Bayesian Modelling to Characterise the Responder Profile to a Novel Cancer Immunotherapy

Supervisors: Professor Reiko Tanaka, Doctor Tara Hameed
v2.2

Alexandre Yann Péré
CID: 01938104
January 5, 2024

Word Count: 4148

Contents

1 Project Specification

1.1 Background

1.1.1 Cancer

Cancer is a large class of diseases that is the second leading cause of death in the United-States [1]. While the immune system has the potential to target and eliminate cancer cells, cancer often finds ways to evade these natural defenses [2]. Traditional methods, such as chemotherapy or surgery, rely on using destructive external agents to kill the cancerous cells. However, introducing foreign agents in the body often results in heavy side effects [3]. This prompted the development of immunotherapies, a type of treatment aimed at countering cancer's ability to escape immune detection, which thus has the potential to be less toxic. Several viable strategies exist for immunotherapy [4]. We will first review the general principles behind these stratgeies.

1.1.2 Immunotherapies

Cytokine-based Therapies

Cytokine-based therapies rely on the injection of specific cytokines (small proteins that act as signalling molecules during the immune response) to control tumour growth [5]. One of the most promising cytokine thus far is the interleukin-12 (IL-12), that was shown to have potent antitumour effects [6]. While it does not directly affect tumour cells, it mediates the production of other molecules or cells that have a more direct effect [7]. First of all, it activates the production of tumour-infiltrating cytotoxic cells, mainly $CD8^+$ [8]. These are a type of T-lymphocytes whose main function is to carry out cytotoxic activity (i.e. killing the malignant cells) after detecting tumoural antigen [9]. Secondly, they induce production of another type of cytokine, called interferon-γ (IFNγ) [10]. IFNγ in turn affects the tumour microenvironment by stimulating production of cytotoxic cells [11], reducing angiogenesis [12] and upregulating antigen-presenting pathways within tumour cells [13]. Lastly, IL-12 facilitates T-cell proliferation (including $CD8^+$) by reducing negative regulatory pathways that lead to immunosuppression [14]. It does so by inhibiting the effect of immune checkpoint Programmed Death 1 (PD1), following a similar strategy to checkpoint inhibitor (CPI) treatments. While these three pathways indicate that IL-12 has a very robust antitumour effect, clinical studies demonstrated that systemic injection of IL-12 is exceedingly toxic as it triggers a large immune response throughout the whole body [15][16]. These severe treatment-related adverse effects (TRAEs) dampened research about IL-12, waiting for a safer, more localised delivery method to be found.

Immune Checkpoint Inhibitors

The usual partner of cytokine-based treatments are checkpoint inhibitors. To understand checkpoint inhibition, we must first review in more detail the negative regulatory pathways of $CD8^+$ T-cells activity. The most potent pathway involves checkpoint molecules, either Cytotoxic T-lymphocyte antigen 4 (CTLA4) or programmed cell death 1 (PD1) [17]. Both molecules are membrane protein receptors that act with some delay to exhaust and deactivate T-cell functions after they are stimulated by antigen-presenting cells (APC). Both CTLA4 and PD1 function in similar ways, the main difference being the type of tissues they affect [18]. Although their original function was shown to be prevention of autoimmunity [19], they lead to immunosuppression in the presence of tumours. The idea behind CPI treatments is to inhibit these regulators to shift the tumour microenvironment away from immunosuppression. Clinical trials demonstrated positive results in several types of cancers , but performed poorly against immunologically

cold tumours, i.e. tumour that do not normally elicit a strong immune response (they escape the immune system very effectively), such as melanoma [20].

1.2 Previous Work on Computational Modelling

To have a better understanding of the immune response to the treatment, T. Miyano, 2019 [23], proposed to use a computational modelling approach. Computational models are common in pharmacodynamics as they can be analysed with mathematical tools, potentially revealing key mechanisms to optimise the treatment. He developped an initial mechanistic model based on Delay-Differential Equations (DDEs), parameterised by 21 parameters representing various relevant biological factors of a given mouse, such as the tumour growth rate or the degradation rate of IFNγ [23]:

$$
\begin{aligned}
\dot{g}(t) &= k_1 + k_2[d_{CBD}(t) + d_{12}(t)] - d_1 g(t) \\
\dot{c}(t, t - t_d) &= k_3 + k_4 g(t - t_d) - d_2 c(t) \\
\dot{p}(t) &= k_5 - [d_3 + d_4 g(t)]p(t) \\
\dot{v}_l(t) &= k_6 \left[1 - \frac{v(t)}{v_{max}}\right] v_l(t) - \left[d_5 + \frac{\frac{d_6 c(t)}{1+s_1 p(t)(1-d_{CPI}(t))} + d_7 g(t)}{1 + s_2 v(t)}\right] v_l(t) \\
\dot{v}_d(t) &= \left[d_5 + \frac{\frac{d_6 c(t)}{1+s_1 p(t)(1-d_{CPI}(t))} + d_7 g(t)}{1 + s_2 v(t)}\right] v_l(t) - d_8 v_d(t)
\end{aligned}
$$

The five state variables (g, c, p, v_l and v_d) are concentration of IFNγ, of CD8$^+$ and of PD1 along with volume of living and dead tumour, respectively. This was motivated by the fact that these are the key players in the immune response, as explained above. The meaning of each parameter is reported in the Appendix. The model was investigated by C. Hines, who showed that the model could successfuly reproduce experimental data by using a Genetic Algorithm for parameter fitting [24]. However, C. Hines also demonstrated in a subsequent analysis that the model had [significant differences] with findings from the biologists in two ways. A positive feedback loop, where IL-12-induced IFNγ in turn produces IL-12, is missing from the model [25]. Additionally, C. Hines showed that the model outcome does not depend much on the initial tumour volume and treatment characteristics (day of injection, number of doses, etc.)[24], which is opposite to results reported in [21]. Rather, he found that the model depends almost exclusively on parameter k_6, d_1 and s_2 (which respectively corresponds to proliferation rate of tumour; degradation rate of IFNγ and tumour-induced immunosuppression strength)

