 Implementation of communication between neighbour cells on different processes

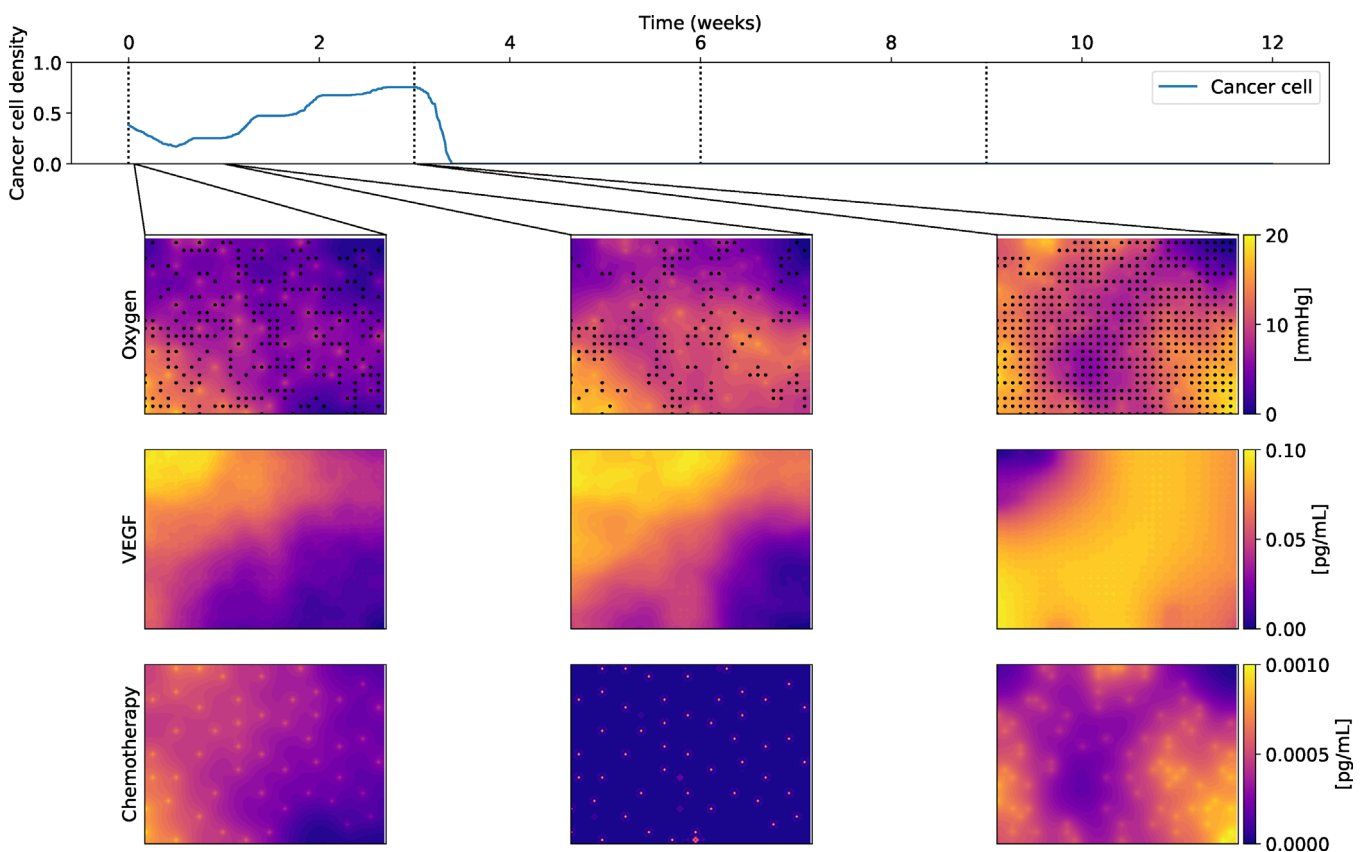
- 1: MPI Initialization
- 2: Make a map containing the coordinates on the mesh and the process of all eight neighbour cells for each cell.
- 3:
- 4: **for** each time step **do**
- 5:     Get oxygen concentration  $\forall \mathbf{x} \in \mathcal{L}(\mathbf{x})$
- 6:     Get and update  $\phi(\mathbf{x})$
- 7:     Gather  $x$  such that  $\phi(\mathbf{x}) > 1$
- 8:     update cells according to cellular automaton rules
- 9:
- 10:     **if** more than one processor attempt to update at the same  $\mathbf{x}$  **then**
- 11:         update  $\mathcal{L}(\mathbf{x})$  from the processor with higher  $\phi(\mathbf{x}')$  value  $\forall \mathbf{x}'$  in the neighbour
- 12:         Set automata, vegf, p53 and cell cycle value of the new cancer cell equal to the parent cancer cells
- 13:     **end if**
- 14: **end for**
- 15: Update map

truly random sequence. The RNG should also be memory-efficient as sequences of the RNG will be the length of any given simulation, and the size of the RNG grows at the rate of the total number of vertices. We have therefore chosen permuted congruential RNG<sup>16</sup> which partitions the main sequence of the generator into sub-sequences for the sake of memory-efficiency, good statistical properties, small state and a more than sufficient period.

### 3 | RESULTS

#### 3.1 | Small-scale simulations

We first run personalized simulations of a patient analyzed previously with a previous version of the model and using a non-parallel solver.<sup>9</sup> These simulations correspond to a  $200\ \mu\text{m} \times 300\ \mu\text{m}$  tissue section, a size that allowed us to use exact initial cell positions available from histopathological slides. We use here the same initial conditions and previously published model parameters<sup>9</sup> (see Table 1). As both the model and the solver are now extended, we do not expect to get the same solutions as in previous work. We tested different perfusion conditions estimated from MRI images and confirmed that the treatment outcome after 12 weeks strongly depends on the perfusion conditions used. In Figure 2, we show a simulation representing a high perfusion condition estimated at the outer edge of the tumor. In the simulation the initial cell density is reduced by half after the first chemotherapy shot at week 0. Chemotherapy is efficiently distributed in space but it is washed out after a few days. As approximately half of the cells are not dividing while the chemotherapy is available, they are not killed and the number of cells grows significantly before the second chemotherapy shot at week 3. After week 3 instead, all cancer cells are killed as they divide when enough chemotherapy is available. Comparing with previous simulations, the final outcome is matched but the transients are slightly different. We



