Targeting Cellular DNA Damage Responses: Predicting *in vivo* treatment responses using an *in vitro*-calibrated agent-based mathematical model

Sara Hamis[1,2,3], James Yates[4], Mark AJ Chaplain[1], Gibin G Powathil[2,3]

[1] School of Mathematics and Statistics, University of St Andrews, St Andrews KY16 9SS, Scotland.
[2] Department of Mathematics, College of Science, Swansea University, Swansea, SA2 8PP, United Kingdom.
[3] Computational Foundry, Swansea University, Swansea, SA2 8PP, United Kingdom.
[4] Oncology R&D, AstraZeneca, Cambridge, United Kingdom.

Abstract

The ATR (ataxia telangiectasia mutated and rad3-related kinase) inhibitor AZD6738 is an anti-cancer drug that potentially hinders tumour proliferation by targeting cellular DNA damage responses. In this study, we combine a systems pharmacology approach with an agent-based modelling approach to simulate AZD6738 treatment responses *in silico*. The mathematical model is governed by a set of empirically observable rules. By adjusting only the rules, whilst keeping the fundamental mathematical framework and model parameters intact, the mathematical model can first be calibrated by *in vitro* data and thereafter be used to successfully predict treatment responses in human tumour xenografts *in vivo* qualitatively, and quantitatively up to approximately 10 days post tumour injection.

Keywords

Cancer, Mathematical Oncology, DNA Damage Response Inhibition, ATR Inhibitor, AZD6738.

1

1   Introduction

The deoxyribonucleic acid (DNA) in human cells is perpetually exposed to influences that are potentially harmful. These influences can be derived from both exogenous and endogenous sources and events [1, 2]. Exogenous sources include ultraviolet radiation, ionising radiation (IR) and chemotherapeutic drugs [1], whilst erroneous DNA replication is an example an endogenous event yielding DNA damage [1]. Regardless of the source, a multitude of intracellular events are triggered when the DNA in a cell is damaged. For example, cells may respond to DNA damage by activating DNA repair mechanisms, cell cycle arrest or, in cases of severe DNA damage, apoptosis [3]. Cellular responses to DNA damage are mainly governed by the DNA damage response (DDR) pathway, which comprises a complex network of signaling pathways [3]. The DDR pathway has many functionalities. Amongst other things, the DDR pathway monitors DNA integrity and repairs DNA damage in order to maintain genomic stability in cells. The DDR pathway also governs DNA replication and cell cycle progression, and it controls apoptosis [1, 4]. From this, it is clear that the DDR pathway is crucial for maintaining cell viability. When DNA repair is needed, the DDR activates effector proteins [1]. Included in the group of effector proteins are approximately 450 proteins associated with the DDR [4], out of which the two main regulators for cell cycle checkpoints are ataxia telangiectasia mutated kinase (ATM) and ataxia telangiectasia mutated and rad3-related kinase (ATR) [2]. ATM and ATR are two proteins belonging to the enzyme family phosphatidyilinositol-3-OH-kinases (PI3K), and they both play central roles when cells respond to DNA damage [3].

The two principle types of DNA lesions are double-strand DNA breaks and single-strand DNA breaks [1]. Double-strand breaks are typically induced by IR [2], and ATM is the central protein involved in repairing double-strand breaks [1]. Single-strand breaks are a common result of replication stress [5] and the repair of single-strand DNA breaks is mainly attributed to ATR activity. ATR is active in the checkpoint in the intra-S phase of the cell cycle, both under undamaged circumstances and in response to DNA damage [3]. Although cross-talk between ATM and ATR regulators has been observed [2, 3], in this study we make the approximation that ATM repairs (typically IR-induced) double strand breaks and ATR repairs (here replication stress-induced) single strand breaks. Cancer cells are correlated with high replication stress and are therefore likely to rely on ATR for survival [1]. Moreover, cells that have a defected DDR pathway, with inhibited ATM or ATR, need to rely on other repair pathways for survival [1]. Due to the fact that ATM and ATR play such central roles in cellular repair post DNA-damage, both ATM and ATR inhibitors have during the last decade been designed and explored as mono and combination anti-cancer therapies [3]. With the premise that inhibiting DNA damage responses should increase the effect of some other main therapy, DDR inhibitors have been explored as both radiotherapy and chemotherapy treatment intensifiers [3]. The focus of this paper is ATR inhibitors, however it should be noted that pre-clinical experiments have validated the clinical potential of ATM inhibitors as ATM inhibitors have been investigated as radiosensitisers in pre-clinical studies, and early phase clinical testing is currently on-going [1]. Certain types of cancers are particularly susceptible to treatment plans involving ATR inhibitors [1]. Cancers with low ATM activity and high replication stress naturally rely a lot on ATR, and are thus rendered especially fragile and susceptible to anti-cancer targeting should their ATR be inhibited [4].

Two well-studied ATR inhibitors are namely AZD6738 and VX-970. AZD6738 is an oral ATR inhibitor, and its anti-tumour potential has been demonstrated in preclinical *vitro* and *in vivo* xenograft studies for various ATM deficient cell lines [2, 6]. Combination treatments with AZD6738 and radiotherapy or chemotherapy have produced synergistic results in preclinical settings [2], and AZD6738 is currently being evaluated in clinical phase I/II trials [1]. The intravenous ATR inhibitor VX-970 has demonstrated tumour controlling effects in a phase I clinical trial, both as a monotherapy and in combination with the chemotherapy drug carboplatin [1]. One cancer type that commonly features disruptions in the DDR is triple negative breast cancer (TNBC) [5]. Tu *et al.* [5] demonstrated that VX-970 can be used as a radiosensetiser and treatment intensifier to treat TNBC in patient-derived xenografts (PDX).

In a previous mathematical study by Checkley *et al.* [7], tumour responses to the ATR inhibitor AZD6738 are simulated using coupled ordinary differential equations, where both AZD6738 monotherapies and combination treatments with IR are regarded. In their study, a pharmacokinetic and pharmacodynamic (PK/PD) model of tumour growth is integrated with a mechanistic cell cycle model. Their model, which is calibrated by *in vitro* experiments, is predictive of *in vivo* xenograft studies. The model is aiding in quantitatively predicting dose and scheduling responses in a clinical Phase I trial design. The *in vitro* and *in vivo* data used in our current study is gathered from this previous work by Checkley *et al.* [7]. The mathematical framework used in our study is an extension of a mathematical model introduced by Powathil *et al.* [8] that has previously been used to study tumour growth, chemotherapy responses, radiotherapy responses, drug resistance and more [8, 9, 10, 11, 12].

Mathematical models, and their corresponding *in silico* experiments, can be used to simulate both *in vitro* and *in vivo* tumour scenarios. However, the microenvironment in an *in vitro* cell culture is significantly different from the microenvironment in a solid tumour, and many details that influence tumour dynamics differ between *in vitro* and *in vivo* settings. These details include cell proliferation, oxygen distribution and drug delivery. It follows that translating quantitative *in vitro* findings to *in vivo* predictions is not straightforward. In this study we use a multiscale, hybrid, agent-based mathematical model, in which one agent corresponds to one cancer cell (*in vitro*) or one group of cancer cells (*in vivo*). The model is governed by a few observable and well documented principles, or rules. To account for differences between the *in vitro* and *in vivo* scenarios, these rules are appropriately adjusted when moving from *in vitro* simulations to *in vivo* simulations. By only adjusting the rules, whilst keeping the fundamental mathematical framework intact, the model can first be calibrated by *in vitro* data and thereafter be used to successfully predict treatment responses in human tumour xenografts *in vivo*. Since *in vitro* data generally is easier to produce than *in vivo* data, this paper describes a modelling approach that facilitates moving towards *in silico* informed drug development and, ultimately, clinical trials.

2   Model

We take a minimal parameter, agent-based modelling (ABM) approach, and use a cellular automaton (CA) to model a population of cancer cells (*in vitro*), or a tumour (*in vivo*), that evolves in time and space. In order to account for differences between *in vitro* and *in vivo* scenarios, model rules are appropriately adjusted, as described throughout Section 2. Differences between the *in vitro* and the *in vivo* model are summarised in Figure 2 in Section 2.8. In the model, one *agent* corresponds to one cancer cell (*in vitro*) or one group of cancer cells (*in vivo*) that is distinct from other agents in the system. The behaviour and fate of each agent is governed by both intracellular and environmental dynamic variables, that are integrated using multiscale and hybrid modelling techniques. At the start of an *in silico* experiment, one initial agent is placed in the centre of the lattice. This initial agent divides and duplicates to gives rise to a heterogeneous population of agents. When the population has reached an appropriate size (determined by the experiments performed by Checkley at al. [7]), AZD6738 anticancer treatments commence.

2.1   Cellular Automaton Lattice

Every lattice point in the CA is either empty or occupied by one agent. If a lattice point is empty, it consists of extracellular solution (*in vitro*) or extracellular matrix (ECM) *in vivo*). The ECM comprises multiple components such as collagen, elastin and fibronectin but we do not distinguish between these components in the mathematical model [13]. Differences between the simulated *in vitro* an *in vivo* lattices are described bellow.

In vitro **lattice:** In the *in vitro* experiments performed by Checkley *et al.* [7], populations of LoVo (human colon carcinoma) cells were plated, and population sizes of up to roughly 4000 cells were investigated. In our corresponding mathematical *in vitro* model, we regard cells on a two-dimensional square lattice with $L \times L = 100 \times 100$ lattice points, where the spacing in both spatial directions, x_1 and x_2, corresponds to one cell diameter so that $\Delta x_1 = \Delta x_2 = 20\mu$m [8]. Thus the lattice simulates a physical space of 2×2 mm^2 [8].

In vivo **lattice:** In the *in vivo* experiments performed by Checkley *et al.* [7], LoVo cells were subcutaneously injected on flanks on Female Swiss nude mice in order to produce human tumour xenografts. Treatments started when the tumours had reached a volume of 0.2-0.3 cm^3 [7]. Approximating the tumour as spherical, we simulate (only) a central cross section of the tumour as an, approximately circular, plane of cells living on a two-dimensional square lattice. This lattice is specifically an $\tilde{L} \times \tilde{L} = 500 \times 500$ square lattice, with a spacing in both spatial directions $\tilde{x}_1$ and $\tilde{x}_2$ equal to 40μm, where the tilde-notations over the variables denote the *in vivo* setting. The dimensions are chisen in order to allow our agent-based model to simulate these physical dimensions, whilst keeping computational costs low, we let one agent correspond to one group of cancer cells. Post simulation time, the two-dimensional cross section of cells is extrapolated to represent a three-dimensional tumour-spheroid. This extrapolation process is outlined in the Supplementary Material.
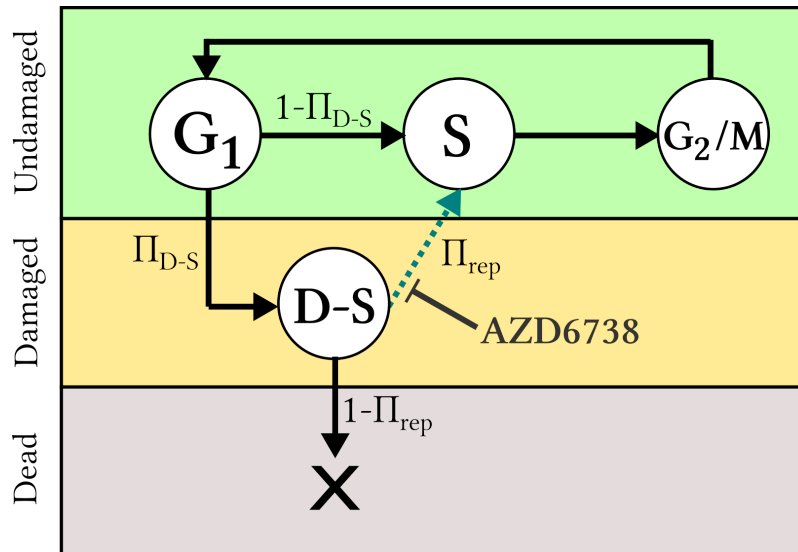
4

Figure 1: cell cycle model: An agent, i.e a cell (*in vitro*) or a group of cells (*in vivo*), progresses through various states of the cell cycle, where the states correspond to cell cycle phases. Viable (undamaged or damaged) states are shown in circles, whilst the dead state is shown as a cross. Paths illustrate transitions between states, and symbols next to the paths denote the probability that the respective path will be taken. The teal, dashed path indicates a replication path that can be inhibited by ATR-inhibitors, such as AZD6738, as is further described in Section 2.6.

2.2 Cell Cycle Model

In order to capture the influence of ATR and the ATR-inhibitor AZD6738 on the cell cycle, we use a probabilistic, rule-based cell cycle model adapted from Checkley *et al.* [7]. In this model, a cell progresses through various states in the cell cycle, where the states correspond to different cell cycle phases. As is illustrated in Figure 1, a cell can be in an undamaged state (G1, S or G2/M), a replication stress-induced DNA damaged state (D-S) or a dead state. The cause of cell death is here unrepaired replication stress. A cell can take different possible paths through the cell cycle, and every time the cell cycle path forks, stochastic 'dice-rolls' determine which path will be taken. Every cell commences its life in the G1 state of the cell cycle, but thereafter a cell can enter either the S state or the damaged S (D-S) state. The probability that a cell enters the D-S state is denoted Π_{D-S} and is fitted from *in vitro* data reported by Checkley *et al.* [7]. If a cell enters the D-S state, it has a chance to repair itself and enter the S state. If there is no drug in the system, this repair is always achieved, however the repair path is inhibited by the presence of AZD6738. The higher the drug-concentration is, the more unlikely it is that a cell in the D-S state will successfully repair itself to the S state. If a cell in the D-S state can not repair, it is sentenced to die. Whether a cell in state D-S repairs or dies is decided by a stochastic 'dice-roll' influenced by the local drug concentration $C(\bar{x}, t)$, as is further described in Section 2.6. A cell that has successfully reached the S state continues to the G2/M state after which it duplicates and starts over in the G1 state again, ready to perform another cycle.