1.3 Motivation

Recent endeavours in this field of immunotherapies led to the development by Ishihara group at Imperial College in 2019 [21] of a new molecule, CBD-IL-12, that demonstrated promising results to treat melanoma. The CBD-IL-12 molecule consists of a collagen-binding protein (or collagen-binding domain, CBD) that is fused onto a IL-12 cytokine. The modified interleukin hence mainly accumulates in collagen-rich regions. As collagen is the main component of cancerous microenvironment [22], this effectively results in an enhanced delivery method that can achieve high concentration of IL-12 specifically in cancerous microenvironments. In mice tumour-models, this novel molecule achieved a CR [must explain CR rate] rate of up to 67% ($n = 15$) for melanoma, and 87% ($n = 15$) for breast cancer when combined with CPI drugs (a mix of both anti-PD1 and anti-CTLA4). While these results are encouraging, the study showed that

such high CR-rates could only be achieved in very specific settings (such as a tumour volume of 70mm^3 upon injection [add details]). Different settings (e.g. volume of 150mm^3) elicited little to no response. This heterogeneity of treatment outcome could not be explained by the authors. The first step to improve efficacy of CBD-IL-12-based treaments would thus be to understand better what are the key parameters that control the treatment outcome.

As outlined above, the aim of the project is to identify biological factors involved in the immune response that determine treatment outcome (CR vs non-CR), along with their critical value that lead to a bifurcation. These biological factors correspond to the parameter of the DDE model. However, the first key issue is that these parameters are kinetic rates that cannot be measured experimentally. We thus have to estimate them solely from the measurable data (i.e., tumour volume). This process is called parameter estimation. The other key challenge is that the data collected concerns multiple mice drawn from the same population. There is hence some shared information between the mice, as well as some variability between that explain the range of output obtained in the lab. Nonlinear mixed-effects models [26] are a standard method to analyse population dataset with both shared information (called fixed-effects) and individual variability (called random-effects).

1.4 Aims and Objectives

The aim of this project is to use computational models to improve our understanding of the immune mechanisms behind the CBD-IL-12 immunotherapy, utlimately to characterise the responder profile for the treatment.

Objective 1: change the initial mechanistic model to a mixed-effect model so that it can reproduce the experimental data while including all the important pathways of the immune response.
Objective 2: analyse the mixed-effect model to identify key biological factors in mice that determine the outcome of the treatment, along with the corresponding threshold that separates complete response (CR) from non-CR

2 Ethical Analysis

Treatments and experiments on the specimens used, mice in this case, were approved by the Institutional Animal Care and Use Committee of the Univerity of Chicago (see Methods section of [21]).

The data derived from these experiments holds immense potential for understanding treatment response in cancer, aiding in the optimization of therapies for human use.

Long-term implications involve the potential for groundbreaking advancements in cancer treatment, benefiting society globally. However, considerations extend beyond treatment efficacy to encompass environmental impact, fostering sustainable practices in research, and ethical ownership and dissemination of findings. Collaboration among colleagues, stakeholders, and the wider community fosters responsible research conduct and knowledge dissemination while ensuring transparency and accountability. Additionally, financial sustainability and social responsibility in implementing these therapies for public benefit are essential considerations.

To ensure that the results are true and reliable, the full analysis along with the code will be published on GitHub.

3 Literature Review

In this section we will give an overview of the mathematical tools relevant for the project that were mentioned in Section 1. Starting by presenting the process of parameter estimation, we will then focus specifically on the general Bayesian method and extend it its hierarchical version.

3.1 Overview of Parameter Estimation

Since the model developed by Miyano is based of DDE, we will focus in the followings sections on DDEs. Let the general definition of a DDE model be:

$$\frac{\mathrm{d}X_i}{\mathrm{d}t} = f_i(t, \boldsymbol{X}(t), \boldsymbol{X}(t-\tau)|\boldsymbol{\theta}), \qquad t \in [t_0, t_{max}], i = 1, \ldots, I$$

where τ denotes a constant delay, so that the rate of change of state X_i depends on both the present state $\boldsymbol{X}(t)$ and a past state $\boldsymbol{X}(t-\tau)$. The subscript i indexes the different state variables of interest, and $\boldsymbol{\theta}$ is the (unknown) vector of the parameters for the DDE model. This parameter vector is different for each treated mouse, as it uniquely characterises its treatement response, and hence we denote with $\boldsymbol{\theta}_j$ the parameter vector that characterises the j-th mouse. The experimentally observed tumour evolution for the j-th mouse is denoted by $\boldsymbol{y}_j$, and each of its elements is the tumour volume observed at a given time, noted $y_j(t)$. The aim of parameter estimation is to retrieve the parameter vector that can reproduce the observed data. Mathematically, this can be expressed as maximising the likelihood of the observations $\mathcal{L}(\theta; \boldsymbol{y})$, where η_j is the random effect for individual j [29].

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{y}) = \prod_j \mathcal{L}(\theta; \boldsymbol{y}_j) = \prod_j p(\boldsymbol{y}_j|\boldsymbol{\theta}) = \prod_j \int p(\boldsymbol{y}_j|\eta_i, \boldsymbol{\theta}) p(\eta_j|\boldsymbol{\theta})\, d\eta_j$$

Many statistical approaches have been developed to perform parameter estimation on differential equation models from noisy data [27]. However, most of them cannot be applied to mixed-effect models. In the context of pharmacodynamics model , Donnet et al, 2013 [28] reviewed different techniques available to perform parameter estimation. For analysis of population data with observational noise, two methods are available: Expectation-Maximisation (either Stochastic [29] or using First-Order Condition [30] (FOCE)) and Bayesian parameter estimation [31]. The authors concluded that Bayesian modelling in particular is the most flexible method, since it does not rely on assumptions and hence work both inidividual or population datasets, with or without noise, which is not the case for alternative methods. Additionally, it benefits from theoretical validity although at the cost of being more computationally intensive, whereas the FOCE method was not proven to converge to the true posterior. Another relevant advantage is that Bayesian approach do not provide point estimate, but rather distributions, which could be key to explain outcome heterogeneity. Hence the Bayesian modelling approach seems to be the most relevant for the project.