**FIGURE 2** Simulation of a small tissue cross section of size  $200\ \mu\text{m} \times 300\ \mu\text{m}$ . The top panel depicts the evolution of the cancer cell density over the course of 12 weeks of treatment. Chemotherapy shots are given at weeks 0, 3, 6 and 9 and are illustrated by a vertical dotted line in the figure. The lower panels depict the spatial distribution of oxygen, VEGF and chemotherapy at weeks 0, 1, and 3 in the tissue section. The presence of cancer cells is shown in black dots in the oxygen panel. Since the vessels are point sources of those molecules, areas with high blood vessel density can be perceived in the oxygen and chemotherapy panels

attribute it mostly to solving time-dependent PDEs instead of the steady-state solver considered previously. This difference in solvers adds a factor of seven to the run time on a single core.

### 3.2 | Solver performance

To evaluate the runtime and scalability of the complete numerical solver, we consider a simulation of a  $1 \text{ mm}^2$  tissue slice ( $D = [0,1] \times [0,1] \text{ mm}^2$ ), with  $n = m = 100$ . Initial cell and vascular densities were assumed to be 0.4 and 0.12 respectively. These settings reflect a real patient simulation scenario.

#### 3.2.1 | Serial runtime profile

The average intra- and inter-update timing of the solver was first carried out on a single CPU core. The total timings, broken down by computational processes and by model modules, are shown in Figures 3 and 4.

We observe that the assembly of the discrete PDE operators dominates the runtime, specifically the time-dependent assembly of the non-linear PDEs (Figure 3A). As expected, the solution of the linear systems (PETSc Krylov solver) is also a substantial component. The other components and in particular the solution of the ODEs represent a nearly negligible cost. Overall, we found similar cost distribution comparing intra- and inter-update, indicating that communication during discrete updates were efficient.

The breakdown by modules shed some further insights on the cost of assembly: the solution of both VAC and Oxygen modules are obtained by solving non-linear PDEs, and the system is re-assembled for each  $\Delta t$ . To reduce the runtime of these modules, an explicit splitting scheme could be used instead (corresponding to a single step of the

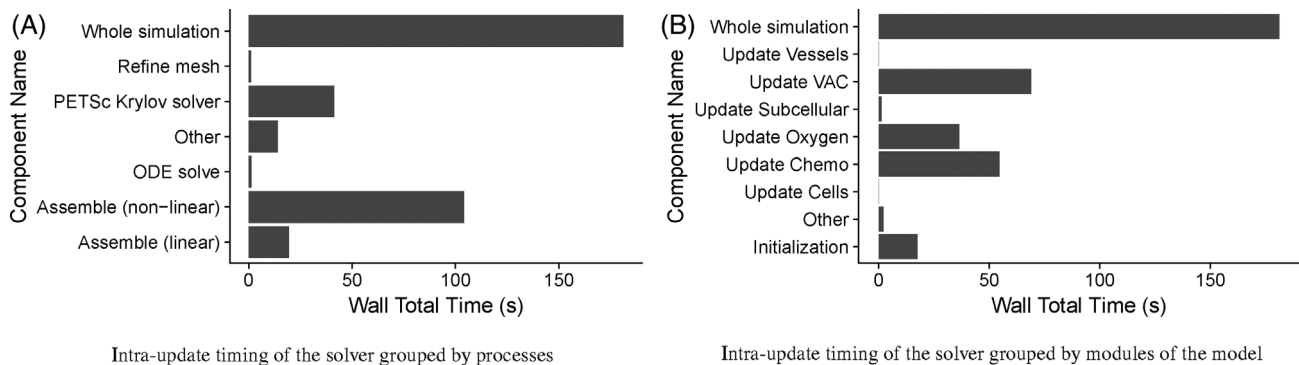


FIGURE 3 Intra-update timing of the model with  $\Delta t = 1 \text{ min}$  solved for  $t = 30 \text{ min}$

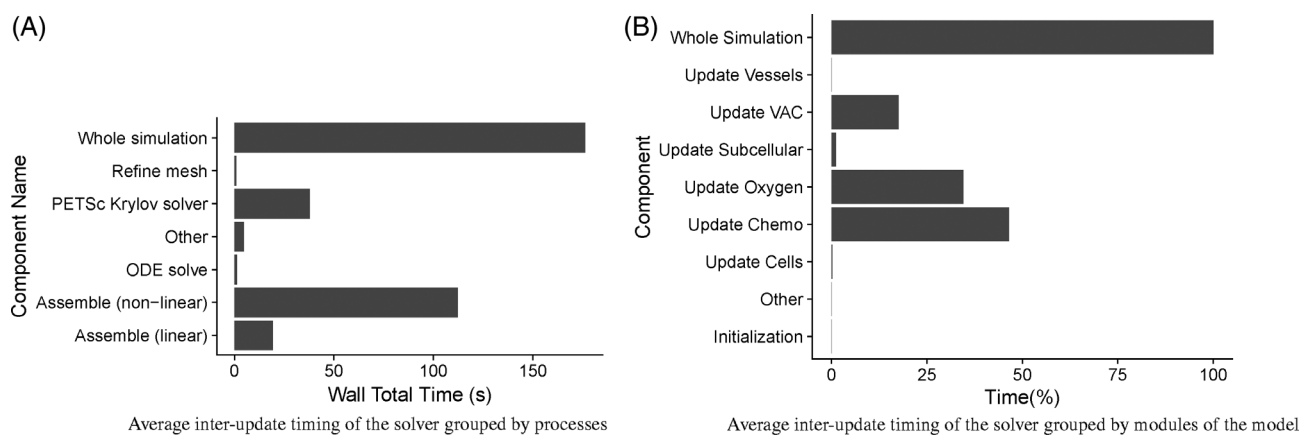


FIGURE 4 Average inter-update timing of the model with  $\Delta t = 30 \text{ min}$  solved over  $t = 120 \text{ min}$

Newton iteration) at the cost of numerical accuracy. Time spent on chemotherapy module comes second as there are three types of chemotherapies to be solved separately. Initialization includes mesh initialization as well as function initialization for each module. This is done only once at the beginning of the whole simulation, and is therefore expected to become negligible for longer simulations.