Each agent i on the lattice is assigned an individual doubling time τ_i, where τ_i is stochastically selected from a normal distribution with mean value μ and standard deviation σ. A normal distribution is chosen in order to model the *in vitro* cells as being *almost* in sync, this is motivated by observations of the *in vitro* data, listed in the Supplementary Material, which implies fairly synchronised cell cycles amongst cells. Each agent is also attributed an individual cell cycle clock, that determines when the agent should progress to a subsequent state in the cell cycle. Every agent that is placed on the lattice commences its life in the G1 state, and progression to a subsequent state (i.e. S or D-S) is scheduled to occur once an agent has spent a certain fraction of its doubling time in the G1 state. The fraction of the doubling time spent in the G1, S (including D-S), G2/M states are respectively denoted Θ_{G1}, Θ_S and $\Theta_{G2/M}$, where these values are approximate, and chosen from literature to match values for typical human cells with a rapid doubling time of 24 hours so that $\Theta_{G1} = 11/24$, $\Theta_S = 8/24$ and $\Theta_{G2/M} = 5/24$ [14]. The fraction of an agent's doubling-time spent in the D-S state, Θ_{D-S}, is on the other hand fitted by *in vitro* data produced by Checkley *et al.* [7]. In order to account for differences between *in vitro* and *in vivo* systems, cell cycle model rules are adjusted as described below.

In vitro **cell cycle model rules:** Each cell has an individual doubling time, and individually (stochastically) determines which paths to take between the various states in the cell cycle model.

In vivo **cell cycle model rules:** Each agent comprises a group of identical cells that simultaneously and uniformly progress through the states in the cell cycle model.

2.3 Cell Proliferation

When an agent has completed the mitoses state in the cell cycle model, a secondary agent, namely a daughter agent, is produced. Each daughter agent is placed on a random lattice point in the (approximately circular) neighbourhood of the mother agent, where up to ν order of neighbourhoods are regarded. After all lattice points in neighbourhood n are occupied, daughter cells may be placed in neighbourhood $n + 1$. To accomplish circular-like growth, the model stochastically alternates between placing daughter agents on Moore and von Neumann neighbourhoods of the mother agent. Modelling rules for the *in vitro* and *in vivo* scenarios are outlined below.

In vitro **proliferation rules:** In the experimental *in vitro* setup, there is no spatial constraint or lack of nutrient deficiency inhibiting cell-division within the time-course of the experiment. Consequently cells are allowed to divide freely in the *in vitro* model and we set $\nu_{in\,vitro} = L/2$, where L corresponds to the size of the lattice, in order to restrict daughter cells to the lattice.

In vivo **proliferation rules:** *In vivo* tumours typically consist of a core with quiescent cells and a shell of proliferating cells. To accommodate for this, a daughter agent (representing a group of daughter cells) is allowed to be placed on up to a third order neighbourhood of the 'mother agent', so that $\tilde{\nu} = 3$, in accordance with previous mathematical models [8]. For the *in vivo* experiment regarded our current study, $\tilde{\nu} = 3$ matches the experimental data. However, for other experiments the value of $\tilde{\nu}$ may be adjusted to fit the specific cell-line and modelling

scenario at hand. When an agent is in the G1 phase of the cell cycle, it scans its environment to see if it has enough resources, in terms of space and nutrients, to commence the process of producing a daughter cell. If not, the cell enters the quiescent phase [15]. Thus in the model, when an agent is in the G1 phase, it continues to progress through the cell cycle model, provided that free space is available on the lattice within in its $\tilde{\nu}$th neighbourhood. If this is not the case, the agent exits the cell cycle to enter a quiescent state G0. Should neighbourhood space be made available again, as a result of anticancer targeting, quiescent agents may re-enter the cell cycle.

2.4 Oxygen Distribution and Influence on Cells

Tumour growth and treatment responses are highly influenced by intratumoural oxygen levels. [16, 17, 18]. Hypoxic (cancer) cells proliferate slower than do well-oxygenated cells [15] and, hypoxic tumour regions express reduced sensitivity to radiotherapy and a plethora of chemotherapeutic drugs [19]. In our model, the hypoxic cells (i.e. cells with a partial pressure of oxygen (pO_2) value of 10 mmHg or less [20]) display arrest (i.e. delay) in the G1 phase of the cell cycle [21]. As is described bellow, details regarding oxygen dynamics are included in the *in vivo* model, but not in the *in vitro* model.

In vitro **oxygen distribution and responses:** In the mathematical *in vitro* model, all cells are assumed to be well-oxygenated in accordance with the experimental *in vitro* setup performed by Checkley *et al.* [7]. Consequently, neither oxygen dynamics nor cellular responses to low oxygen levels are incorporated in the mathematical *in vitro* model.

In vivo **oxygen distribution and responses:** Within solid tumours, oxygen concentrations typically vary and hypoxic regions are common tumour features [21, 18, 22]. Avoiding complicated details of vasculature in the model, we here approximate the oxygen as diffusing in from 'outside the tumour'. Oxygen dynamics across the CA lattice is here described using a mechanistic diffusion equation, where the oxygen concentration in location $\bar{x}$ at time t is denoted by $K(\bar{x},t)$ where

$$\frac{\partial K(\bar{x},t)}{\partial t} = \nabla \cdot (D_K(\bar{x},t)\nabla K(\bar{x},t)) + r_K m(\bar{x},t) - \phi_K K(\bar{x},t)cell(\bar{x},t). \tag{1}$$

$D_K(\bar{x},t)$ denotes the oxygen diffusion coefficient, and r_K and ϕ_K are source and consumption coefficients respectively. The diffusion coefficient for oxygen is known from literature to be 2.5×10^{-5} cm^{2}s^{-1} [8]. Assuming that oxygen diffuses slower over cells than in the ECM, the oxygen diffusion coefficient is divided by a factor 1.5 if there is a cell in location $\bar{x}$ at time t. The binary factor $m(\bar{x},t)$ is 1 if the regarded location $\bar{x}$ is outside the tumour at time t and 0 otherwise. Similarly, the binary factor $cell(\bar{x},t)$ is 1 if there is a viable cell in location $\bar{x}$ at time t, and 0 otherwise [8]. Equation 1 is coupled with no-flux boundary conditions so that the total amount of oxygen in the system fluctuates over time [23]. A scaled oxygen variable $\hat{K}(\bar{x},t)$ is introduced in order to express oxygenation levels in units of mmHg. $\hat{K}(\bar{x},t)$ is obtained by

$$\hat{K}(\bar{x},t) = \frac{K(\bar{x},t)}{max_{\bar{x},t}K(\bar{x},t)} \cdot h, \tag{2}$$

7

where $max_{\bar{x},t}K(\bar{x},t)$ denotes the maxima occurring $K(\bar{x},t)$-value at time t [23] and h is a scaling factor. The scaling factor is included in order to achieve a lattice on which lattice points have oxygenation levels ranging between 0 and 100 mmHg.

Low cellular oxygen levels have been shown to delay cell cycle progression by inducing arrest in, particularly, the G1 phase of the cell cycle [15]. In mechanistic Tyson-Novak type cell cycle models [24, 25, 26], the cell cycle is governed by a system of ordinary differential equations (ODEs) in which the G1 phase is inherently elongated under hypoxic conditions by incorporating hypoxia-induced factors into the ODEs [8]. In this model, we use a clock to model cell cycle progression and thus we introduce a G1 delay factor (G1DF) in order to achieve a longer G1-phase under hypoxic conditions where

$$G1DF(\hat{K}(x,t)) = \begin{cases} a_1 + \frac{a_2}{a_3+\hat{K}(\bar{x},t)} & \text{if } \hat{K}(x,t) \leq 10.5 \text{ mmHg }, \\ 1 & \text{otherwise}, \end{cases} \tag{3}$$

The G1DF is an approximation for how much the G1 phase is expanded in time as a function of oxygen pressure. It is matched to fit data points extracted from a previous mathematical study by Alarcon *et al.* [15], in which a Tyson-Novak cell cycle model is extended to incorporate the action of p27, a protein that is up-regulated under hypoxia and delays cell cycle progression. Data-fitting yields the parameter values $a_1 = 0.9209, a_2 = 0.8200, a_3 = -0.2389$ [21]. Thus the fraction of an agent's doubling time spent in the G1 state is $G1DF(\hat{K}(\bar{x},t)) \cdot \Theta_{G1}$, where $G1DF(\hat{K}(\bar{x},t)) = 1$ for normoxic cells.

2.5 Drug Distribution Across the Lattice

Drug distribution significantly varies between *in vitro* and *in vivo* settings, as described below. In the regarded *in vitro* setup, the drug concentration can be regarded as homogeneous, whilst heterogeneous drug concentrations must be accounted for *in vivo*.

In vitro **drug distribution:** In the *in vitro* experiments performed by Checkley *et al.* [7], plated cell populations of roughly 1000 cells were treated with AZD6738 in the solvent dimethyl-sulfoxide (DMSO). In the mathematical model, we thus approximate the drug distribution across the lattice to be instantaneous (occurring at treatment time T_0) and homogeneous. We furthermore assume that the drug has a half-life time that exceeds the time course of the experiment, and note that there is no other drug elimination from the *in vitro* system (unless drug wash-out is applied). In our mathematical model, this is equivalent to there being no drug decay. Hence in the mathematical *in vitro* model, the drug concentration $C(\bar{x},t)$, in location $\bar{x}$ at time t is simply given by

$$C(\bar{x},t) = \begin{cases} 0 \text{ everywhere} & \text{if } t < T_0, \\ \mathbb{C} \text{ everywhere} & \text{if } t \geq T_0, \end{cases} \tag{4}$$

where $\mathbb{C}$ denotes the applied drug concentration (in units of molarity).

8

In vivo **drug distribution:** In the *in vivo* experiments performed by Checkley *et al.* [7], AZD6738, or vehicle in the control case, were administered via oral gavage once per day to Female Swiss nude mice. In the mathematical *in vivo* model, we consider the drug to diffuse through the tumour from its surrounding, creating a smooth drug gradient within the tumour. In the mathematical model, this drug dynamics is modelled using a mechanistic partial differential equation (PDE), where the concentration of the drug AZD6738 in lattice point n at location $\bar{x}$ at time t is denoted by $C(\bar{x}, t)$ and

$$\frac{\partial C(\bar{x},t)}{\partial t} = \nabla \cdot \big(D_{AZD}(\bar{x},t)\nabla C(\bar{x},t) \big) + p(\bar{x},t) - \eta_{AZD}C(\bar{x},t), \tag{5}$$

where D_{AZD} is the diffusion coefficient of the drug AZD6738, and the production coefficient $p(\bar{x}, t)$ is greater than zero at drug administration times only for lattice points outside the tumour, i.e. in the extracellular matrix. The value of $p(\bar{x}, t)$ corresponds to the amount of administered drug. Assuming first order kinetics for drug elimination, the drug decay constant η_{AZD} is matched to the reported half-life time of 6 hours for AZD6738 *in vivo* [27]. Note that the drug decay term here represents all drug elimination from the system, both metabolic and that caused by excretion.

The diffusion rate of a drug is predominately affected by the molecular size of the drug. More specifically, the diffusion coefficient of a drug is inversely proportional to the square root of the molecular weight of the drug, so that large molecules diffuse more slowly than do small molecules [28]. Using this assumption, the drug diffusion coefficient is set in relation to the oxygen diffusion coefficient, as is done in previous mathematical studies [8]. Thus the relationship between the diffusion coefficients for the drug (AZD) and oxygen (O_2) corresponds to the square of the inverse relationship between the corresponding molecular weights, such that

$$\frac{D_{AZD}}{D_{0_2}} = \frac{\sqrt{\text{molecular weight}(O_2)}}{\sqrt{\text{molecular weight}(AZD)}} = \sqrt{\frac{31.998\text{g/mol}}{412.512\text{g/mol}}} \approx 0.27851, \tag{6}$$

where the molecular weights are collected from the PubChem database [29]. Details regarding pharmacokinetics are outside the scope of this study, bioavailability is instead calibrated using the extreme scenario in which the maximum drug dose is administered *in vivo*.

2.6 Drug Responses

AZD6738 inhibits the repair from the D-S state to the S state in the cell cycle, as illustrated in Figure 1, and maximal drug effect corresponds to complete repair inhibition. The drug effect is modelled using an agent-based adaptation of the sigmoid Emax model [30], in which the drug effect on a cell in position $\bar{x}$ at time t is given by

$$E(\bar{x},t) = E_{max}\frac{C(\bar{x},t)^{\gamma}}{EC_{50}^{\gamma} + C(\bar{x},t)^{\gamma}}. \tag{7}$$

E_{max} denotes the maximal drug effect, here corresponding to complete repair inhibition ($E_{max} = 1$), EC_{50} denotes the concentration required to achieve half of the maximal drug effect,

9

Section	Parameter	Calibrated Value
2.2	μ, σ	24 h, 0.5 h
	Π_{D-S}, θ_{D-S}	0.75, 0.03
2.6	EC_{50}, γ	1 μM, 2
	$T_{L \to D}$	τ_i

Table 1: *In vitro* calibrated parameters.