3.2 Bayesian Parameter Estimation

Bayesian Parameter Estimation is a method to estimate $\boldsymbol{\theta}_j$ given an observation vector $\boldsymbol{y}_j$. Contrary to frequentist approach, estimations are in the form of probability distributions (called posteriors, denoted $p(\boldsymbol{\theta}|\boldsymbol{y})$) rather than point estimates.

For a situation where data about only one individual was gathered, the posterior distribution is defined as follows [32]:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto p(\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{\theta})$$

This formula is the direct application of Bayes' theorem. It is the product of the prior distribution $p(\boldsymbol{\theta})$, which represents our knownledge of the problem, and the likelihood $p(\boldsymbol{y}|\boldsymbol{\theta})$. Before further defining these distributions, we must extend this definition of the posterior distribution to work for mixed-effect models, called hierarchical models in Bayesian statistics.

3.3 Hierarchical Modelling

We seek to estimate the probability distribution of the parameter vector $\boldsymbol{\theta}$ for each individual, however these vectors are not independent from each other since each individual is sampled from a common population (mixed-effect model). To analyse mixed-effect model, hierarchical Bayesian modelling is a standard method [28][31]. For example, Rosenbaum et al., 2019 [31], studied models of predator-prey systems, which also display radically different behaviours depending on the values of certain kinetic rates that cannot be directly measured. By fitting times series of measurable data to a hierarchical Bayesian model, they could not only extract a specific set of parameters for each individual system; but also determine the regions in parameter space that led to radically different types of behaviour accross the population. As this study present many similarities with the current project, it shows that hierarchical Bayesian modelling is a promising tool that is worth exploring.

Hierarchical modelling enables us to formulate that the parameter vector $\boldsymbol{\theta}_j$ is sampled from an population-level distribution characterised by the (also unknown) hyperparameters $\boldsymbol{\phi}$. The objective is hence to find the distribution of both $\boldsymbol{\theta}_j$ $\forall j$ and $\boldsymbol{\phi}$. The Bayesian parameter estimation framework integrates this additional assumption by changing the posterior to [32]:

$$p(\boldsymbol{\theta}, \boldsymbol{\phi}|\boldsymbol{y}) \propto p(\boldsymbol{\phi})p(\boldsymbol{\theta}|\boldsymbol{\phi})p(\boldsymbol{y}|\boldsymbol{\theta})$$

This expression is a product of the hyperprior $p(\boldsymbol{\phi})$, the population distribution $p(\boldsymbol{\theta}|\boldsymbol{\phi})$ and the likelihood $p(\boldsymbol{y}|\boldsymbol{\theta})$.

3.3.1 Hierarchical Priors and Hyperpriors

In hierarchical Bayesian models, there are two types of parameters: hyperparameters $\boldsymbol{\phi}$ and individual parameters $\boldsymbol{\theta}_j$ [32]. The key feature is that the simulated data is directly conditioned on the regular parameters, which are themselves drawn from population-level distributions characterised by hyperparameters. Hence, this results in two types of priors: hierarchical priors, that specify how to sample $\boldsymbol{\theta}$ using the hyperparameters $\boldsymbol{\phi}$; and hyperpriors that convery our knowledge about the potential values of $\boldsymbol{\phi}$.

For the case of Bayesian parameter estimation on sets of time series, Rosenbaum et al, 2019 [31], proposed to sample each ODE parameter from a Normal distribution. As Normal distribution are characterised by two values (mean μ and standard deviation σ), this resulted in two hyperparameters per ODE parameter.

This can be summarised as follows, for a given scalar parameter θ:

$$\begin{aligned} \theta \sim p(\theta|\phi) \Leftrightarrow \theta \sim \mathcal{N}(\phi_\mu, \phi_\sigma) \quad & \textit{hierarchical prior} \\ \phi_\mu \sim p(\phi_\mu) \quad & \textit{hyperprior for the hyper-mean} \\ \phi_\mu \sim p(\phi_\mu) \quad & \textit{hyperprior for the hyper-standard deviation} \end{aligned}$$

By adding additional hyperparameters, this definition could be modified to allow for bimodal distribution instead of a simple normal distribution for the hierarchical priors.

3.3.2 Likelihood function

The likelihood function in the Bayesian framework is how we model the dynamics of the underlying process. In pharmacodynamics, it is the mechanistic model of the immune response. Assuming that the observational noise is a white Gaussian noise with zero-mean and a standard deviation σ_{err} that is common to all experiments, the likelihood can then be defined as follows [27], [33] (for a given tumour evolution $\boldsymbol{y}_j$):

$$\mathcal{L}(\boldsymbol{\theta}_j) = \prod_{t=t_0}^{t_{max}} \frac{1}{\sigma_{err}} \exp\left(-\frac{(y_j(t) - Y_j(t|\boldsymbol{\theta}_j))^2}{2\sigma_{err}^2}\right)$$

Where $Y_j(t|\boldsymbol{\theta}_j)$ is the simulated time series using the DDE model parameterised by $\boldsymbol{\theta}_j$, t is the time index for the time series and σ_{err} is the experimental error. This is usually obtained by numerical methods. Without additional information about the measurement methods, this is the approach suggested by Rosenbaum et al [31].

To find the marginal posterior distribution for the hyperparameter vector $\boldsymbol{\phi}$, we can marginalise out θ by integrating over θ the full joint posterior distribution (ie. accross all time series):

$$p(\phi|y) = \int \prod_{j=1}^{J} p(\theta_j, \phi|y_j)\, \mathrm{d}\theta$$

3.4 Reduction of the Computational Burden

3.4.1 Sensitivity Analysis

As pharmacodynamics models can generally contain a high number of parameters, the Bayesian approach, which is computationally intensive [28], can result in intractable computations. To this end, Vasquez-Cruz et al, 2012 [34], proposed a method to reduce the number of parameters in an ODE model. By taking the example of a crop growth ODE model with 17 parameters, they used a senstivity analysis (namely, eFAST and Sobol' method) to identify the most influencial parameters, and used these results to design a reduced model with only 7 free parameters. More specifically, they followed a procedure highlighted by Varela et al, 2010 [need paper], according to which parameters that account for less than 10% of the total effects can be considered as fixed parameters. Then, using a parameter fitting algorithm, they were able to show that this reduced model could still replicate the experimental data with minimal error. Additionally, they showed that the two sensitivity analysis methods (eFAST and Sobol' methods) yield different results. Hence, to ensure that all the important parameters are detected, both methods should be used.