### 3.2.2 | Parallel scalability

We also investigated the parallel (weak) scalability of the assembly and solution of the PDEs. Weak scalability can be examined for a series of problem sizes by assigning a constant problem size to each processing element—typically for

**TABLE 2** Timing of temporal chemotherapy update ( $t = 30$  min)

| $N_{\text{proc}}$ | $N_{\text{it}}$ | $N_{\text{dofs}}$ | $\bar{N}_{\text{dof}}$ | $t_A$  | $\text{WSE}_A$ | $t_S$  | $\text{WSE}_S$ |
|-------------------|-----------------|-------------------|------------------------|--------|----------------|--------|----------------|
| 1                 | 5               | 1                 | 159,197                | 2.9177 | 1              | 6.2043 | 1              |
| 8                 | 6               | 8                 | 1,279,157              | 3.1415 | 1.0767         | 9.8944 | 1.5948         |
| 16                | 6               | 16                | 2,556,797              | 3.543  | 1.2143         | 12.911 | 2.081          |
| 24                | 6               | 24                | 3,837,677              | 3.6139 | 1.2386         | 14.752 | 2.3777         |
| 32                | 6               | 32                | 5,121,165              | 3.8919 | 1.3339         | 15.248 | 2.4577         |
| 40                | 6               | 40                | 6,405,957              | 3.7023 | 1.2689         | 15.502 | 2.4986         |
| 80                | 6               | 80                | 12,809,237             | 3.7708 | 1.2924         | 15.784 | 2.544          |

*Note:*  $N_{\text{proc}}$  is the number of processors,  $N_{\text{it}}$  is the number of Krylov iterations,  $N_{\text{dofs}}$  is the total number of degrees of freedom,  $\bar{N}_{\text{dof}}$  is the average number of degrees of freedom per process,  $t_A$  is assembly runtime,  $\text{WSE}_A$  is the weak scaling efficiency for the assembly,  $t_S$  is the linear solver runtime, and  $\text{WSE}_S$  the linear solver weak scaling efficiency.  $\Delta t$  was kept constant during the update and the results shown are the minimum of five repeated runs of the updates with random initialization of cell and vessel configurations. Intravascular concentration was kept constant during the simulation.

**TABLE 3** Timing of temporal oxygen update ( $t = 30$  min)

| $N_{\text{proc}}$  | $N_{\text{it}}$ | $N_{\text{dofs}}$ | $\bar{N}_{\text{dof}}$ | $t_A$  | $\text{WSE}_A$ | $t_S$  | $\text{WSE}_S$ |
|--|-----------------|-------------------|------------------------|--------|----------------|--------|----------------|
| <b>(a) Timing of temporal oxygen update (<math>t = 30</math> min)</b>            |                 |                   |                        |        |                |        |                |
| 1  | 12              | 159,197           | 159,197                | 13.112 | 1              | 44.012 | 1              |
| 8  | 12              | 1,279,157         | 159,894                | 15.211 | 1.1601         | 82.138 | 1.8663         |
| 16   | 12              | 2,556,797         | 159,799                | 15.528 | 1.1843         | 164.76 | 3.7435         |
| 24   | 11              | 3,837,677         | 159,903                | 15.802 | 1.2052         | 277.56 | 6.3065         |
| 32   | 11              | 5,121,165         | 160,036                | 15.753 | 1.2014         | 296.43 | 6.7352         |
| 40   | 10              | 6,405,957         | 160,036                | 18.601 | 1.419          | 322.40 | 7.3253         |
| 80   | 10              | 6,405,957         | 160,036                | 18.485 | 1.410          | 399.94 | 9.0871         |
| <b>(b) Timing of temporal linearized oxygen update (<math>t = 30</math> min)</b> |                 |                   |                        |        |                |        |                |
| 1  | 1               | 159,197           | 159,197                | 2.6657 | 1              | 3.0532 | 1              |
| 8  | 8               | 1,279,157         | 159,894                | 3.0377 | 1.1396         | 3.9054 | 1.2791         |
| 16   | 16              | 2,556,797         | 159,799                | 3.3315 | 1.2498         | 5.6732 | 1.8581         |
| 24   | 24              | 3,837,677         | 159,903                | 3.5559 | 1.3339         | 6.2294 | 2.0403         |
| 32   | 32              | 5,121,165         | 160,036                | 3.5187 | 1.32           | 6.6993 | 2.1942         |
| 40   | 40              | 6,405,957         | 160,148                | 3.7548 | 1.4086         | 6.6725 | 2.1854         |
| 80   | 80              | 12,809,237        | 160,115                | 3.5985 | 1.3499         | 6.8429 | 2.2412         |

*Note:*  $N_{\text{proc}}$  is the number of processors,  $N_{\text{it}}$  is the number of Krylov iterations,  $N_{\text{dofs}}$  is the total number of degrees of freedom,  $\bar{N}_{\text{dof}}$  is the average number of degrees of freedom per process,  $t_A$  is assembly runtime,  $\text{WSE}_A$  is the weak scaling efficiency for the assembly,  $t_S$  is the linear solver runtime, and  $\text{WSE}_S$  the linear solver weak scaling efficiency.  $\Delta t$  was kept constant during the update and the results shown are the minimum time of five repeated runs of the updates with random initialization of cell and vessel configurations.