$0.5 \cdot E_{max}$ and γ is the Hill-exponent [30]. EC_{50} and γ are fitted from the *in vitro* data. When an agent is scheduled to progress from the D-S state in the cell cycle, it has a probability Π_{rep} to repair which is determined by the local drug concentration such that

$$\Pi_{rep}(\bar{x}, t) = 1 - E(\bar{x}, t). \tag{8}$$

Note that in the absence of drugs, the repair probability is 1. When a cell dies, it is transformed into a membrane-enclosed 'cell-corpse' [28]. In the *in vivo* setting, the cellular debris is digested by macrophages but in the *in vitro* setting such 'cell-corpses' linger on the lattice. Post the lethal event (i.e. the D-S to S repair failure) a cell is declared 'dead' in the model after a time $T_{L \to D}$ has passed (where L stands for 'lethal event' and D stands for 'death'). The parameter $T_{L \to D}$ is calibrated by *in vitro* experiments. The differences between modelling rules for *in vitro* and *in vivo* drug responses are described below.

In vitro **drug responses:** The survival probability of a cell is modelled using the Emax model. After failure to repair from the D-S state, a cell is considered to be dead after a time $T_{L \to D}$ has passed. However, a dead cell is never physically removed from the lattice.

In vivo **drug responses:** An agent's survival probability is modelled using the Emax model. An agent is declared to be dead and removed from the lattice after an amount of time $T_{L \to D}$ post the lethal event (failure to repair).

2.7   Parameters

sec:ddrparams The parameters used in the mathematical framework are calibrated by *in vitro* data and are listed in Table 1. The *in vitro* and *in vivo* data produced by Checkley *et al.* [7] is listed in the Supplementary Material. The Supplementary Material also includes information regarding the model calibration process. In the context of quantitative pharmacology, knowledge about a model's robustness is crucial [31]. Accordingly, results from the uncertainty and sensitivity analysis are also included in the Supplementary Material. Three different uncertainty and sensitivity analyses techniques, suitable for agent-based models with stochastic elements, are used as described in the Supplementary Material. These three techniques are namely (*i*) Consistency Analysis, (*ii*) Robustness Analysis and (*iii*) Latin Hypercube Analysis [32, 33].

2.8   Differences Between *in vitro* and *in vivo* Modelling Rules

The differences between the *in vitro* and *in vivo* rules used in the mathematical framework are pictorially summarised in Figure 2.
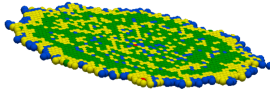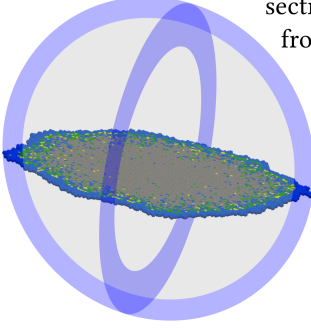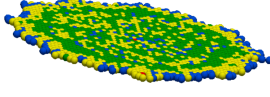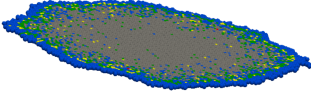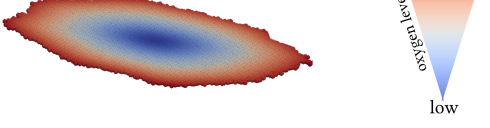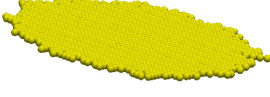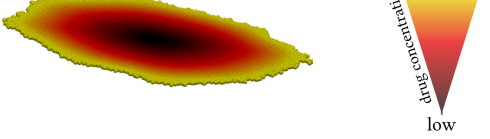
	in vitro rules	*in vivo rules*
Lattice setup	A monolayer of cells is simulated. One agent represents one cancer cell.	A center tumour cross-section is simulated, from which a tumour spheroid is extrapolated post simulation time. One agent represents a group of cancer cells.
Cell cycle heterogeneity	All cells are cycling. 🔵G1 🟡S 🔴D-S 🟢G2/M	Agents that are surrounded by other agents in ν or more orders of their circular neighbourhoods are quiescent. 🔵G1 🟡S 🔴D-S 🟢G2/M ⚪G0
Oxygen dynamics	All cells are well oxygenated, oxygenation levels are constant in time and space on the lattice. 🔴Oxygenation level	Oxygen diffuses in from outside the tumour. Hypoxia induces G1 arrest. high / oxygen level / low
Drug dynamics	🟡Drug concentration Post drug administration time, drug concentration is constant in time and space on the lattice.	high / drug concentration / low Drugs diffuse in from outside the tumour, and drugs are eliminated from the system.
Cell removal	Dead agents are *not* removed from the lattice.	Dead agents are removed from the lattice after a time $T_{L->D}$.

Figure 2: A summary of the differences between the *in vitro* and *in vivo* rules used in the mathematical framework.

2.9 Implementation

The mathematical model is implemented in an in-house C/C++ framework in order to produce *in silico* results. Partial differential equations are solved using explicit finite difference methods. Simulations are visualised using ParaView [34]. Results reported in Section 3 are based on 100 simulation runs as, according to results from the uncertainty analysis, 100 runs are sufficient for mitigating uncertainty originating from intrinsic model stochasticity. Uncertainty and sensitivity analyses are performed using MATLAB [35].

3 Results

The mathematical framework is first calibrated by *in vitro* data, and is thereafter used to predict treatment responses in human tumour xenografts *in vivo*. By only adjusting the modelling *rules* in the mathematical framework, whilst keeping the *in vitro* calibrated model parameters intact, the mathematical framework is able successfully predict treatment responses in human tumour xenografts *in vivo*.

3.1 Simulating *in vitro* Experiments

The mathematical framework used in this study is calibrated by results from an *in vitro* experiment performed by Checkley *et al.* [7] in which populations of LoVo (human coloncarcinoma) cells are exposed to the ATR-inhibiting drug AZD6738. The *in silico* results in Figure 3 simulate the evolution of the *in vitro* cell population over time in terms of cell damage (Figure 3 Left) and in terms of cell count (Figure 3 Right). AZD6738 drugs are given at 0 hours, when the cell population has reached a size of approximately 1000 cells. Response curves for six different drug concentrations, including the zero-drug concentration control case, are shown. Each response curve is based on a mean values from 100 simulation runs, according to the uncertainty analysis described in the supplementary material. Also shown in Figure 3 are simulation standard deviations and *in vitro* results produced by Checkley et al. [7]. Using a minimal-parameter modelling approach, the mathematical framework is calibrated to best fit all *in vitro* data points without introducing any auxiliary scaling variables or similar. The calibration process is described in the Supplementary Material. Our results demonstrate that, after calibration, our mathematical framework is able to capture *in vitro* LoVo cell population growth and drug (AZD6738) responses.

3.2 Simulating *in vivo* Experiments

Post the *in vitro* calibration, our mathematical framework is used to simulate *in vivo* experiments performed by Checkley *et al.* [7] in which LoVo xenografts, that are injected in mice flanks, are treated with AZD6738 once daily for 14 days. The results in Figure 4 simulate AZD6738 drug responses in terms of the percentage of damaged cells (Figure 4 Left) and tumour volume (Figure 4 Right). Response curves to three different drug doses (0, 25 and 50 mg/kg) are shown, where the curves represent mean values for 100 simulation runs. Simulation standard deviations and *in vivo* data are also provided in Figure 4. These result graphs demonstrate that our mathematical framework quantitatively agrees with the *in vivo* results reported by Checkley et al. [7] for up

to approximately 10 days post tumour injection. This empirically demonstrates the predictive ability of our mathematical framework and modelling approach.
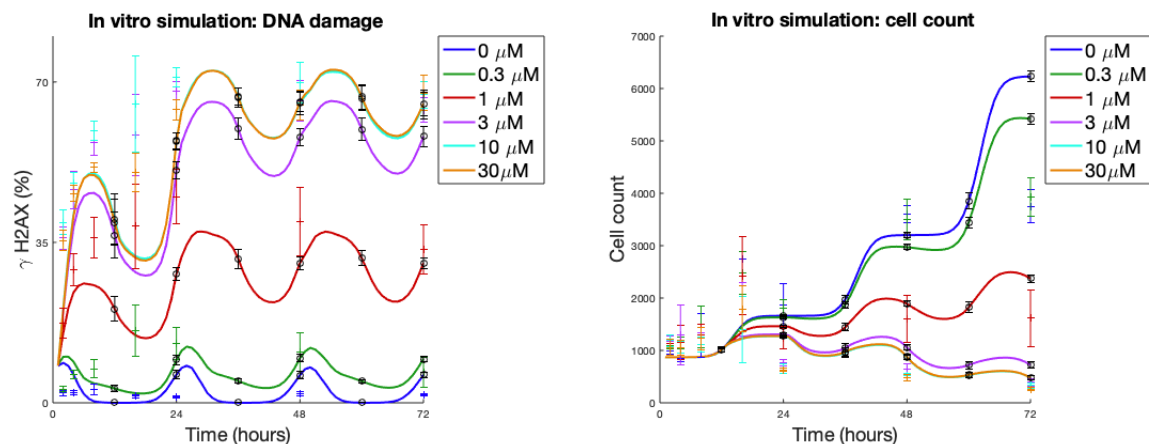

Figure 3: Simulated *in vitro* drug response curves. LoVo cells are exposed to drug (AZD6738) at 0 hours. Left: The percentage of γH2AX-positive (i.e. DNA-damaged) cells in the system over time. Right: Cell count over time. Simulation mean values for 100 *in silico* (coloured lines), simulation standard deviations (black error bars) and *in vitro* data with standard deviation (coloured error bars) [7] are shown.
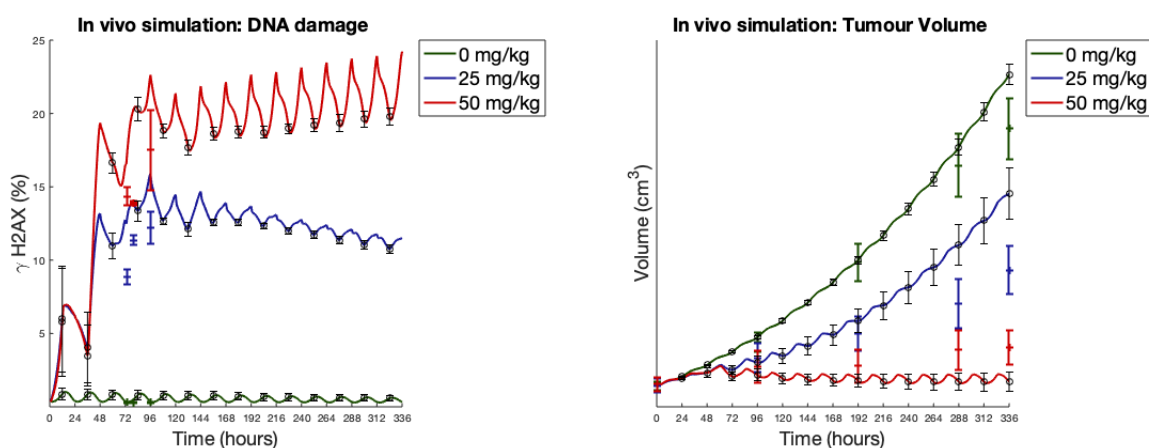

Figure 4: Simulated *in vivo* drug response curves. LoVo xenografts are exposed to drug (AZD6738) once daily (i.e. once every 24 hours) for 14 days. Left: The percentage of γH2AX-positive (i.e. DNA-damaged) cells in the xenograft over time. Right: Tumour volume over time. Simulation mean values for 100 simulation runs (coloured lines), simulation standard deviations (black error bars) and *in vivo* data with standard errors (coloured error bars) [7] are shown.

13

4   Discussion

We demonstrated that the mathematical framework used in this study can successfully simulate growth and drug (AZD6738) responses both in cancer cell populations (*in vitro*) and in tumour xenografts (*in vivo*). This versatile mathematical framework is an extension of a multiscale, hybrid, agent-based, on-lattice model that has previously been used to study tumour growth and treatment responses to chemotherapy, radiotherapy, hypoxia-activated prodrugs and more [8, 9, 10, 11, 12]. By only adjusting the rules in the mathematical framework, whilst keeping model parameters intact, the mathematical model can first be calibrated by *in vitro* data and thereafter be used to successfully predict treatment responses in human tumour xenografts *in vivo*. After comparing *in vivo*-scenario *in silico* results to *in vivo* data, one could chose to further fine-tune the modelling parameters in order to obtain an even better fit.

Data-driven modeling, exploitation of existing data and proof-of-concept studies are important steps involved in current and future procedure for enabling multiscale modeling in systems medicine, as argued in a report by Wolkenhauer *et al.* [36]. In the data-driven, minimal-parameter, 'proof-of-concept' study discussed in this article, we demonstrate that our mathematical framework is able to produce *in silico* results that match both *in vitro* and *in vivo* data produced in a previous study by Checkley *et al.* [7]. In their study, Checkley at al. [7] used a pharmacokinetic/pharmacodynamic (PK/PD) model to predict *in vitro* and *in vivo* AZD6738 treatment responses. The mathematical results produced by Checkley *et al.*'s [7] PK/PD model agree with the mathematical result produced by our multiscale, hybrid, agent-based, on-lattice model agrees. Thus our mathematical framework is validated by, not only *in vitro* and *in vivo* data, but also by another mathematical model that is based on modelling techniques different from ours. Despite the fact that mathematical modelling is, in general, is becoming increasing popular in the pharmaceutical industry, there are not that many agent-based models present on the pharmaceutical scene [37]. We argue this is a missed opportunity in the context of oncology, as agent-based models naturally capture the heterogeneous nature of tumours, that is known to complicate treatment matters. Moreover, multiscale models enable the integration of knowledge and data on different scales in time and space. By formulating agent-based modelling rules using 'low' principles (our ideological equivalent to 'first principles' in this context) regarding how the agents in the system behave, the produced *in silico* data can be handled as classical *in vitro* and *in vivo* data post simulation [37].