3.4.2 High-performance Computational Methods

As the estimation of the posterior distributions cannot be evaluated analytically, Luengo et al., 2020 [35], pointed out that Markov-Chains Monte Carlo approach to approximate the posterior is the only feasible approach in most applications. While many different MCMC methods exist [36], it was shown by Nishio et al, 2019 [37] that the No U-Turn Sampler (NUTS) led to lower skewness of the posterior and more decorrelated samples compared to the other two popular algorithms, Gibbs and Hamiltonian Monte Carlo samplers. However, traditional MCMC methods, including NUTS, can have convergence problem [38]. Since we hypothesised that the posterior distribution might be multimodal, we need also need a robust sampler for this case. Liu et al, 2018 [27], proposed a new sampler called Stochastic Approximation Monte Carlo (SAMC), which is based on the idea of partitioning the parameter space in a finite number of subspaces and forcing the algorithm to explore each subspace. By running the algorithm on a parameter estimation problem for an ODE model of predator-prey system (defined in [39]), they showed that SAMC could successfully retrieve the correct parameter value, while traditional MCMC methods such as the Metropolis-Hastings method were stuck in local modes.

3.4.3 Transforms

Another finding from the study of predator-prey systems [31] is that log-transformation of the residuals make the inference much more robust, resulting in faster convergence and more accurate posteriors. This changes the likelihood function to the following:

$$\mathcal{L}(\boldsymbol{\theta}_j) = \prod_{t=t_0}^{t_{max}} \frac{1}{\sigma_{err}} \exp\left(-\frac{[\ln(y_j(t)) - \ln(Y_j(t|\boldsymbol{\theta}_j))]^2}{2\sigma_{err}^2}\right)$$

4 Risk Register

The main risks are :

5 Evaluation

Below we present a list of the key components of the cancer immunotherapy project, along with a way to verify that they function correctly.

Bayesian Model

To build and validate the Bayesian model, we will follow the Bayesian Workflow procedure outlined in [40]. It consists of three different tests:

- **Prior Predictive Check:** we sample 1,000 sets of parameter from the priors and simulate tumour growth for each of them. The 95% credible interval of the resulting collection of time series should contain our expected range of curves we can expect. Evaluation of this step is mostly qualitative.
- **Fake Data Check:** we first need to generate a artificial dataset using known values of parameters, and then fit the Bayesian model to these fake datasets. By comparing the results of the model (i.e. the estimated parameter value in the form of the posterior distribution) to the true values, we can conclude whether the model can successfully perform parameter estimation or not. The approach proposed by Gelman et al. [40] is to verify that the 95% credible interval of the posterior

Table 1: Table of the different risks associated with the project's objectives

Risk	**Likelihood**	**Impact**	**Mitigation Strategy**
MCMC chains do not converge	High	Very high	Two alternative methods can be used in this case. **Approximate Bayesian Computations,** which is a approach that does not rely on a likelihood function. It is relatively easy to implement but introduces additional approximation error. **Model simplification**, an approach proposed in [40]. It consists of simplifying the likelihood function, and does not introduce additional error.
Model does not pass validation protocol	High	Very high	Modify the immune response model to ensure that all key interactions are translated in the mechanistic model. We especially plan to implement the positive feedback loop mentionned in Section 1.2.

includes the true parameter value. Additionally, another possible approach detailed in [31] involves substracting the true parameter value from the posterior. This results in a "error" distribution that should theoretically be zero-mean.

- **Posterior Predictive Check:** this is analogous to the prior predictive check, except that parameter are drawn from the marginal posterior distributions instead of the prior distributions. This results in a collection of simulated time series. The 95% credible interval should contain the experimental time series obtined in the lab.

Convergence of the MCMC Chains

As mentioned above, convergence of the MCMC chains is a critical element that needs to be evaluated, as it reflects the quality of the approximation of the posterior distribution [cite T. Moins et al]. The most popular method to assess convergence is the potential scale factor reduction, usually termed $\hat{R}$, developed by Gelman et al., 1992 [41]. It can be calculated as follows:

$$\hat{R} = \frac{m+1}{m}\frac{\hat{\sigma}_+^2}{W} - \frac{n-1}{mn}$$
$$W = \frac{1}{m(n-1)}\sum_{j=1}^{m}\sum_{t=1}^{n}(\psi_{jt} - \bar{\psi}_j)^2$$

Where m is the number of chains used to explore the posterior in parallel, n is their length, $\hat{\sigma}_+^2$ is the variance of the chain with the highest variance, ψ_{jt} is the value of the j-th chain at the t-th iteration, and $\bar{\psi}_j$ the mean value of the j-th chain. The authors additional highlight that a value close to 1 usually indicates convergence, usually the criterion is $\hat{R} < 1.05$ for convergence.

Responder Profile

Once the responder profile has been characterised, its accuracy can be evaluated by comparing the pre-

dicted treatment outcome against the true outcome on a new batch of cancerous mice. The procedure to inoculate skin cancer, inject the immunotherapy and measure tumour volume should be the same as defined in [21] and [23]. To be robust against all cases, only half of the batch of cancerous mice should be predicted as complete responder according to the responder profile. This ensures that we can collect data about the accuracy for both true positives and of true negatives.

6 Preliminary Results

6.1 Numerical Analysis

6.1.1 Numerical Stability Analysis

The very first aspect of Miyano's model that we wanted to verify was its ability to capture two specific treatment outcome: CR vs non-CR. As these behaviours can essentially be characterised by the fixed-points of the model (if the steady-state behaviour of the model converges to high values of tumour volume, it is a non-CR behaviour, and vice-versa. See definition in Section 1). We opted for a grid-search stability analysis, meaning that we sample reguarly-spaced points in parameter space and classify them as either CR or non-CR. Fig. 1 shows the resuls of this analysis, where each axis of the cube corresponds to the value of a k_6, s_2 or d_1 respectively. A can be seen, there seem the be a clear boundary between the two response modes, with very little "mixing". This suggests that it would be possible to predict how a given patient would respond to the treatment, by knowning on which side of the boundary he is.