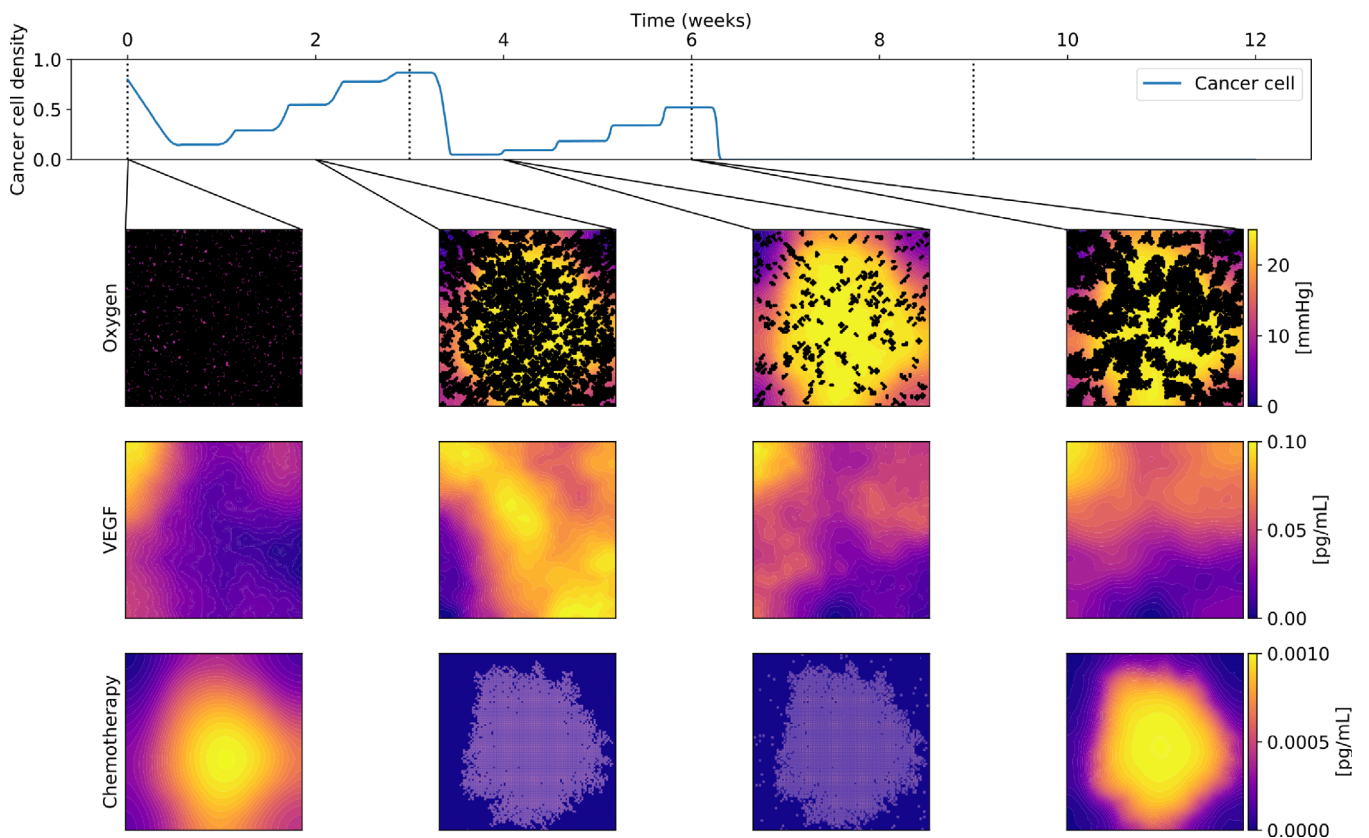
an increasing amount of processor cores. Perfect weak scaling is achieved if the run time stays constant while the workload is increased in direct proportion to the number of processors. The weak scaling factor was defined as

$$\text{WSF} = \frac{t_N}{t_1}, \quad (22)$$

where  $t_1$  and  $t_N$  is the runtime when using 1 and  $N$  processors, respectively, and was calculated for the assembly and solve times separately.

All weak scaling experiments were conducted on Saga, the high performance computing facility placed at NTNU in Trondheim, composed of Intel Xeon-Gold 6138 running at 2.0 GHz. Each Saga compute node consists of 40 physical compute cores with 192 GiB memory each, giving approximately 4 GiB memory per physical core. We used FEniCS version 2019.1.0 built with an Intel compiler. These scalability tests were run on two compute nodes totaling 80 physical cores, and with exclusive access to minimize communication overhead. For the sake of comparison, the end time of all temporal updates of PDEs in the model was set to  $t = 30$  min, while cell and vessel configurations were initialized randomly. The parallel runtime and scalability of the chemotherapy equations are presented in Table 2. We observe a nearly optimal WSF ( $\leq 1.3$ ) for the finite element assembly up to 80 cores. For the linear solver, we observe a moderate increase in the WSF up to 2 for 16 cores and 2.5 for 80 cores. The number of iterations of the Krylov solver stay constant however.

For the solution of the oxygen concentration equation, the parallel scalability timings are shown in Table 3. We note that the solution time is around four times that of the chemotherapy equations and thus a dominant contribution. The finite element assembly scales well up to 80 cores ( $\text{WSE}_A \leq 1.4$ ) for both non-linear and the linearized solver. However, by linearizing the oxygen equation, the weak scalability of the linear solver is improved from a WSF of up to 9 in 80 cores to be comparable to the results for the chemotherapy equations (WSE less than 2.24). We also investigated the