Moving drug-response investigations from *in vitro* to *in vivo* settings is a key step involved in the process of moving a drug from *bench-to-bedside*. Data from *in vivo* experiments is often sparse, as gathering *in vivo* data is associated with practical, financial and ethical constraints. Plentiful and adaptable *in silico* data are however easy to produce, and thus sparse *in vivo* data can be complemented and validated by *in silico* data. Thus mathematical frameworks, such as the one used in this study, can be used as an epistemic contribution to sparse experimental data. Since our *in vitro*-calibrated mathematical framework is able to predict *in vivo* responses, it can be used as a tool in the process of moving a drug-response investigation from an *in vitro* to an *in vivo* setting. Furthermore, the corresponding *in silico* experiments can be extended to investigate various dose-schedule scenarios in order to, not only validate, but also, guide *in vitro* and *in vivo*

experiments. Ultimately, following interdisciplinary collaborations between clinicians, biologists and mathematicians, mathematical oncology may be used to further personalised cancer care in clinical settings according to *bench-to-bedside* and *blackboard-to-bedside* approaches [7, 11, 36, 31]. As multiscale, agent-based models organically enable the integration of data across various scales in time and space, it follows that they are useful to the interdisciplinary team that wishes to combine the knowledge and data of its team members.

Funding

Competing Interests

No competing interests to declare.

Authors' Contributions

All authors conceived of the study and coordinated the study. All authors contributed to model design. SH performed the calibration, *in silico* experiments, sensitivity analysis and drafted the manuscript. All authors critically revised the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

References

[1] A. Minchom, C. Aversa, and J. Lopez, "Dancing with the DNA damage response: next-generation anti-cancer therapeutic strategies," *Ther Adv Med Oncol* **10** (2018) 1758835918786658.

[2] R. Sundar, J. Brown, A. Ingles Russo, and T. A. Yap, "Targeting ATR in cancer medicine," *Curr Probl Cancer* **41** no. 4, (2017) 302–315.

[3] L. Carrassa and G. Damia, "DNA damage response inhibitors: Mechanisms and potential applications in cancer therapy," *Cancer Treat. Rev.* **60** (Nov, 2017) 139–151.

[4] A. R. Nam, M. H. Jin, J. E. Park, J. H. Bang, D. Y. Oh, and Y. J. Bang, "Therapeutic Targeting of the DNA Damage Response Using an ATR Inhibitor in Biliary Tract Cancer," *Cancer Res Treat* (Dec, 2018) .

[5] X. Tu, M. M. Kahila, Q. Zhou, J. Yu, K. R. Kalari, L. Wang, W. S. Harmsen, J. Yuan, J. C. Boughey, M. P. Goetz, J. N. Sarkaria, Z. Lou, and R. W. Mutter, "ATR Inhibition Is a Promising Radiosensitizing Strategy for Triple-Negative Breast Cancer," *Mol. Cancer Ther.* **17** no. 11, (Nov, 2018) 2462–2472.

[6] K. M. Foote, J. W. M. Nissink, T. McGuire, P. Turner, S. Guichard, J. W. T. Yates, A. Lau, K. Blades, D. Heathcote, R. Odedra, G. Wilkinson, Z. Wilson, C. M. Wood, and P. J. Jewsbury, "Discovery and Characterization of AZD6738, a Potent Inhibitor of Ataxia Telangiectasia Mutated and Rad3 Related (ATR) Kinase with Application as an Anticancer Agent," *J. Med. Chem.* **61** no. 22, (11, 2018) 9889–9907.

[7] S. Checkley, L. MacCallum, J. Yates, P. Jasper, H. Luo, J. Tolsma, and C. Bendtsen, "Bridging the gap between in vitro and in vivo: Dose and schedule predictions for the ATR inhibitor AZD6738," *Sci Rep* **5** (Aug, 2015) 13545.

[8] G. G. Powathil, K. E. Gordon, L. A. Hill, and M. A. Chaplain, "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.* **308** (Sep, 2012) 1–19.

[9] G. G. Powathil, M. Swat, and M. A. Chaplain, "Systems oncology: towards patient-specific treatment regimes informed by multiscale mathematical modelling," *Semin. Cancer Biol.* **30** (Feb, 2015) 13–20.

[10] S. Hamis, P. Nithiarasu, and G. G. Powathil, "What does not kill a tumour may make it stronger: In silico insights into chemotherapeutic drug resistance," *J. Theor. Biol.* **454** (Jun, 2018) 253–267.

[11] S. Hamis, G. G. Powathil, and M. A. J. Chaplain, "Blackboard to Bedside: A Mathematical Modeling Bottom-Up Approach Toward Personalized Cancer Treatments," *JCO Clin Cancer Inform* **3** (02, 2019) 1–11.

[12] S. Bruningk, G. Powathil, P. Ziegenhein, J. Ijaz, I. Rivens, S. Nill, M. Chaplain, U. Oelfke, and G. Ter Haar, "Combining radiation with hyperthermia: a multiscale model informed byin vitroexperiments," *J R Soc Interface* **15** no. 138, (Jan, 2018) .

[13] L. Preziosi and A. Tosin, "Multiphase modelling of tumour growth and extracellular matrix interaction: mathematical tools and applications," *J Math Biol* **58** no. 4-5, (Apr, 2009) 625–656.

[14] G. Cooper and R. Hausman, "The Cell-A Molecular Approach 4th edition,".

[15] T. Alarcon, H. M. Byrne, and P. K. Maini, "A mathematical model of the effects of hypoxia on the cell-cycle of normal and cancer cells," *J. Theor. Biol.* **229** no. 3, (Aug, 2004) 395–411.

[16] J. Hu, D. R. Handisides, E. Van Valckenborgh, H. De Raeve, E. Menu, I. Vande Broek, Q. Liu, J. D. Sun, B. Van Camp, C. P. Hart, and K. Vanderkerken, "Targeting the multiple myeloma hypoxic niche with TH-302, a hypoxia-activated prodrug," *Blood* **116** no. 9, (Sep, 2010) 1524–1527.

[17] V. Liapis, A. Labrinidis, I. Zinonos, S. Hay, V. Ponomarev, V. Panagopoulos, M. DeNichilo, W. Ingman, G. J. Atkins, D. M. Findlay, A. C. Zannettino, and A. Evdokiou, "Hypoxia-activated pro-drug TH-302 exhibits potent tumor suppressive activity and cooperates with chemotherapy against osteosarcoma," *Cancer Lett.* **357** no. 1, (Feb, 2015) 160–169.

[18] S. G. Peeters, C. M. Zegers, R. Biemans, N. G. Lieuwes, R. G. van Stiphout, A. Yaromina, J. D. Sun, C. P. Hart, A. D. Windhorst, W. van Elmpt, L. J. Dubois, and P. Lambin, "TH-302 in Combination with Radiotherapy Enhances the Therapeutic Outcome and Is Associated with Pretreatment [18F]HX4 Hypoxia PET Imaging," *Clin. Cancer Res.* **21** no. 13, (Jul, 2015) 2984–2992.

[19] R. M. Phillips, "Targeting the hypoxic fraction of tumours using hypoxia-activated prodrugs," *Cancer Chemother. Pharmacol.* **77** no. 3, (Mar, 2016) 441–457.

[20] C. T. Lee, M. K. Boss, and M. W. Dewhirst, "Imaging tumor hypoxia to advance radiation oncology," *Antioxid. Redox Signal.* **21** no. 2, (Jul, 2014) 313–337.

[21] S. Hamis, A. Yaromina, L. Dubois, M. Kohandel, P. Lambin, and G. G. Powathil, "Combining Hypoxia-Activated Prodrugs and Radiotherapy *in silico*: Impacts of Treatment Scheduling and the Intra-Tumoural Oxygen Landscape," *(To Appear)* .

[22] J. D. Sun, Q. Liu, J. Wang, D. Ahluwalia, D. Ferraro, Y. Wang, J. X. Duan, W. S. Ammons, J. G. Curd, M. D. Matteucci, and C. P. Hart, "Selective tumor hypoxia targeting by hypoxia-activated prodrug TH-302 inhibits tumor growth in preclinical models of cancer," *Clin. Cancer Res.* **18** no. 3, (Feb, 2012) 758–770.

[23] G. Powathil, M. Kohandel, M. Milosevic, and S. Sivaloganathan, "Modeling the spatial distribution of chronic tumor hypoxia: implications for experimental and clinical studies," *Comput Math Methods Med* **2012** (2012) 410602.

[24] J. J. Tyson and B. Novak, "Regulation of the eukaryotic cell cycle: molecular antagonism, hysteresis, and irreversible transitions," *J. Theor. Biol.* **210** no. 2, (May, 2001) 249–263.

[25] B. Novak and J. J. Tyson, "Modelling the controls of the eukaryotic cell cycle," *Biochem. Soc. Trans.* **31** no. Pt 6, (Dec, 2003) 1526–1529.

[26] B. Novak and J. J. Tyson, "A model for restriction point control of the mammalian cell cycle," *J. Theor. Biol.* **230** no. 4, (Oct, 2004) 563–579.

[27] F. P. Vendetti, A. Lau, S. Schamus, T. P. Conrads, M. J. O'Connor, and C. J. Bakkenist, "The orally active and bioavailable ATR kinase inhibitor AZD6738 potentiates the anti-tumor effects of cisplatin to resolve ATM-deficient non-small cell lung cancer in vivo," *Oncotarget* **6** no. 42, (Dec, 2015) 44289–44305.

[28] M. M. Dale and H. P. Rang, "Rang Dale's pharmacology ," *Edinburgh: Churchill Livingstone* (2007) .

[29] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, and E. E. Bolton, "PubChem 2019 update: improved access to chemical data," *Nucleic Acids Res.* **47** no. D1, (Jan, 2019) D1102–D1109.

[30] N. Holford, "Pharmacodynamic principles and the time course of immediate drug effects. ," *Transl Clin Pharmacol* **4** (2017) 157–161.

[31] S. A. Visser, D. P. de Alwis, T. Kerbusch, J. A. Stone, and S. R. Allerheiligen, "Implementation of quantitative and systems pharmacology in large pharma," *CPT Pharmacometrics Syst Pharmacol* **3** (Oct, 2014) e142.

[32] S. Hamis, S. Stratiev, and G. Powathil, "Uncertainty and sensitivity analyses methods for agent-based models: A review," *(To Appear)* .

[33] K. Alden, M. Read, J. Timmis, P. S. Andrews, H. Veiga-Fernandes, and M. Coles, "Spartan: a comprehensive tool for understanding uncertainty in simulations of biological systems," *PLoS Comput. Biol.* **9** no. 2, (2013) e1002916.

[34] Ayachit and Utkarsh, "The ParaView Guide: A Parallel Visualization Application, ISBN: 978-1930934306," *Kitware* (2015) .

[35] MATLAB, *version 1.8.0_202 (R2019n)*. The MathWorks Inc., Natick, Massachusetts, 2019.

[36] O. Wolkenhauer, C. Auffray, O. Brass, J. Clairambault, A. Deutsch, D. Drasdo, F. Gervasio, L. Preziosi, P. Maini, A. Marciniak-Czochra, C. Kossow, L. Kuepfer, K. Rateitschak, I. Ramis-Conde, B. Ribba, A. Schuppert, R. Smallwood, G. Stamatakos, F. Winter, and H. Byrne, "Enabling multiscale modeling in systems medicine," *Genome Med* **6** no. 3, (2014) 21.

[37] J. Cosgrove, J. Butler, K. Alden, M. Read, V. Kumar, L. Cucurull-Sanchez, J. Timmis, and M. Coles, "Agent-Based Modeling in Systems Pharmacology," *CPT Pharmacometrics Syst Pharmacol* **4** no. 11, (Nov, 2015) 615–629.

Targeting Cellular DNA Damage Responses: Predicting *in vivo* treatment responses using an *in vitro*-calibrated agent-based mathematical model

Sara Hamis[1,2,3], James Yates[4], Mark AJ Chaplain[1], Gibin G Powathil[2,3]

[1] School of Mathematics and Statistics, University of St Andrews, St Andrews KY16 9SS, Scotland.
[2] Department of Mathematics, College of Science, Swansea University, Swansea, SA2 8PP, United Kingdom.
[3] Computational Foundry, Swansea University, Swansea, SA2 8PP, United Kingdom.
[4] Oncology R&D, AstraZeneca, Cambridge, United Kingdom.

– Supplementary Material

1 Experimental Data

The experimental *in vitro* and *in vivo* data used in our current study are gratefully gathered from a previous study performed by Checkley *et al.* [1]. *In vitro* data are listed in Table 1 and *in vivo* data are listed in Tables 2 and 3.