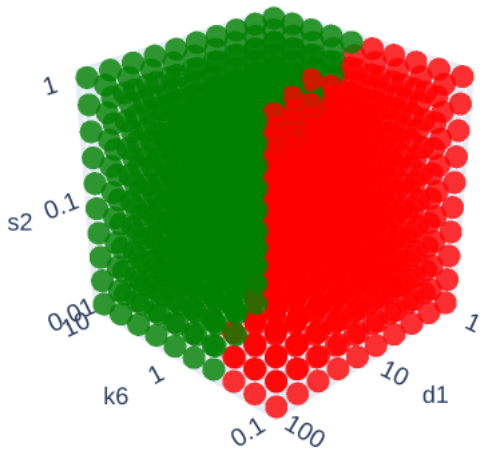


Figure 1: Stability analysis shows that there is a clear boundary in parameter space between CR and non-CR

6.1.2 Sensitivity Analysis

In order to restrict the parameter space for subsequent analysis, and also to understand to main mechanisms behind the immune response, we performed a eFAST sensitivity analysis, which is a variance decomposition method. As it can only decompose variance of a scalar metric, it does not natively support time-series. Hence we chose to apply it to the integral of the tumour growth curve simulated by Miyano's model. This metric simply reflects the cumulative tumour volume over time, which is ultimately the

quantity that we want to minimse. Other metrics which capture more specific feature of the tumour growth might be considered in the future, such as the derivative of tumour growth on the last simulated day to capture the patient's potential to go into CR. Results are shown in Fig. 2. The total height of the bar represent the fraction of the variance that is imputable to the corresponding parameter. The first observation we can make is that model is mostly sensitive to k_6, d_1, d_7 and s_2. However, we can see that the main effect indices (in blue) are almost always negligible compared to the interaction indices (orange). According to a study by Vazquez-Cruz et al. (2012), this is a typical sign of non-identifiability [34] that will significantly hinder Bayesian inference. Additionally, results indicate that the treatment characteristics (labelled `t_d`, `t_delay` and `t_last`) have almost no impact on the treatment outcome, which is conflicting with the results experimentally obtained in the CBD-IL-12 study [21]. Hence we suspect that the current model does not correctly model the immune response mechanisms.

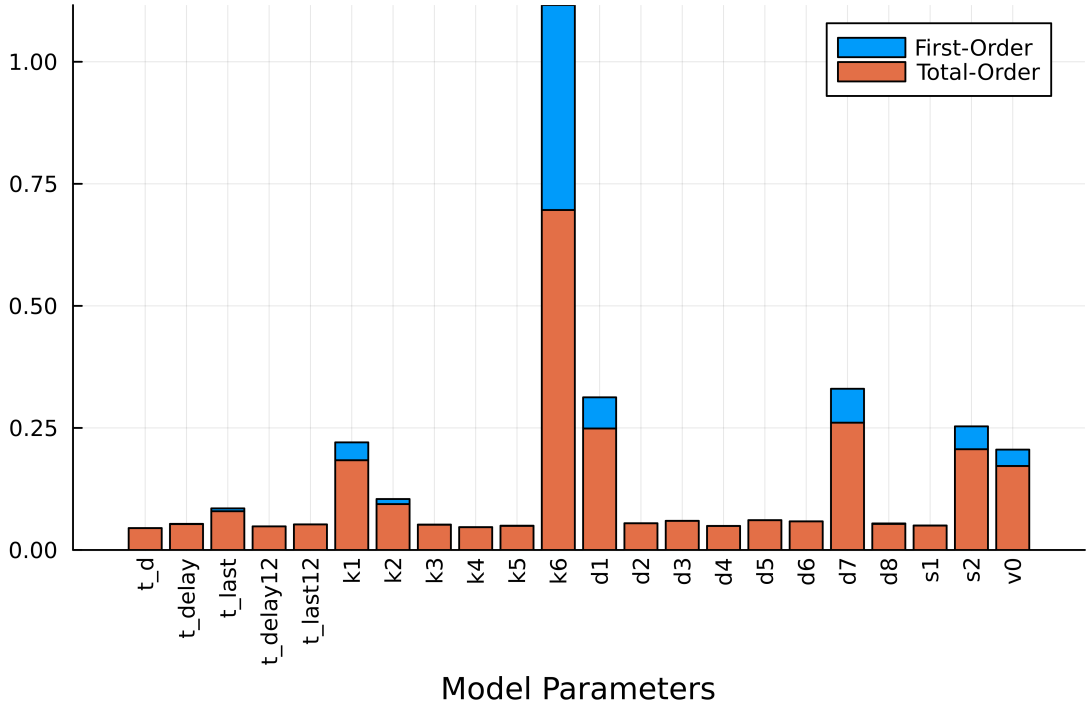


Figure 2: eFAST sensitivity analysis on Miyano's model shows that k_6 is the most influencial parameter by far, followed by d_1, d_7 and s_2. This is conflicting with findings from Ishihara group

6.2 Bayesian Model Validation

In this section we show how the Bayesian model was validated, following the procedure highlighted in Section 5. We focus on a reduced model with only three free parameters, k_6, d_1, s_2, which were identified by C. Hines are the most impactful ones [24]. All other parameters of the model were fixed to their respective value estimated by C. Hines to model the mean response curve. This ensures that each parameter is assigned a realistic yet arbitrary value.