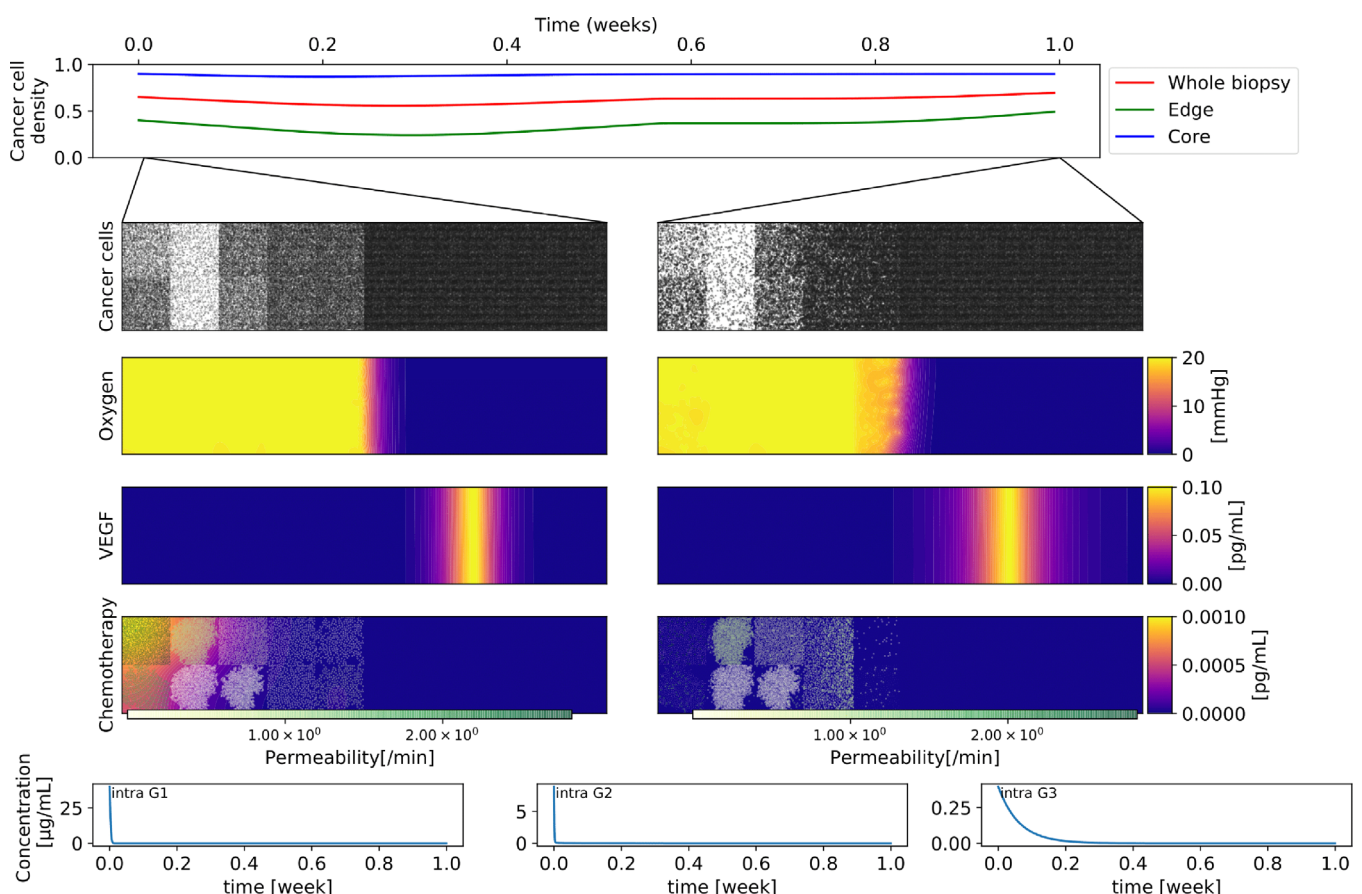
**FIGURE 5** Simulation of a patient biopsy of size  $1 \text{ mm} \times 1 \text{ mm}$ , equivalent to the size of a MRI voxel. The top panel depicts the evolution of the cancer cell density over the course of 12 weeks of treatment. Chemotherapy shots are given at weeks 0, 3, 6 and 9 and are illustrated by a vertical dotted line in the figure. The lower panels depict the spatial distribution of oxygen, VEGF and chemotherapy at weeks 0, 2, 4 and 6 in the tissue section. The presence of cancer cells is shown in black dots in the oxygen panel. Since the vessels are point sources of those molecules, areas with high blood vessel density can be perceived in the oxygen and chemotherapy panels

weak parallel scalability of the VAC equations. The results (data not shown) were highly comparable to the scalability of the oxygen concentration solver.

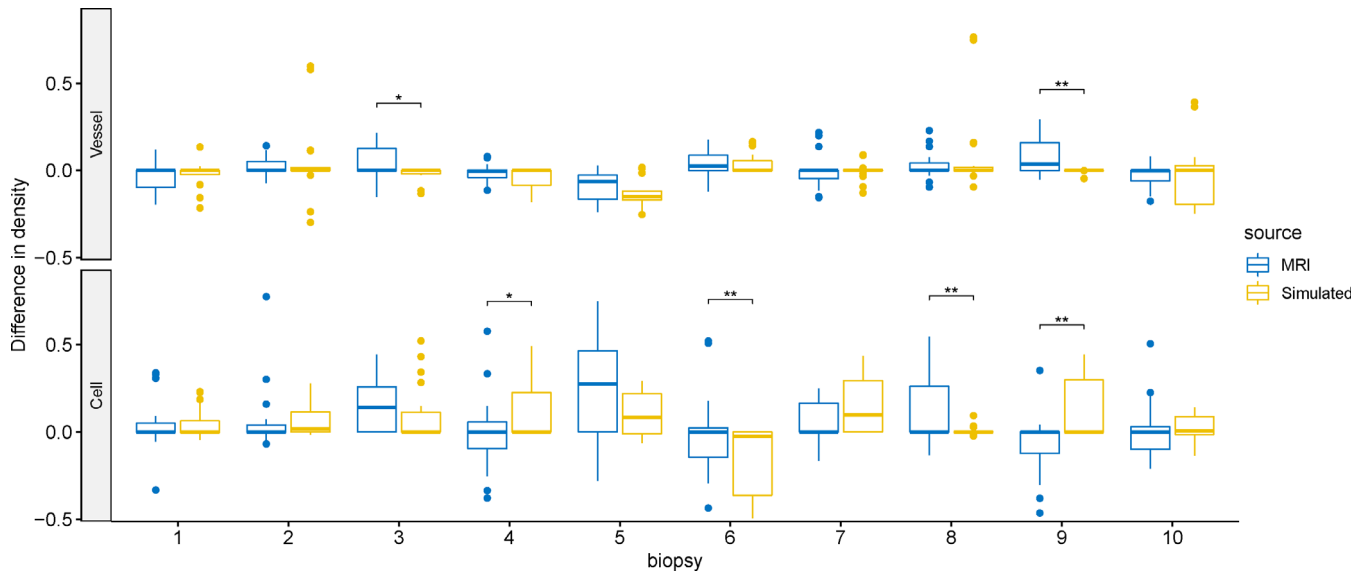
### 3.3 | Simulating a cross-section of a MRI voxel of tissue

The spatial resolution of the available MRI data is defined by the imaging voxels of size  $1 \text{ mm}^3$ . Thus, important biological information, such as the perfusion of the tumor is only given per voxel of tissue. As an intermediate step to run multi-voxel simulations, here we demonstrate the ability of our parallel solver to run a grid-size equivalent to a cross-section of a MRI voxel.