Time (hours)	Cell count	Std.Dev (count)	γH2AX positive (%)	Std. Dev (%)
0μM (control)				
2	996	59.72	2.14	0.56
4	850	62.30	2.20	0.45
8	1287.5	417.59	2.90	1.22
16	2742.75	439.69	1.44	0.33
24	1857.5	409.39	1.29	0.32
48	3605.25	167.38	1.93	0.44
72	3753	311.17	1.71	0.21
0.3 μM				
2	1081.88	53.63	2.93	0.59
4	1040.75	217.96	6.15	1.00
8	1447.25	392.45	7.41	1.99
16	2479.5	414.02	15.68	5.56
24	1805.63	161.41	12.91	3.42
48	3497.63	385.19	11.08	4.18
72	3928.25	376.08	6.57	3.30
1 μM				
2	1129.63	58.26	17.35	3.31
4	1153.63	331.31	29.12	3.47
8	1303.88	199.72	36.05	4.35
16	2420.25	744.38	38.51	9.25
24	1226.38	185.58	45.01	6.01
48	1600.38	456.80	39.47	7.47
72	1612.88	540.55	33.47	5.46
3 μM				
2	1171.14	97.71	36.01	2.42
4	1291.38	567.63	46.47	4.09
8	1224.63	113.30	56.72	2.62
16	1784.38	513.06	58.41	8.81
24	765.75	70.76	68.07	2.05
48	638.75	112.54	65.90	4.40
72	392.63	67.64	63.82	2.67
10 μM				
2	1191.13	110.15	39.38	2.62
4	1056.63	106.72	47.98	2.32
8	1113.63	144.42	59.35	1.99
16	1396	633.86	65.21	10.48
24	654.5	100.26	71.02	2.10
48	525.29	43.93	69.75	4.42
72	326.63	47.73	67.25	2.87
30 μM				
2	1055.13	155.16	35.37	2.21
4	1049.13	147.96	45.66	1.75
8	1228.75	211.96	51.37	1.11
16	1794.88	435.42	50.35	4.19
24	629	27.12	63.92	2.15
48	469.63	61.26	64.92	3.25
72	265.13	22.26	67.63	3.96

Table 1: *In vitro* data gathered from a previous study by Checkley *et al.* [1].

Time (hours)	γH2AX positive (%)	Std. Error (%)
0 mg/kg (control)		
74	0.3200	0.0320
80	0.2770	0.0300
96	0.2970	0.0340
25 mg/kg QD		
74	8.8579	0.5364
80	11.3692	0.3272
96	12.1945	1.0949
50 mg/kg QD		
74	14.3417	0.6278
80	13.8967	0.1401
96	17.4986	2.7558

Table 2: *In vivo* data for DNA damage gathered and adapted from a previous study published by Checkley *et al.* [1].

Time (hours)	Volume (cm^3)	Std. Error (cm^3)
0 mg/kg (control)		
168	0.3028	0.0219
264	0.5189	0.0465
360	0.9095	0.0934
456	1.3857	0.1554
504	1.5646	0.1483
25 mg/kg QD		
168	0.3037	0.0342
264	0.4411	0.0704
360	0.5617	0.0840
456	0.7064	0.1221
504	0.8701	0.1187
50 mg/kg QD		
168	0.3106	0.0332
264	0.3971	0.0768
360	0.4002	0.0817
456	0.4783	0.0966
504	0.4923	0.0846

Table 3: *In vivo* data for tumour volume gathered and adapted from a previous study published by Checkley *et al.* [1].

2 Calibrating the model using *in vitro* data

Using a minimal-parameter approach, seven model parameters are calibrated using the *in vitro* data previously produced by Checkley *et al.* [1], as listed in Table 1 in Section 2.7 in the main paper. Parameter sensitivity is explored in the sensitivity analysis described later on in the Supplementary Material. The calibration process is outlined in Sections 2.1 through to 2.4. The *in vivo* calibration is described in Section 2.5.

2.1 Cell doubling

In the model, the doubling time of a cell i is denoted τ_i, where τ_i is stochastically picked from a normal distribution with mean value μ and standard deviation σ. Thus μ corresponds to the

average cell doubling time and σ corresponds to how synchronised the cells are. If σ is zero, then all cells have perfectly synchronised cell cycles and duplicate at the same time. Higher σ values achieve less synchronised cell cycles amongst cells and smoother cell count growth curves over time. The choice of using a normal distribution from which to pick τ_i is motivated by the observation that the cells in the *in vitro* experiment are 'fairly synchronised', as can be seen by observing the cell count data listed in Table 1. The control case (i.e. no drug) cell count data is used to estimate μ and σ.

By observing the control case cell count data, we conclude that the average doubling time for cells should be between 22 and 26 hours, hence σ-values in the parameter range [22,26] hours are investigated *in silico*. Note that in the last 24-hour interval, between 48 and 72 hours, the control population increases by less than 5%. However, according to cell count data from earlier time poins, we assume that cells are on the brink of cell division at the end time of the experiment and thus we disregard the influence of the 72-hour data point in the overall model calibration. Due to the synchronised nature of the cell count data, σ-values between 0 and 2.5 hours were investigated *in silico*, where $\sigma = 0$ h corresponds to completely synchronised cells and $\sigma = 2.5$ h achieves a smooth cell count growth curve. After an iterative process of tuning parameters and running *in silico* experiments, the calibrated values are set to be $\mu = 24$ hours and $\sigma = 0.5$ hours.

2.2   Cell cycle progression

The *in vitro* data provides information on how many cells are in the damaged S state via the biomarker γH2AX. For the control case, the number of γH2AX positive cells in our mathematical model depends on two variables: (1) the probability (Π_{D-S}) that a cell enters the D-S state and (2) the amount of time ($\Theta_{D-S} \cdot \tau_i$) spent in the D-S state prior to repairing. Recall that Θ_{D-S} is the fraction of a cell's doubling time (τ_i) spent in the D-S state. As a first step, *in silico* experiments are performed in which we find various parameter pairs (Π_{D-S}, Θ_{D-S}) that agree with the control data. We thereafter note that the *in vitro* drug effect saturates for concentrations 3, 10 and 30 μM and assume that the maximal dose (30 μM) yields 100% D-S to S repair inhibition. Thus a second step we test the variable pairs (Π_{D-S}, Θ_{D-S}) for this 'maximal drug and no repair' scenario *in silico*, and we match these *in silico* results to the 30 M *in vitro* data. Here, we only use data from early time points (time < 12 hours) in order to avoid the influence that dying cells have on the data and model outputs. After iterative *in silico* testing, the variable pair (Π_{D-S}, Θ_{D-S}) that best fits these both extreme cases is $\Pi_{D-S} = 0.75$ and $\Theta_{D-S} = 0.03$. The first extreme case refers to the 'no drug' *in silico* experiment matched to the *in vitro* control data, where we assume that all D-S cells repair to state S. The second extreme case refers to the 'maximum drug' *in silico* experiment matched to the 30μM control data, where we assume that no D-S cells repair to state S.

2.3   Drug Response

Drug effects are modelled using the sigmoid E-max model [2], where the drug effect E is a function of the drug concentration C, so that

$$E(C) = E_{max} \cdot \frac{C^\gamma}{EC_{50}^\gamma + C^\gamma},$$

where E_max denotes the maximal drug effect. Here we set $E_{max} = 1$ to corresponds to total D-S to S repair inhibition. EC_{50} denotes the drug concentration that achieves half of the maximal drug effect and γ is the Hill-coefficient. If drug effect is plotted over time, the EC_{50}-value determines the asymptotic behaviour of the effect whilst the γ-value determines how quickly the asymptotic value is reached.

From the *in vitro* data, we note that the drug concentration 1 μM achieves roughly half of the total drug effect, and further that the drug concentrations 0.3 and 3 μM respectively achieve less than, and more than, the half of the total drug effect. With this in mind, EC_{50} values over 0.3 and under3 μM are investigated with various Hill coefficients to fit *in vitro* data for all (non-control) drug concentrations. In order to avoid the impact that dying cells have on the data used parameterise EC_{50} and γ, only early *in vitro* data (time < 12 hours) is used to guide the calibration. After iterative *in silico* testing, the best variable pair (EC_50, γ) is determined to be $EC_{50} = 1$ μM and $\gamma = 2$.

2.4 Cell death

In the *in vitro* experiments, cells that are damaged (but not yet dead) are γ-H2AX positive. In the model, the time it takes between the 'lethal event' (i.e. a cell's failure to repair) and a cell being 'dead' is denoted $T_{L \to D}$ and is matched from the *in vitro* experiment. After noting the asymptotic behaviour of the *in vitro* data, both in terms of cell damage and cell count, we estimate that the rate of cell elimination should roughly correspond to the rate of cell production, and thus $T_{L \to D}$ should be in the same order of magnitude as the doubling time. Consequently, values of $T_{L \to D}$ between 0 and 2 τ_i are explored *in siico* after which $T_{L \to D} = \tau_i$ is chosen as it best matches the *in vitro* data for all tested (non-control) drug concentrations.

2.5 *In vivo* calibration

For the control case, the *in vivo* model is directly calibrated by the *in vitro* data, and no further calibration is needed. For drug concentrations larger than 0 μM, we use the *in vivo* data for the highest administered drug dose to calibrate the model in order to disregard details concerning pharmacokinetics and bioavailability. In future work, our model can be integrated with pharmacokinetic modelling techniques.

3 Cross-Section to Tumour Spheroid Extrapolation

When implementing our mathematical *in vivo* model, only a central cross-section of the tumour is actually simulated *in silico* and post simulation time this cross-section area (that is approximately circular) is extrapolated to a tumour volume (that is approximately spherical). From the extrapolated tumour spheroid, the two outputs $\tilde{X}_1$ (percentage of γH2AX positive cells) and

$\tilde{X}_2$ (tumour volume) are gathered. This is done by using simulated areas to compute the total tumour volume,

$$\tilde{X}_2 = \text{Total Tumour Volume} = \frac{4\pi}{3} \left(\frac{\text{Total Simulated Area}}{\pi} \right)^{3/2}, \tag{1}$$

and the quiescent tumour volume,

$$\text{Quiescent Tumour Volume} = \frac{4\pi}{3} \left(\frac{\text{Quiescent Simulated Area}}{\pi} \right)^{3/2}. \tag{2}$$

From the above, the volume of cycling, or proliferating cells, is obtained by

$$\text{Cycling Tumour Volume} = \text{Total Tumour Volume} - \text{Quiescent Tumour Volume}. \tag{3}$$

Now the output $\tilde{X}_1$ can be computed where,

$$\tilde{X}_1 = \text{Percentage of } \gamma\text{H2AX positive cells in sphere} =$$
$$\frac{\text{Number of simulated } \gamma\text{H2AX positive cells}}{\text{Number of simulated cycling cells}} \times \frac{\text{Cycling Tumour Volume}}{\text{Total Tumour Volume}}. \tag{4}$$

6

4   Uncertainty and Sensitivity Analyses

To evaluate the *in silico* findings obtained in our *in vitro* study, three uncertainty and sensitivity analyses techniques are performed. The three techniques are namely: (1) Consistency Analysis, which is used to determine how many *in silico* runs should be performed before defining results in terms of statistical metrics in order to mitigate uncertainty originating from intrinsic model stochasticity, (2) Robustness Analysis, which investigates model sensitivity to local parameter perturbations and (3) Latin Hypercube Analysis, which investigates model sensitivity to global parameter perturbations. These techniques are thoroughly described in a review/methodology paper [3] which provides background information concerning the origin of these techniques as well as detailed information on how to implement them. To perform uncertainty and sensitivity analyses we need to specify a set of inputs and outputs. Here, the output variables are X_1: the percentage of γH2AX-positive (i.e. damaged) cells at the end time of the experiment (72 hours), and X_2: the cell count (i.e. the number of non-dead cells) at the end of the experiment. The input variables are the seven model parameters listed in Table 1, in the main article, that we calibrate using *in vitro* data. These inputs are namely μ, σ, Π_{D-S}, Θ_{D-S}, EC_{50}, γ and $T_{L \to D}$.

4.1   Consistency Analysis

Results from the Consistency Analysis are provided in Figures 1, 2, 3, 4, 5 which show the $\hat{A}$-measures, in both computed and scaled forms, for the distribution sizes $n = 1, 5, 50, 100, 300$ respectively. By observing Figures 1 through to 5, it is clear that the statistical significance decreases with increasing distribution size n, as is shown in Figure 6 and Table 4.1 which show the maximal scaled $\hat{A}$-values for all tested distribution sizes. These results demonstrate that the distribution size $n = 100$ is the smallest tested distribution size that yields a small statistical significance (i.e. a maximum scaled $\hat{A}$-value smaller than 0.56) for both regarded output variables X_1 and X_2. From this we decide to base every *in silico* result (here in terms of mean values and standard deviations) on 100 simulation runs.

distribution size output	n=1	n=5	n=50	n=100	n=300
X_1	1	0.92	0.61	0.55	0.54
X_2	1	0.84	0.59	0.55	0.54

Table 4: Maximal scaled $\hat{A}$-values produced in the Consistency Analysis for various distribution sixes n. The output variables are X_1, corresponding to the percentage of γH2AX positive (i.e. damaged) cells, and X_2, corresponding to the cell count.

4.2 Robustness Analysis

We use Robustness Analysis to investigate how sensitive the output is to *local* parameter perturbations, that is to say when input parameters are varied one at a time. Figures 7, 8, 9, 10, 11, 12, 13 provide boxplots and $\hat{A}$-measures that demonstrate the effect that local perturbations of the input variables μ, σ, Π_{D-s}, Θ_{D-S}, EC_{50}, γ and $T_{L\rightarrow D}$ respectively have on the output variables X_1 and X_2. Key findings are listed below, discussing the impact of one input parameter at a time.

▶ Remarks regarding input parameter μ: Figure 7 shows that, for small parameter perturbations, increasing the average doubling times of cells, μ, overall decreases the percentage of γH2AX positive cells and increases the cell count, however this decrease/increase is not linear. This indicates that the results of the *in vitro* simulation (and of the *in vitro* experiment nonetheless) are sensitive to the *timing* of the drug administration. In other words, Robustness Analyses demonstrates that treatment responses depend on how many cells are in the susceptible cell-cycle state at time of drug administration.