6.2.1 Prior Predictive Check

In this case, we do not have much data on the typical values of the parameters since it is impossible to measure it (we only know that it is a positive number whose typical value is between 0 and 1, as evidenced by [24]), so we aim to design an uninformative prior. Fig. 3 shows a plot of 1,000 tumour growth time-series. Each of them was simulated using a set of parameters drawn from the following prior distribution:

$$\ln(k_6) \sim \text{Cauchy}^-(0, 1)$$
$$\ln(d_1) \sim \text{Cauchy}^+(0, 1)$$
$$\ln(s_2) \sim \text{Cauchy}^-(0, 1)$$

We chose Cauchy distributions since that carry less information that standard Gaussian distributions, allowing for value far from their center of mass. This is critical since we do not have information about the true values of the parameters. As we exponentiate the Cauchy distribution, it means that $0 < k_6 < 1$. The blue shade represents the 95% credible interval, and the dark green line is the median growth curve. As we can see, the 95% credible interval is can virtually contain our expected range of curves, since it ranges from 0 (minimum volume) to 600 (maximum possible volume according to the equation), meaning that they are relatively uninformative priors. The median curve has the shape of the typical growth curve, as observed in the labs. Hence, we can say that the prior distribution is satisfying, as it could explain any potential growth curve while restricting the values of the parameters to a smaller subset of $\mathbb{R}$.

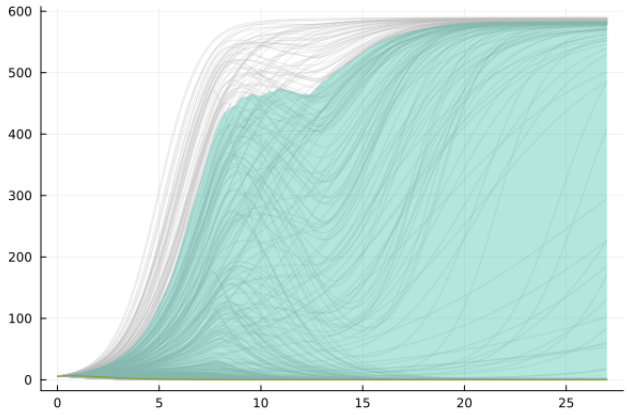


Figure 3: ODE solution for 1,000 parameter values sampled from the prior ($\boldsymbol{\theta} \in \mathbb{R}^1$)

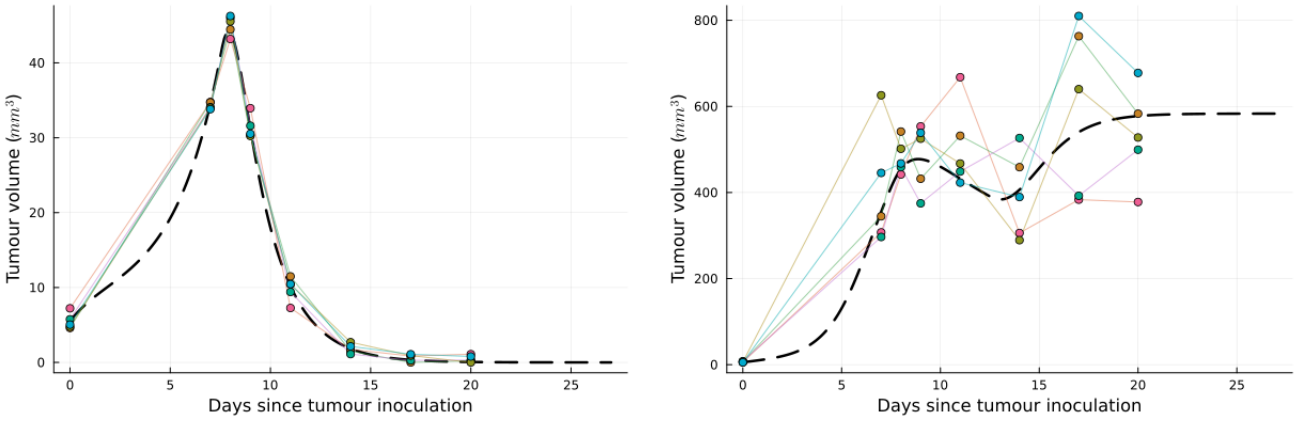
6.2.2 Fake Data Check

Each fake growth curve was generated by sampling a value of θ from the prior distribution, and then simulating the tumour growth in the same way as for the Prior Predictive Check. However, as biological data is always noisy, we also added some noise to make the fake dataset closer to what we would actually expect from the labs. This was done in two different ways, resulting in two distinct datasets. For dataset A, we simply added a white standard Gaussian noise to the simulation. For dataset B, we use added white noise to the log of the simulated curve (i.e. multiplicative noise). Using a standard Gaussian would

result in too large noise values, so we chose a standard deviation of 0.3 to achieve similar levels of noise compared to datasete A. The generation process is summarized in Table 3, where x_* denotes a noiseless data point. The reason for using two different noise generation is that we observed, in the experimental data from the labs, that data points are usually more dispersed when they have a high value, suggesting an exponential relationship.
Additionally, each dataset contains 10 time series. Fig. 4 plots the fake data points against the original curve. For clarity, only 5 time series, selected at random, were shown.

Table 3: Summary of the generation process for the two datasets A and B

Dataset	Generation Process
A	$x_A = x_* + \mathcal{N}(0, 1)$
B	$x_B = x_* \times e^{\mathcal{N}(0, 0.3)}$



(a) Dataset A

(b) Dataset B

Figure 4: Plot of the fake data points (colored lines and scatter plot) along with the original growth curve (dashed line)

Results

Before checking if the estimated values match the true ones, we first assess convergence of the MCMC chains. Th $\hat{R}$ values (see Section 5) are reported in Table. 4 as the average $\hat{R}$ value accross the 3 parameters. It must be noted that often, out of the 5 chains per inference, some chains get trapped (assessed by visual inspection). In that case, they are excluded from the $\hat{R}$ calculation, and this is reported in the "Number of Chains" column. If all chains are exlcuded, meaning that none of them converged, we simply report N/A.

Looking at Table 4, we can hence conclude that the chains did not converged, meaning that the model cannot make inference with $\boldsymbol{\theta} \in \mathbb{R}^3$ and uninformative priors. To further diagnose the model, we performed

Table 4: Assessement of convergence for the MCMC chains for uninformative priors (3 free parameters)

Data set	Pooling Type	$\hat{R}$ diagnostic	Number of Chains	Convergence
A	None	116.08	N/A	No
	Complete	298.67	N/A	No
B	None	1.092	3/5	No
	Complete	3.11	N/A	No

another set of inferences, except that the priors where highly informative:

$$\ln(k_6) \sim \text{Cauchy}^-(\theta_{k_6}, 1)$$
$$\ln(d_1) \sim \text{Cauchy}^+(\theta_{d_1}, 1)$$
$$\ln(s_2) \sim \text{Cauchy}^-(\theta_{s_2}, 1)$$

where θ_x represent the true value of parameter x. Convergence of this new set of inferences is shown in Table 5. As we can see, convergence of the MCMC chains are still very poor, even with highly informative priors centered on the true parameter values.