Figure 5 illustrates the cancer cell trajectory under 12 weeks of chemotherapy treatment. Since digitalized biopsy data detailing the exact distribution of cancer and stroma cells are available only for the small tumor section of size  $200 \mu\text{m} \times 300 \mu\text{m}$ , we used the same cell density to approximate the number of cancer and stroma cells in the larger one, but they were randomly distributed in the grid to initialize the simulation. We again use the same model parameters detailed in Table 1. Comparing Figure 5 to Figure 2, we observe similar patterns. Immediately following the initial administration of chemotherapy, the number of cancer cells starts to decrease, reaching a minima about 2.5 days later. At that point, the chemotherapy treatment loses its effect and the cancer cells start to grow back again. This pattern is repeated after the second chemotherapy shot at 3 weeks, while the third shot at 6 weeks kills all existing cancer cells. The panels at weeks 0, 2, 4 and 6 in Figure 5 show the spatial distribution of cells, oxygen, VEGF, blood vessels and



**FIGURE 6** Simulation of a digital biopsy sample of size  $2 \text{ mm} \times 1 \text{ cm}$  from week 0 to week 1. The top panel depicts the evolution of the cancer cell density of the whole, the edge and the core of the simulated biopsy over the course of 1 week of treatment. This is followed by spatial comparative view of cancer cells distribution and three extracellular variables, from top to bottom, oxygen, VEGF, and chemotherapy between week 0 (left panels) and week 1 (right panels). In the first row, locations of simulated cancer cells were marked in black, while locations of simulated vessels were overlaid in color to represent their permeability in chemotherapy. Intravascular concentration of the three chemotherapies between week 0 and week 1 are shown at the bottom



**FIGURE 7** Differences in cell densities between weeks 0 and 1 are depicted as box-plots. Simulated and actual cell densities from magnetic resonance imaging (MRI) were seen in blue and yellow, respectively. The x-axis shows the comparison of 10 biopsy slices. The first box-plot from the left compares the simulation shown in Figure 6 to MRI data. Paired Wilcoxon signed rank test were performed between simulated and actual cell densities for each biopsy. Asterisks were placed above the boxes of comparisons of p-values less than 0.05. \* $p < .05$ , \*\* $p < .01$

**TABLE 4** Results from the fitted mixed effect model. Intercept and the source of the cell/vessel density (simulated or actual) were considered as the fixed effects; the location of each MRI voxel and the biopsy sample ID was considered as nested random effects to reflect the variation in biopsy samples in either simulated or actual vessel and cell densities. The regression coefficients estimates, standard error as well as the significance of the two fixed effects were shown in the table

| Cells  | Estimate    | Std. error | t-Value    | p-Value |
|--|-------------|------------|------------|---------|
| (Intercept)                                  | 0.04917292  | 0.02014897 | 2.4404685  | 0.0151  |
| Simulated density compared to actual density | -0.01138260 | 0.02021387 | -0.5631087 | 0.5737  |
| Vessels                                      | Estimate    | Std. error | t-Value    | p-Value |
| (Intercept)                                  | 0.003049063 | 0.01643971 | 0.1854693  | 0.8530  |
| Simulated density compared to actual density | 0.002905701 | 0.01412006 | 0.2057853  | 0.8371  |

chemotherapy in the simulated tumor cross-section. As expected, the tissue oxygen concentration increases and VEGF concentration decreases when the number of cancer cells decreases. Due to a good tissue perfusion with a large number of vessels and high permeability, chemotherapy reaches the whole tissue at effective concentrations. The total runtime of this 12-week simulation on Saga took approximately 12 h with 80 cores.

### 3.4 | Simulating a cross-section of a digital biopsy sample

To ultimately demonstrate the impact of heterogeneous perfusion conditions on the dynamics of cell growth and killing, we constructed a mesh of size equivalent to  $2 \times 10$  MRI voxels. This matches with the size of the core needle biopsy taken at the beginning of the treatment in the clinical trial and we will refer to it as a digital biopsy sample. Perfusion related parameters were applied in each of the 20 simulated voxels of a digital biopsy sample as follows. First, the MRI slice with the widest tumor diameter is selected. This image reveals a difference in perfusion conditions between the tumor edge and the core. Starting from the boundary between the core and the edge of the tumor, 5 voxels were chosen radially inwards and outwards respectively, totaling 10 voxels. Another 10 voxels were chosen in the same fashion from their immediate neighbor, resulting in a simulated biopsy of  $2 \times 10$  voxels. Estimated perfusion parameters, specifically



$k_{\text{trans}}$  and  $v_p$ , from the extended Tofts model were then calculated individually to each voxel. Since exact distribution of vessels, cancer and stroma cells are currently available only for small tumor sections, we used  $v_p$  and  $v_c = 1 - (v_p + v_e)$  (where  $v_e$  is the estimated fractional extravascular extracellular space volume), from applying the extended Tofts model to the MRI data as estimates of the vessel and cell density respectively. Vessels and cells were then randomly distributed accordingly within each voxel in the grid ( $100 \times 100$  grid points) to initialize the simulation.

Figure 6 shows 1-week evolution of one digital biopsy sample under therapy. We can clearly see the complex effect of the heterogeneous perfusion condition illustrated by the heterogeneity in the oxygen concentration. As chemotherapy reaches the tissue, active cancer cells are killed at different rates in various areas of the simulated biopsy. In areas with high perfusion, cells are killed more rapidly, while in those areas with low perfusion, chemotherapy cannot be delivered sufficiently, and more cancer cells survive. In areas with no vessels, cell growth is stalled and bounded by the capacity of the computational grid and the oxygen. In areas with lower initial cell density, cells were dividing albeit at a lower rate. In addition, a combined effect in areas between high and low perfusion is observed. Drug delivery to areas with low perfusion reaches neighboring areas of high perfusion by diffusion instead of via the vasculature. Moreover, dense cell populations from low perfusion areas can grow into areas with high perfusion where more cancer cells were killed by the chemotherapy. Total runtime of the simulation was on average 18 h with 80 cores.