▶ Remarks regarding input parameter σ: Figure 8 demonstrates that the level of cell cycle synchronisation amongst cells, quantified by the input σ, affects *in silico* outputs for small parameter perturbations. The results indicate that for highly asynchronised cells (i.e. high σ-values) the smoother growth curves yield higher cell counts at certain time-points (such as the end time 72 hours) and a lower percentage of γH2AX-positive cells. As discussed in the remark above, the timing between cell cycles and drug administration affect treatment responses.

▶ Remarks regarding input parameter Π_{D-S}: Figure 9 illustrates that increasing the probability that a cell enters the damaged S state, i.e. the variable Π_{D-S}, increases the percentage of γH2AX cells and decreases the cell count, as expected.

▶ Remarks regarding input parameter Θ_{D-S}: Figure 10 shows how the amount of time that damaged cells spend in the D-S state before attempting to repair, and thus the Θ_{D-S}-value, affects the output. Results show that the percentage of γH2AX positive cells increases with increasing values of Θ_{D-S}, as more damaged cells will accumulate in the D-S state. However, this does not affect the probability of cells repairing, so the cell count is not as sensitive to small perturbations of Θ_{D-S}. The value of $\Theta D - S$ implicitly affects the measured cell count at the end time of the experiment as a decreased/increased $\Theta D - S$-value yields a slightly decreased/increased time lag between a cell entering the D-S state and dying.

▶ Remarks regarding input parameter EC_{50}: Figure 11 demonstrates that output variables are highly sensitive to perturbations of EC_{50}. Increasing EC_{50} results in a higher percentage of γH2AX positive cells and a lower cell count. Thus the input parameter EC_{50} should be regarded as a highly influential on quantitative results.

▶ Remarks regarding input parameter γ: Figure 12 illustrates that output variables measured at the end time of the experiment are not very sensitive to small perturbations of γ. This

8

can be understood as the γ parameter inherently corresponds to 'how quickly' a drug achieves asymptotic behaviour in the E_{max} model, the model used in our mathematical framework to express cellular drug response.

▶ Remarks regarding input parameter $T_{L\rightarrow}$: Figure 13 shows how output variables change as a result of perturbations to the input variable $T_{L\rightarrow}$, that describes how long it takes for a cell that has failed to repair to die (i.e. how long a cell with a 'death-sentence' is picked up as γH2AX positive in the *in vitro* experiment). Results show that both the percentage of γH2AX positive cells and the cell count increases with increasing values of $T_{L\rightarrow}$, as dying cells will categorised as γH2AX positive longer before being categorised as dead. When calibrating the model, we avoid the effect of this input parameter by only regarding *in vitro* data at time points that are early enough to correspond to systems with no (or a negligible amount of) dead cells.

4.3 Latin Hypercube Analysis

Latin Hypercube Analysis is here used to investigate how sensitive output responses are to *global* parameter perturbations. We here investigate parameter values within ranges that we consider to be 'plausible' from the calibration process and the Robustness Analysis. Figures 14, 15, 16, 17, 18, 19, 20 provide scatter-plots that demonstrate correlations between the output variables X_1 and X_2 and the input variables μ, σ, Π_{D-s}, Θ_{D-S}, EC_{50}, γ and $T_{L\rightarrow D}$ respectively. The Pearson Product Moment Correlation Coefficients between the various input-output pairs are listed in Table 4.3. To determine threshold values for correlation coefficient descriptors, we compromise between suggested values by other authors [3], and take into account the fact that we are only regarding parameter values within 'plausible' ranges. With this as a guide, we here decide that our obtained correlation coefficients with a magnitude in [0,0.12] corresponds to the linear input-output relationship being 'negligible', [0.19, 0.35] 'weak', [0.48,0.59] 'moderate' and 0.84 'strong'. Key findings from the Latin Hypercube Analysis are listed below, where the impact of one input parameter is discussed one at a time.

▶ Remarks regarding input parameter μ: Figure 14 and the first column in Table 4.3 show that, for the allowed parameter range, μ and X_1 are moderately, negatively correlated as the correlation coefficient is -0.48 and the scatterplot displays an overall trend of the output (X_1) decreasing with increasing values of the input μ. The relationship between μ and the other output variable X_2 is, on the other hand, negligible. We explain this by the fact that treatment responses are sensitive to the timing of the drug administration, but there is a time-lag $T_{L\rightarrow D}$ between a cell's lethal event (failure to repair) and its death. As damaged (but not dead) cells are included in the cell count, the (μ,X_2)-relationship is more weakly linearly correlated than the (μ,X_1)-relationship.

▶ Remarks regarding input parameter σ: Figure 15 and the second column in Table 4.3 demonstrate that the linear relationships between input variable σ and the output variables X_1 and X_2 are both negligible, within the regarded input parameter value range.

▶ Remarks regarding input parameter Π_{D-S}: Figure 16 and the third column in Table 4.3 indicate that the relationships between the input variable Π_{D-S} and the output variables

X_1 and X_2 are, respectively, positively and negatively weakly linearly correlated. This agrees with the intuitive notion that if the probability that a cell enters the D-S state increases, cell damage (X_1) increases whilst the cell count (X_2) decreases.

▶ Remarks regarding input parameter Θ_{D-S}: Figure 17 and the fourth column in Table 4.3 show that the input variable Θ_{D-S} is has a negligible linear correlation with the output variables X_1 and X_2.

▶ Remarks regarding input parameter EC_{50}: Figure 18 and the fifth column in Table 4.3 demonstrate that the input variable EC_{50} impacts the output responses more than do other input variables, within the regarded ranges for input variables. EC_{50} is negatively, moderately linearly correlated with X_1 and EC_{50} is strongly, positively linearly correlated with X_2. These relationships are visually apparent in the regarded scatterplots.

▶ Remarks regarding input parameter γ: Figure 19 and the sixth column in Table 4.3 indicate negligible linear correlations between the input parameter γ and both output variables X_1 and X_2.

▶ Remarks regarding input parameter $T_{L \to D}$: Figure 20 and the last column in Table 4.3 demonstrate that the input variable $T_{L \to D}$ is positively, weakly, linearly correlated with the output X_1, whilst the linear correlation between $T_{L \to D}$ and X_2 is negligible.

input / output	μ	σ	Π_{D-S}	Θ_{D-S}	EC_{50}	γ	$T_{L \to D}$
X_1	-0.48	0.06	0.19	0.06	-0.59	0.05	0.35
X_2	0.12	0.01	-0.24	-0.02	0.84	0.12	0.00

Table 5: Pearson Product Moment Correlation Coefficients between input and output variables obtained in the Latin Hypercube Analysis.
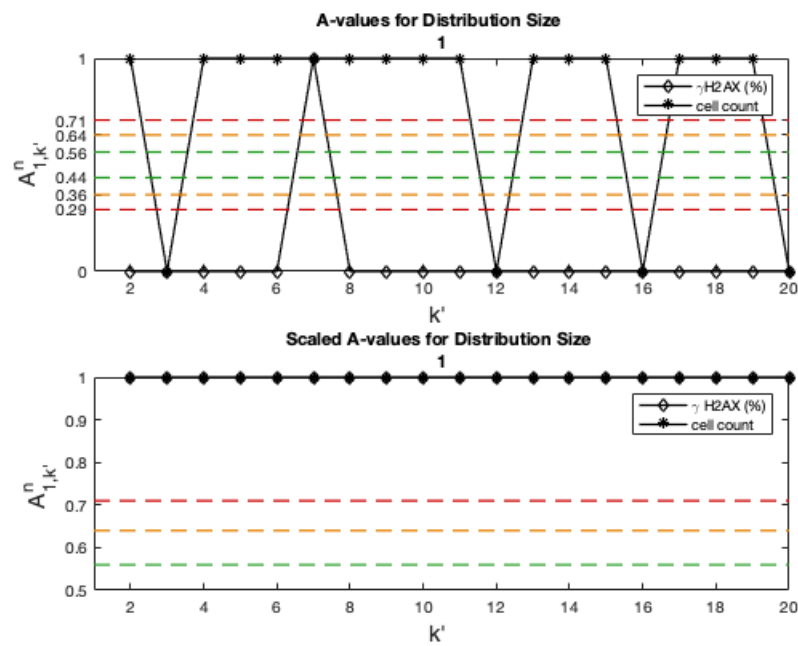
Figure 1: Consistency Analysis. $\hat{A}$-values in initial (top) and scaled (bottom) form for distribution size $n = 1$.
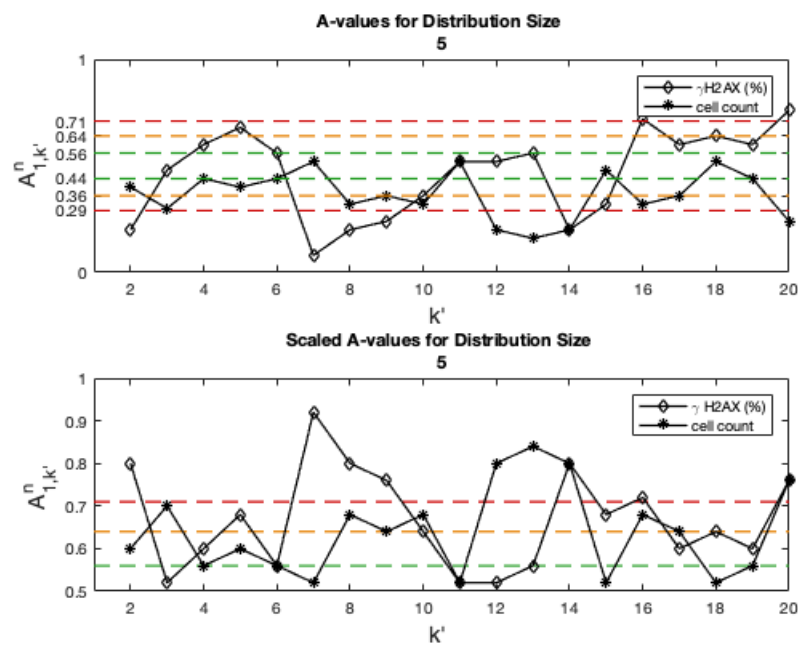
Figure 2: Consistency Analysis. $\hat{A}$-values in initial (top) and scaled (bottom) form for distribution size $n = 5$.