Table 5: Assessement of convergence for the MCMC chains for uninformative priors (3 free parameters)

Data set	Pooling Type	$\hat{R}$ diagnostic	Number of Chains	Convergence
A	None	18.15	N/A	No
	Complete	45.96	N/A	No
B	None	1.505	3/5	No
	Complete	1.014	2/5	Yes

It might be objected that Cauchy distributions are by definition not too informative since a non-negligible portion of their mass stretches well beyond their standard deviation, contrary to normal distributions. This hence motivated us to perform one last fake data check, using the normal priors shown below to be even more informative:

$$\ln(k_6) \sim \mathcal{N}^-(\theta_{k_6}, 0.3)$$
$$\ln(d_1) \sim \mathcal{N}^+(\theta_{d_1}, 0.3)$$
$$\ln(s_2) \sim \mathcal{N}^-(\theta_{s_2}, 0.3)$$

Convergence results are shown in Table. 6. The main result is that chains converged or were close to convergence only for dataset D, showing that a log-normal transformation is key to make exploration of the posterior easier to perform. Whilst the overall convergence rate is still very low given the informative normal priors, this series of fake data checks for the case of three free parameters highlighted the key role of the log-transformation.

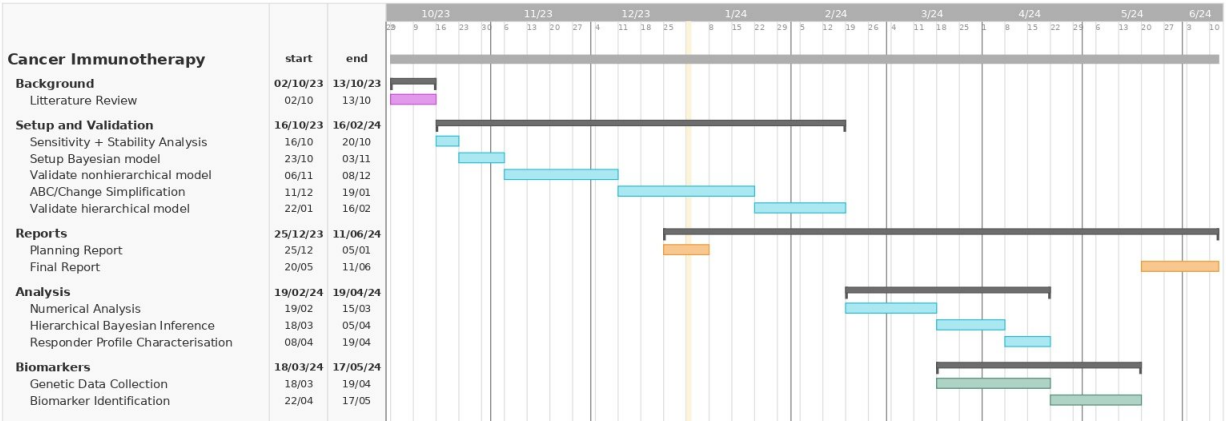
Conclusion

Table 6: Assessement of convergence for the MCMC chains for uninformative priors (3 free parameters)

Data set	Pooling Type	$\hat{R}$ diagnostic	Number of Chains	Convergence
A	None	7.387	N/A	No
	Complete	39.17	N/A	No
B	None	1.066	3/5	Yes
	Complete	1.111	5/5	No

Even with informative priors, the MCMC chains do not even converge. This suggests that the likelihood function is too difficult to explore and might contain discontinuities. As suggested by Gelman et al. (2020) in their *Bayesian Workflow* document, the first step to take to address this issue would be to drastically simplify the likelihood function and re-assess performance of the model of fake datasets. Another approach that we are currently exploring would be to use Approximate Bayesian Computation.

7 Implementation Plan



References

[1] Jiaquan Xu et al. Mortality in the United States. *NCHS Data Brief*, 2021.

[2] Kim S.Ka. and Cho S.W. The evasion mechanisms of cancer immunity and drug intervention in the tumor microenvironment. *Front Pharmacol.*, 13(868695), 2022.

[3] V. Schirrmacher. From chemotherapy to biological therapy: A review of novel concepts to reduce the side effects of systemic cancer treament (review). *Int J Oncol.*, 54(2), 2019.

[4] Alex D. Waldman, Jill M. Fritz, and Michael J. Lenardo. A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nature Reviews*, 20:651–69, 2020.

[5] I.L. Jr Bennet and P.B. Beeson. Studies on the pathogenesis of fever - characterization of fever-producing substances from polymorphonuclear leukocytes and from the fluid of sterile exudates. *J Exp Med.*, 98:493–508, 1953.

[6] R. Mortarini, A. Borri, G. Targni, et al. Peripheral burst of tumor-specific cytotoxic T lymphocytes and infiltration of metastatic lesions by memory CD8+ T cells in melanoma patients receiving interleukin 12. *Cancer Res.*, 60:3559–68, 2000.

[7] J.E. Portielje, C.H. Lamers, W.H. Kruit, A. Sparreboom, R.L. Bolhuis, G. Stoter, et al. Repeated administrations of interleukin (IL)-12 are associated with persistently elevated plasma levels of IL-10 and declining IFN-gamma, tumor necrosis factor-alpha, IL-6, and IL-8 responses. *Clin Cancer Res*, 9, 76-83.

[8] A. Mukhopadhyay, J. Wright, S. Shirley, D.A. Canton, C. Burkart, R.J. Connolly, et al. Characterization of abscopal effects of intratumoral electroporation-mediated IL-12 gene therapy. *Gene Ther.*, 26:1–15, 2019.

[9] B.V. Kumar, T.J. Connors, and D.L. Farber. Human T cell development, localization, and function throughout life. *Immunity*, 48:202–213, 2018.