To assess the validity of our model, we extracted 10 biopsy samples using the same method described above. We use data from week 0 of each digital biopsy to initialize the model and differences in cell density between week 0 and week 1 are evaluated between the simulation and the MRI data for each voxel within the 10 biopsy samples (Figure 7). Given that all digital biopsies were collected from the same tumor, a mixed effect model was used to verify if there is any statistical difference between simulated and actual densities. We found no significant difference, thus proving that the numerical simulations can reproduce the observed drug outcome in large, clinically relevant heterogeneous tumor portions (Table 4).

## 4 | DISCUSSION

In this work, we present a multiscale HCA model for personalized breast cancer growth under the effect of therapy, together with a parallel numerical algorithm aimed at large scale, clinically-relevant simulation sizes. We discovered that the PDE solvers are the most computationally intensive component of the system, with non-linear coupled equation updates dominating the runtime. In terms of solver performance, the assembly of discrete PDE operators scales almost linearly, while a moderate increase in weak scaling efficiency was observed in linear solver runtime. We simulated tumor dynamics and drug therapies of tumor cross-sections in three different system sizes: a small tumor portion, an MRI voxel of tissue, and a full biopsy sample consisting of 20 MRI voxels. Our results show that the parallel implementation of our model can account for tumor heterogeneity at a clinically-relevant system size while correctly predicting treatment outcome.

Previously, multiple cancer modeling studies have used HCA models that couple CA with ODEs and PDEs, see for instance.<sup>3,4,17,5,9</sup> However, they typically considered small 2D or 3D tumor portions and account for only small number of cells. Multi-scale models representing clinically relevant tumor portions, like the one considered in this study, require parallel computing, specially if they include many PDEs describing the tumor microenvironment or very large ODE systems for cell signaling. A number of computational frameworks for simulating multi-scale models with parallel computing capabilities are available,<sup>18,7</sup> including Morpheus,<sup>19</sup> CompuCell3D,<sup>20</sup> PhysiCell,<sup>21</sup> CellSys,<sup>22</sup> Chaste,<sup>23</sup> Biocellion<sup>24</sup> or Timothy.<sup>25</sup> Each of these tools has their own specific features regarding modeling formalism, implementation, usability and performance. Morpheus and CompuCell3D use cellular Potts model as cell-based model formalism while CellSys, PhysiCell, Biocellion and Timothy use other off-lattice/cell centered approaches. Chaste has more flexibility and allows the user to choose their own cell-based formalism, including CA like in our case. In terms of performance, several simulation frameworks allow multithreading via OpenMP, including Morpheus, CompuCell3D, PhysiCell and CellSys. Biocellion and Timothy, instead, were designed for intense parallelisation between nodes and allow domain decomposition techniques as our solver. These tools can be used for simulating large numbers of PDEs in large domains and up to billions of biological cells in high performance supercomputers. In order to run personalized simulation of breast cancer therapy considered in our previous work<sup>9</sup> but considering clinically relevant tumor portions, we have taken this later approach. Specifically, we build on the domain decomposition capabilities of FEniCS for solving PDEs and efficiently link the stochastic CA models with the continuous PDE and ODE models. To the best of our knowledge,

this is the first time that the popular finite element framework FEniCS is extended for simulating a multi-scale HCA model.

In terms of limitations, we used estimated cell density from MRI rather than actual cell distribution captured from a core biopsy staining for cell initialization. This is due to the MRI dataset's inability to properly identify cancer and stroma cells. We do not take into account local differences in cell density in each voxel. As a result, in some cases, simulated treatment outcomes can vary from the actual outcomes. We leave for future work the digitization of the pathological sample and integration of cell identification. Modeling other cell types and their interactions is possible but we did not account for it in this work because the current clinical data cannot inform them on a patient-specific basis. Each MRI voxel contains a large number of vessels, each with a different surface permeability and vascular flow, both of which are important for drug delivery and response. In this study, a voxel was treated as a region of uniform vascular flow and vascular permeability. Improving on this will necessitate MRI data with a higher spatial resolution than that available in the clinical trial under consideration here.

Steady-state models were used in our previous implementation. They are often understood as a simplification of the time evolution ones, with the assumption that the time-dependent solution stabilizes around the steady-state as  $t \rightarrow \infty$ . In our study, where temporal accuracy is essential, this assumption is invalid for slow perfusion scenarios. However, since it is beyond the reach of our work, we have only put in a limited amount of effort to tune the multigrid preconditioners for this problem in order to improve simulation run-time of 1 week of actual treatment, which is roughly 6 h using 80 cores in cluster).

The ability to simulate clinically relevant pieces of breast tumors under therapy is a significant step toward in silico design of patient-specific treatment plans. However, we have so far only applied this methodology to five patients and much more testing and validation is needed before this methodology can assist oncologist in finding effective therapies for each patient. Once models become sophisticated enough to predict treatment outcomes for a wide cancer patient population, one possibility to test them in a clinical setting is in silico-guided clinical trials for personalized cancer medicine. The goal of such trials is to compare a standard approved therapy with personalized drug schedules optimized through computer simulations.

Improved statistical inference methods that allow successful estimation of patient-specific parameters are needed to achieve this using complex models like the one presented here. We have taken the first steps in that direction by employing Approximate Bayesian computation,<sup>26,27,10</sup> a very computationally intensive technique that certainly benefits from the scalable solver presented in this study.

## ACKNOWLEDGMENTS

The simulations were performed on resources provided by UNINETT Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway. This project received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 847912. The project received funding from the UiO: Life Science initiative through the convergence environment grant PerCaThe. X.L., A.K.L. and A.F. were supported by the center for research-based-innovation BigInsight under project number 237718. We acknowledge funding from the Research Council of Norway with project number 294916 and 311188. The authors also acknowledge the Centre for Digital Life Norway for supporting the partner projects PerCaThe, PINpOINT and BigInsight. M.E.R has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement 714892.