11
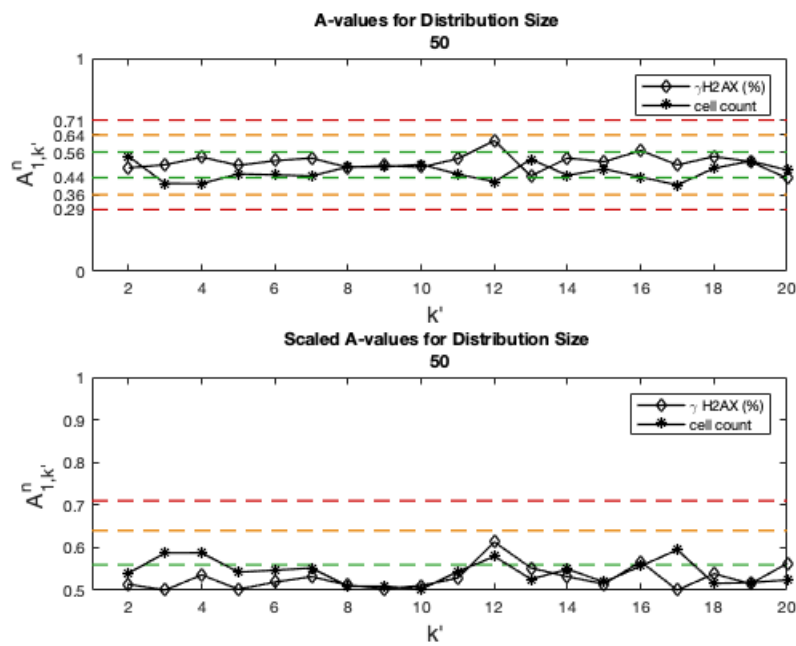
Figure 3: Consistency Analysis. $\hat{A}$-values in initial (top) and scaled (bottom) form for distribution size $n = 50$.
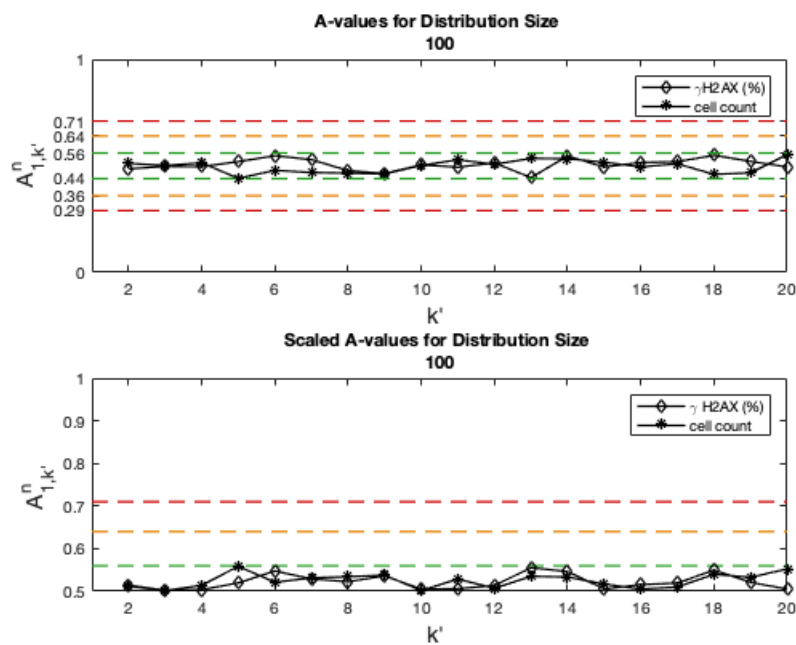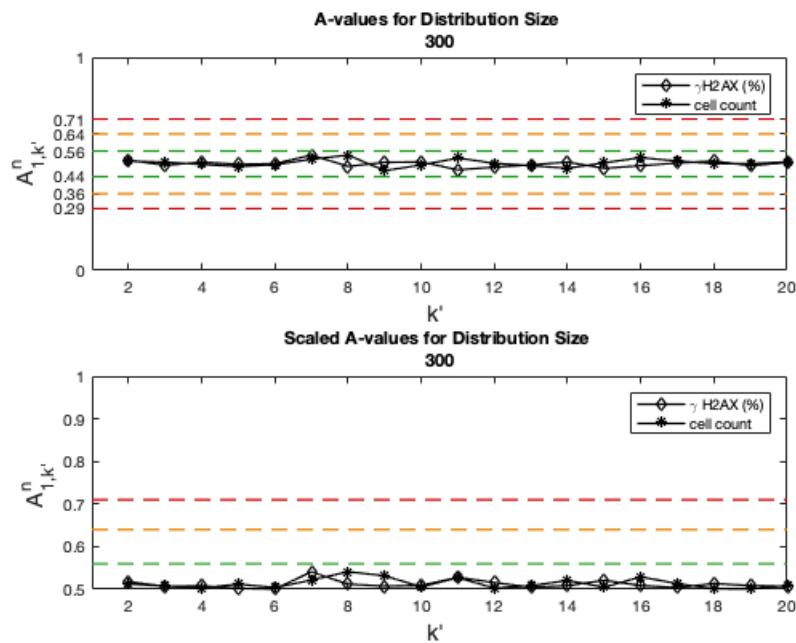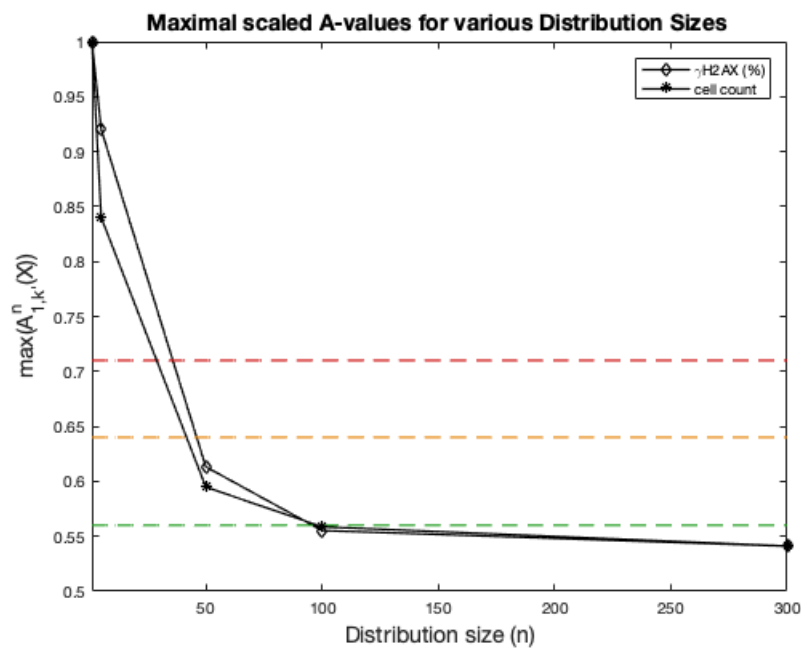

Figure 4: Consistency Analysis. $\hat{A}$-values in initial (top) and scaled (bottom) form for distribution size $n = 100$.

Figure 5: Consistency Analysis. $\hat{A}$-values in initial (top) and scaled (bottom) form for distribution size $n = 300$.

Figure 6: Consistency Analysis. Scaled $\hat{A}$-values for various distribution sizes tested in the Consistency Analysis.

13
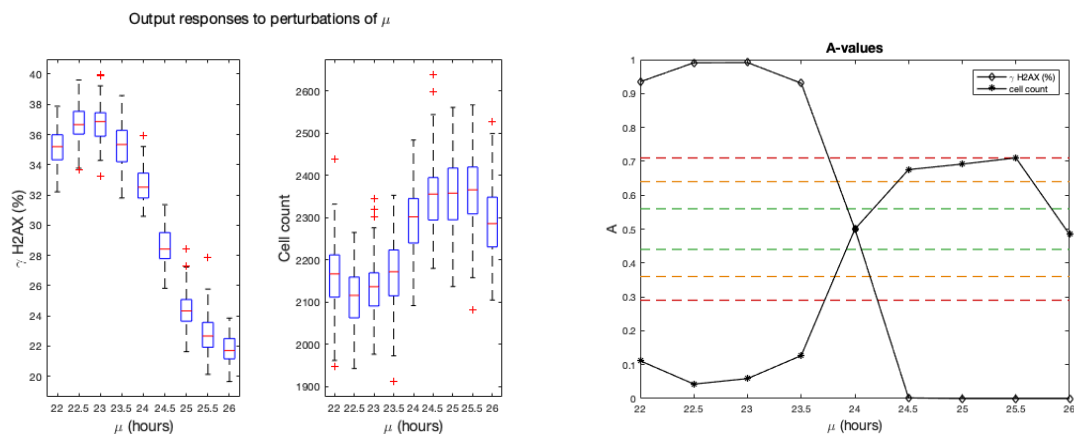
Figure 7: Robustness Analysis. Left: Output responses, in terms of percentage of γH2AX positive (i.e. damaged) cells, and cell count as a result of perturbations to the input variable μ. Right: Maximal $\hat{A}$-values resulting from comparisons between distributions with perturbed data and a distribution with calibrated (unperturbed) data.
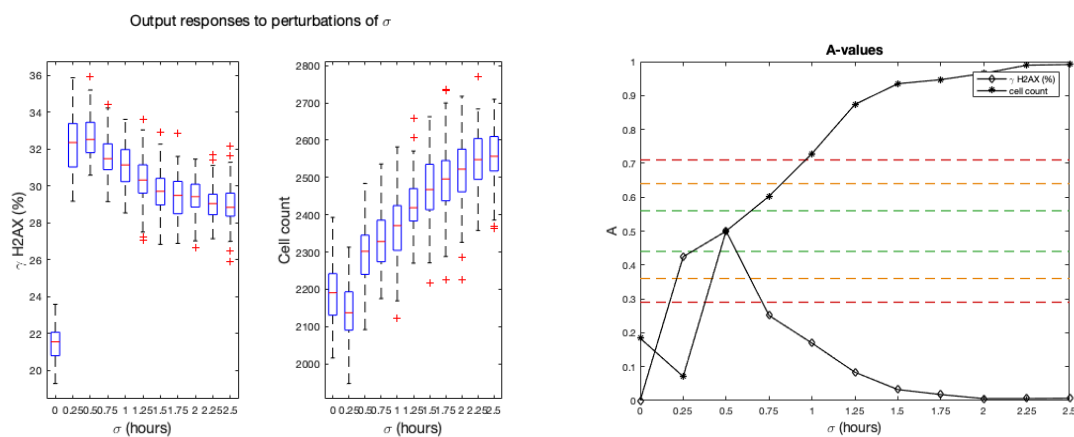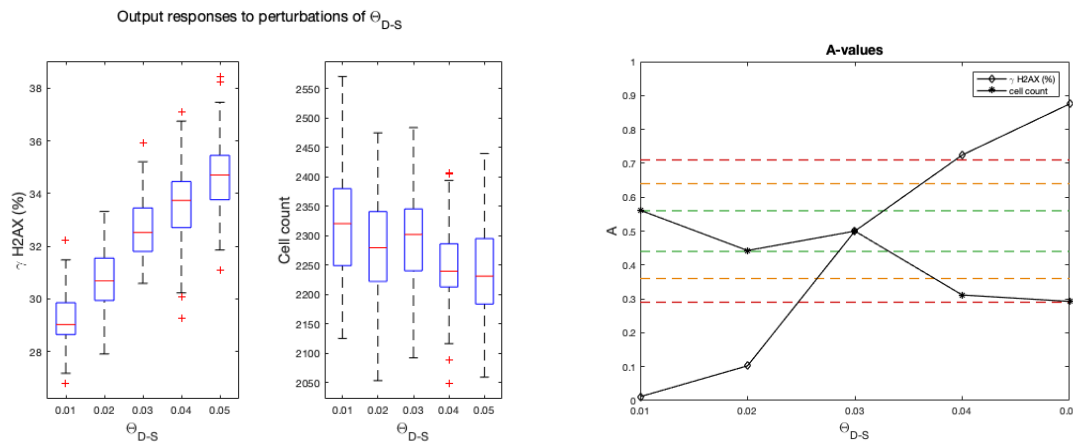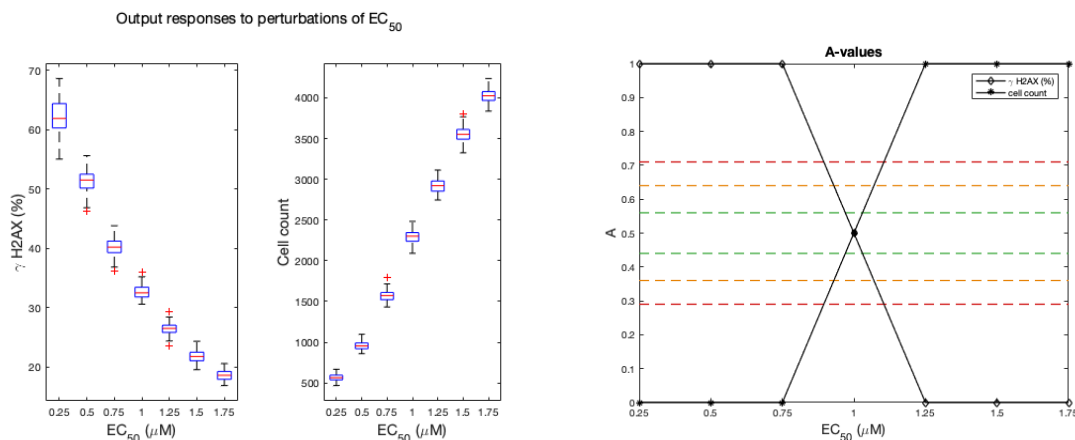

Figure 8: Robustness Analysis. Left: Output responses, in terms of percentage of γH2AX positive (i.e. damaged) cells, and cell count as a result of perturbations to the input variable σ. Right: Maximal $\hat{A}$-values resulting from comparisons between distributions with perturbed data and a distribution with calibrated (unperturbed) data.

14
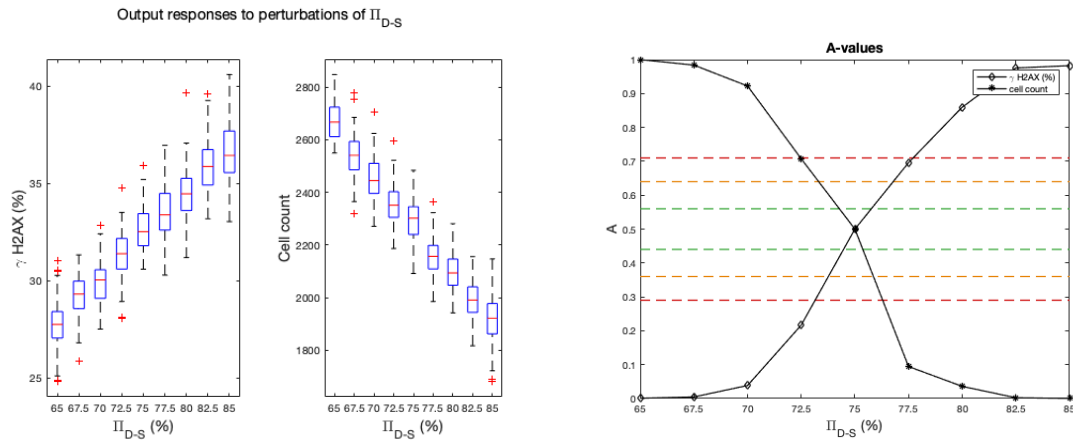
Figure 9: Robustness Analysis. Left: Output responses, in terms of percentage of γH2AX positive (i.e. damaged) cells, and cell count as a result of perturbations to the input variable Π_{D-s}. Right: Maximal $\hat{A}$-values resulting from comparisons between distributions with perturbed data and a distribution with calibrated (unperturbed) data.
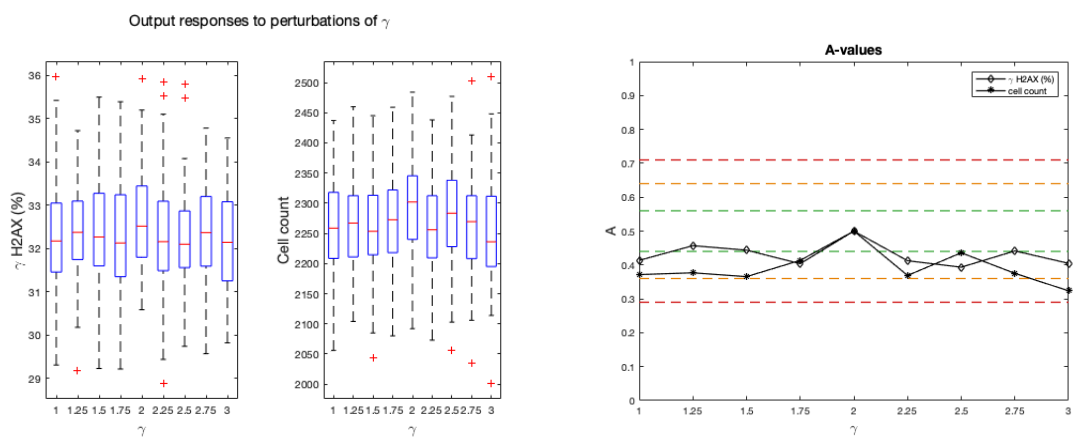