[10] J.E. Portielje, C.H. Lamers, Kruit W.H., A. Sparreboom, R.L. Bolhuis, G. Stoter, et al. Repeated administration of interleukin IL-12 are associated with persistently elevated plasma levels of IL-10 and declining ifn-gamma, tumour necrosis factor-alpha, IL-6 and IL-8 responses. *Clin Cancer Res*, 9:76–83, 2003.

[11] L.K. Chen, B. Tourvieille, G.F. Burns, F.H. Bach, D. Mathieu-Mahul, M. Sasportes, and other. Interferon: a cytotoxic T lymphocyte differentiation signal. *Eur J Immunol*, 17, 767-70.

[12] Y. Hayakawa, K. Takeda, H. Yagita, M.J. Smyth, L. Van Kaern, K. Okumura, et al. IFN-gamma-mediated inhibition of tumor angiogenesis by natural killer T-cell ligand, alpha-galactosylceramide. *Blood*, 100, 2002.

[13] F.M. Rosa, M.M. Cochet, and M. Fellous. Interferon and major histocompatibility complex genes: a model to analyse eukaryotic gene regulation? *Interferon*, 7:47–87, 1986.

[14] S.A. Rosenberg, B.S. Packard, P.M. Aebersold, D. Solomon, S.L. Topalian, S.T. Toy, et al. Use of tumor-infiltrating lymphocytes and interleukin-2 in the immunotherapy of patients with metastatic melanoma. *N Engl J Med*, 319:1676–80, 1988.

[15] D.R. Leach, M.F. Krummel, and J.P. Allision. Enhancement of antitumor immunity by CTLA-4 blockade. *Science*, 271:1734–1736, 1996.

[16] E.D. Kwon et al. Manipulation of T cell costimulatory and inhibitory signals for immunotherapy of prostate cancer. *Proc. Natl Acad. Sci. USA*, 94:8099–8103, 1997.

[17] B.T. Fife and J. A Bluestone. Control of peripheral T-cell tolerance and autoimmunity via the CTLA-4 and PD-1 pathways. *Immunol. Rev*, 224:166–182, 2008.

[18] B.T. Fife and J.A. Bluesone. Control of peripheral T-cell tolerance and autoimmunity via the CTLA-4 and PD-1 pathways. *Immunol. Rev.*, 224, 166-182.

[19] H. Nishimura, N. Minato, T. Nakano, and T. Honjo. Immunological studies on PD-1 deficient mice: implication of PD-1 as a negative regulator for B cell responses. *Int Immunol*, 10:1563–72, 1998.

[20] A. van Elsas, A. A. Hurwitz, and J.P. Allison. Combination immunotherapy of B16 melanoma using anti-cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) and granulocyte/macrophage colony-stimulating factor (GM-CSF)-producing vaccines induces rejection of subcutaneous and metastatic tumors accompanied by autoimmune depigmentation. *J Exp Med*, 190:355–366, 1999.

[21] A. Mansurov, J. Ishihara, P. Hosseinchi, et al. Collagen-binding il-12 enhances tumour inflammation and drives the complete remission of established immunologically cold mouse tumours. *Nature Biomedical Engineering*, 4:531–543, 2020.

[22] Shuaishuai Xu, Huaxiang Xu, Wenquan Wang, et al. The role of collagen in cancer: from bench to bedside. *J Trans Med*, 17(309), 2019.

[23] Takuya Miyano. Mathematical modeling of complete remission of immunologically cold tumor by tumor-matrix targeted interleukin-12, 2019.

[24] Christian Hines. Cancer immunotherapy meeting notes, 2022.

[25] Christian Hines. Positive feedback between IL-12 and IFNg, 2022.

[26] Marie Davidian. *Nonlinear Mixed Effects Models*, pages 947–950. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[27] Baisen Liu, Liangliang Wang, and Jiguo Cao. Bayesian estimation of ordinary differential equation models when the likelihood has multiple local modes. *Monte Carlo Methods and Applications*, 24(2):117–127, 2018.

[28] Sophie Donnet and Adeline Samson. A review on estimation of stochastic differential equations for pharmacokinetic/pharmacodynamic models. *Advanced Drug Delivery Review*, 65, 2013.

[29] Emmanuelle Comets, Audrey Lavenu, and Marc Lavielle. Parameter estimation in nonlinear mixed effect models using saemix, an R implementation of the SAEM algorithm. *Journal of Statistical Software*, 80(3), 2017.

[30] Mary J. Lindstrom and Douglas M. Bates. Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46(3), 1990.

[31] B. Rosenbaum, M. Raats, et al. Estimating parameters from multiple time series of population dynamics using bayesian inference. *Frontiers in Ecology and Evolution*, 6(234), 2019.

[32] Andrew Gelman, John B Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. 3rd edition, 2021.

[33] Valderrama-Bahamóndez, Gloria I., and Holger Fröhlich. Mcmc techniques for parameter estimation of ode based models in systems biology. *Frontiers in Applied Mathematics and Statistics*, 5, 2019.

[34] M.A. Vazquez-Cruz, R. Guzman-Cruz, I.L. Lopez-Cruz, et al. Global sensitivity analysis by means of efast and sobol' methods and calibration of reduced state-variable tomgro model using genetic algorithms. *Computers and Electronics in Agriculture*, 100, 2014.

[35] D. Luengo, L. Martino, M. Bugallo, et al. A survey of monte carlo methods for parameter estimation. *EURASIP J. Adv. Signal Process*, 25, 2020.

[36] C.P. Robert and W. Changye. Markov chain monte carlo methods, a survey with some frequent misunderstandings, 2020.

[37] Motohide Nishio and Aisaku Arakawa. Performance of hamiltonian monte carlo and no-u-turn sampler for estimating genetic parameters and breeding values. *Genetics Selection Evolution*, 51(73), 2019.

[38] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[39] G.F. Fussmann, Ellner S.P., K.W. Shertze, and N.G. Hairston Jr. Crossing the Hopf bifurcation in a live predatory-prey system. *Science*, 290(5495):1358–60, 2000.

[40] Andrew Gelman et al. Bayesian Workflow, 2020.

[41] Stephen P. Brooks and Andrew Gelman. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4), 1992.