## ORCID

Xiaoran Lai  <https://orcid.org/0000-0002-7364-3993>

Simon W. Funke  <https://orcid.org/0000-0003-4709-8415>

## REFERENCES

1. Rockne RC, Hawkins-Daarud A, Swanson KR, et al. The 2019 mathematical oncology roadmap. *Phys Biol*. 2019;16(4):041005.
2. Ribba B, Alarcón T, Marron K, Maini PK, Agur Z. *The Use of Hybrid Cellular Automaton Models for Improving Cancer Therapy*. Springer.; 2004:444-453.
3. Alarcón T, Byrne HM, Maini PK. A multiple scale model for tumor growth. *Multiscale Model Simul*. 2005;3(2):440-475.
4. Gerlee P, Anderson AR. An evolutionary hybrid cellular automaton model of solid tumour growth. *J Theor Biol*. 2007;246(4):583-603.
5. Powathil GG, Gordon KE, Hill LA, Chaplain MA. Modelling the effects of cell-cycle heterogeneity on the response of a solid tumour to chemotherapy: biological insights from a hybrid multiscale cellular automaton model. *J Theor Biol*. 2012;308:1-19.

6. Cilfone NA, Kirschner DE, Linderman JJ. Strategies for efficient numerical implementation of hybrid multi-scale agentbased models to describe biological systems. *Cell Mol Bioeng*. 2015;8(1):119-136.
7. Fletcher AG, Osborne JM. Seven challenges in the multiscale modeling of multicellular tissues. *WIREs Mechanisms of Disease*. 2021; e1527. <http://dx.doi.org/10.1002/wsbm.1527>
8. Alnæs MS, Blechta J, Hake J, et al. The FEniCS project version 1.5. *Arch Numer Softw*. 2015;3(100):9-23.
9. Lai X, Geier OM, Fleischer T, Garred Ø, Borgen E, Funke SW, Kumar S, Rognes ME, Seierstad T, Børresen-Dale A, Kristensen VN, Enggebraaten O, Köhn-Luque A, Frigessi A. Toward Personalized Computer Simulation of Breast Cancer Treatment: A Multiscale Pharmacokinetic and Pharmacodynamic Model Informed by Multitype Patient Data. *Cancer Research*. 2019;79(16):4293-4304. <http://dx.doi.org/10.1158/0008-5472.can-18-1804>
10. Lai X. *Modelling, inference and simulation of personalised breast cancer treatment*. PhD thesis. Univeristy of Oslo; 2019.
11. Mann RM, Balleyguier C, Baltzer PA, et al. Breast MRI: EUSOBI recommendations for women's information. *Eur Radiol*. 2015;25(12):3669-3678.
12. Owen MR, Stamper IJ, Muthana M, et al. Mathematical modeling predicts synergistic antitumor effects of combining a macrophage-based, hypoxia-targeted gene therapy with chemotherapy. *Cancer Res*. 2011;71(8):2826-2837.
13. Dubois A, Bertrand J, Mentré F. Mathematical expressions of the pharmacokinetic and pharmacodynamic models implemented in the PFIM software. Université Paris Diderot and INSERM; 2011.
14. Tofts PS, Brix G, Buckley DL, et al. Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols. *J Magn Reson Imaging*. 1999;10(3):223-232.
15. Falgout RD, Yang UM. *Hypre: A Library of High Performance Preconditioners*. Springer; 2002:632-641.
16. O'Neill ME. PCG: A Family of Simple Fast Space-Efficient Statistically Good Algorithms for Random Number Generation. Tech. Rep. HMC-CS-2014-0905, Harvey Mudd College; Claremont, CA: 2014. *Tech. Rep.* 2014.
17. Perfahl H, Byrne HM, Chen T, et al. Multiscale modelling of vascular tumour growth in 3D: the roles of domain size and boundary conditions. *PLoS One*. 2011;6(4):e14790.
18. Metzcar J, Wang Y, Heiland R, Macklin P. A review of cell-based computational modeling in cancer biology. *JCO Clin Cancer Inform*. 2019;2:1-13.
19. Starruß J, Back DW, Brusch L, Deutsch A. Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics*. 2014;30(9):1331-1332.
20. Swat MH, Thomas GL, Belmonte JM, Shirinifard A, Hmeljak D, Glazier JA. *Multi-Scale Modeling of Tissues Using CompuCell3D*. Elsevier; 2012:325-366.
21. Ghaffarizadeh A, Heiland R, Friedman SH, Mumenthaler SM, Macklin P. PhysiCell: an open source physics-based cell simulator for 3-D multicellular systems. *PLoS Comput Biol*. 2018;14(2):e1005991.
22. Hoehme S, Drasdo D. Acell-based simulation software for multi-cellular systems. *Bioinformatics*. 2010;26(20):2641-2642.
23. Mirams GR, Arthurs CJ, Bernabeu MO, et al. Chaste: an open source C++ library for computational physiology and biology. *PLoS Comput Biol*. 2013;9(3):e1002970.
24. Kang S, Kahan S, McDermott J, Flann N, Shmulevich I. Biocellion: accelerating computer simulation of multicellular biological system models. *Bioinformatics*. 2014;30(21):3101-3108.
25. Cytowski M, Szymanska Z. Large-scale parallel simulations of 3d cell colony dynamics. *Computing in Science & Engineering*. 2014;16(5):86-95.
26. Pesonen H, Frazier D, Frigessi A, et al. ABC for the future.
27. Köhn-Luque A, Lai X, Frigessi A. Towards personalized computer simulations of breast cancer treatment. *arXiv*. 2020;q-bio.TO (2007.15934).

**How to cite this article:** Lai X, Taskén HA, Mo T, et al. A scalable solver for a stochastic, hybrid cellular automaton model of personalized breast cancer therapy. *Int J Numer Meth Biomed Engng*. 2022;38(1):e3542. doi:10.1002/cnm.3542



 **NTNU**

Norwegian University of  
Science and Technology