Figure 10: Robustness Analysis. Left: Output responses, in terms of percentage of γH2AX positive (i.e. damaged) cells, and cell count as a result of perturbations to the input variable Θ_{D-S}. Right: Maximal $\hat{A}$-values resulting from comparisons between distributions with perturbed data and a distribution with calibrated (unperturbed) data.
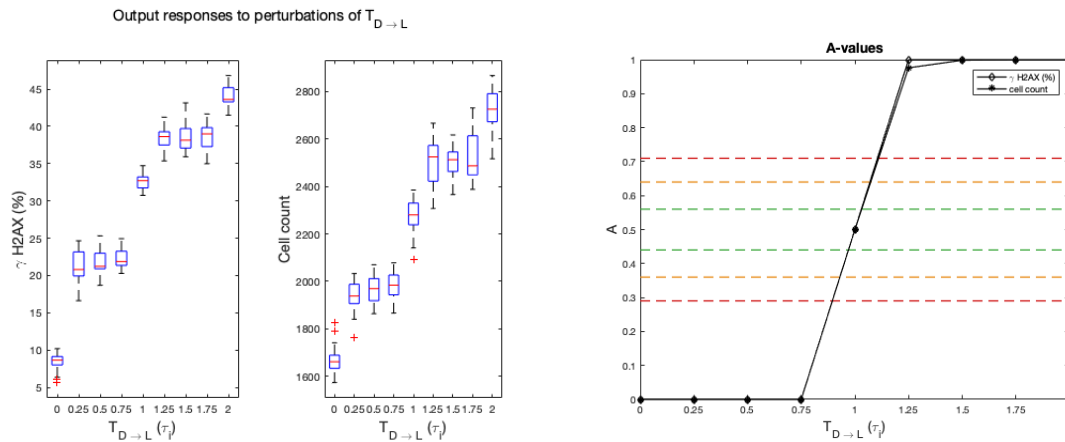
15

Figure 11: Robustness Analysis. Left: Output responses, in terms of percentage of γH2AX positive (i.e. damaged) cells, and cell count as a result of perturbations to the input variable EC_{50}. Right: Maximal $\hat{A}$-values resulting from comparisons between distributions with perturbed data and a distribution with calibrated (unperturbed) data.


Figure 12: Robustness Analysis. Left: Output responses, in terms of percentage of γH2AX positive (i.e. damaged) cells, and cell count as a result of perturbations to the input variable γ. Right: Maximal $\hat{A}$-values resulting from comparisons between distributions with perturbed data and a distribution with calibrated (unperturbed) data.

Figure 13: Robustness Analysis. Left: Output responses, in terms of percentage of γH2AX positive (i.e. damaged) cells, and cell count as a result of perturbations to the input variable $T_{D \to L}$. Right: Maximal $\hat{A}$-values resulting from comparisons between distributions with perturbed data and a distribution with calibrated (unperturbed) data.
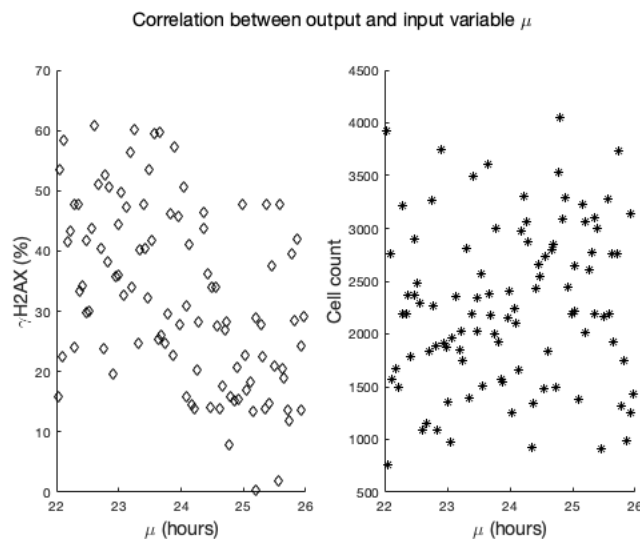
Figure 14: Latin Hypercube Analysis. Outputs in terms of 2AX positive cells (left) and number of viable cells (right) when global parameter perturbations are performed. The scatter-plots show the correlation between outputs and the input variable μ.
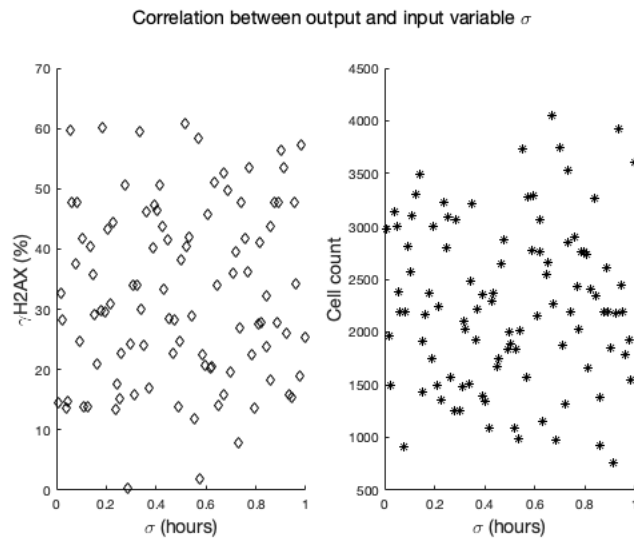
17

Correlation between output and input variable σ

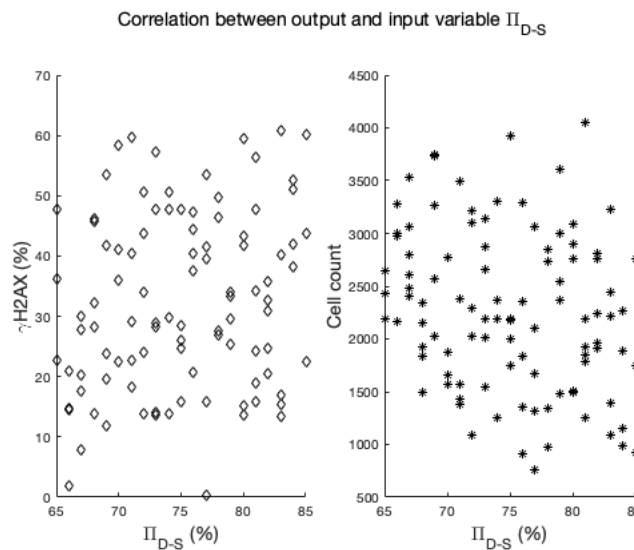

Figure 15: Latin Hypercube Analysis. Outputs in terms of γH2AX positive cells (left) and number of viable cells (right) when global parameter perturbations are performed. The scatterplots show the correlation between outputs and the input variable σ.

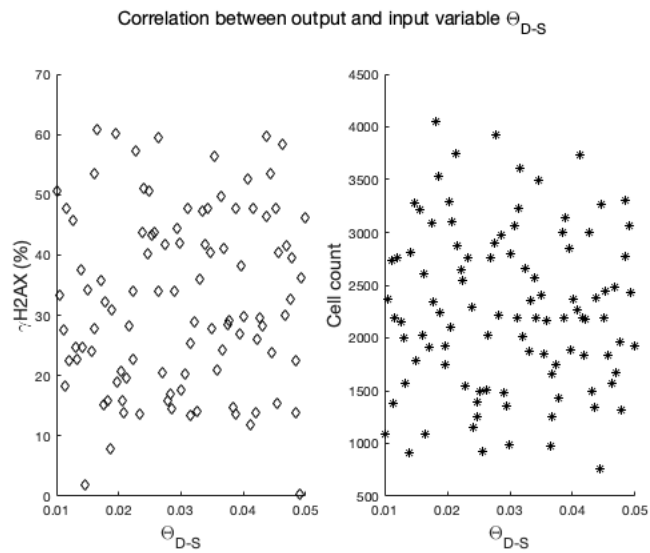Correlation between output and input variable $\Pi_{D\text{-}S}$

Figure 16: Latin Hypercube Analysis. Outputs in terms of γH2AX positive cells (left) and number of viable cells (right) when global parameter perturbations are performed. The scatterplots show the correlation between outputs and the input variable Π_{D-S}.

18

Correlation between output and input variable $\Theta_{D\text{-}S}$

Figure 17: Latin Hypercube Analysis. Outputs in terms of γH2AX positive cells (left) and number of viable cells (right) when global parameter perturbations are performed. The scatterplots show the correlation between outputs and the input variable Θ_{D-S}.
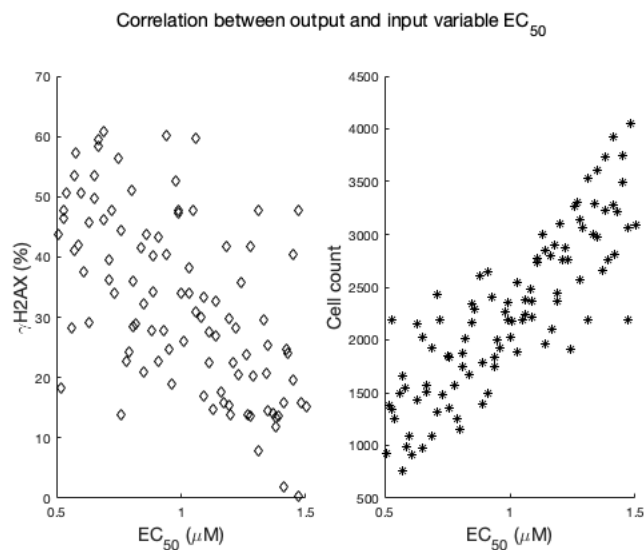
Correlation between output and input variable EC_{50}

Figure 18: Latin Hypercube Analysis. Outputs in terms of γH2AX positive cells (left) and number of viable cells (right) when global parameter perturbations are performed. The scatterplots show the correlation between outputs and the input variable EC_{50}.
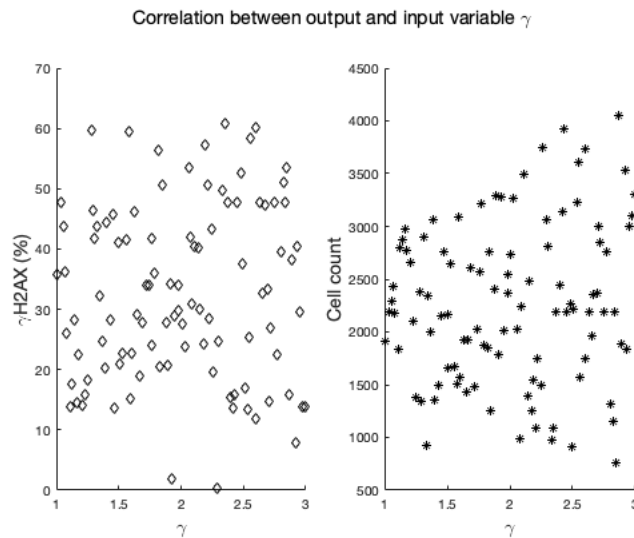
19

Figure 19: Latin Hypercube Analysis. Outputs in terms of γH2AX positive cells (left) and number of viable cells (right) when global parameter perturbations are performed. The scatterplots show the correlation between outputs and the input variable γ.
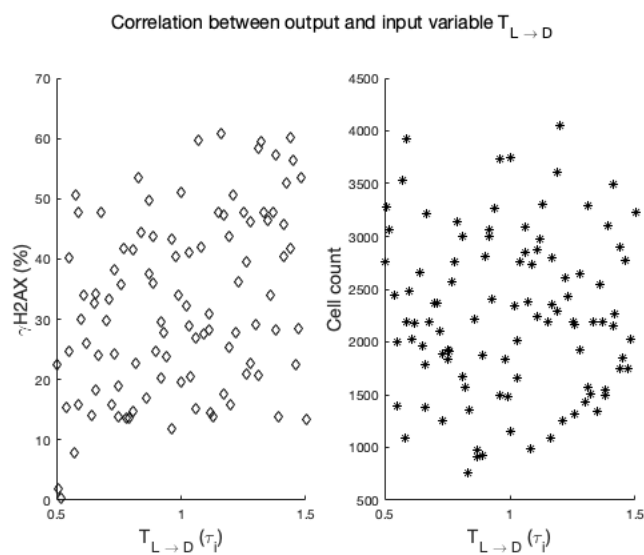

Figure 20: Latin Hypercube Analysis. Outputs in terms of γH2AX positive cells (left) and number of viable cells (right) when global parameter perturbations are performed. The scatterplots show the correlation between outputs and the input variable $T_{L \to D}$.

References

[1] S. Checkley, L. MacCallum, J. Yates, P. Jasper, H. Luo, J. Tolsma, and C. Bendtsen, "Bridging the gap between in vitro and in vivo: Dose and schedule predictions for the ATR inhibitor AZD6738," *Sci Rep* **5** (Aug, 2015) 13545.

[2] N. Holford, "Pharmacodynamic principles and the time course of immediate drug effects. ," *Transl Clin Pharmacol* **4** (2017) 157–161.

[3] S. Hamis, S. Stratiev, and G. Powathil, "Uncertainty and sensitivity analyses methods for agent-based models: A review," *(To Appear)* .