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

Supervisors: Professor Reiko Tanaka, Dr Tara Hameed
v2.4.5

Alexandre Yann Péré
CID: 01938104
February 2, 2024

Contents

1 Background

1.1 Cancer Immunotherapies

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 defences [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 [4]. Several viable strategies exist for immunotherapy [5]. We will first review a specific strategy, cytokine-based therapies, as this will enable us to introduce the CBD-IL-12 treatment in the next section, which is the main focus of the project.

Cytokine-based therapies rely on the injection of specific cytokines (small proteins that act as signalling molecules specifically during the immune response) to control tumour growth [6]. One of the most promising cytokine thus far is the interleukin-12 (IL-12), that was shown to have potent antitumour effects [7]. While it does not directly affect tumour cells, it mediates the production of other molecules [8], especially the cytokine interferon-γ (IFNγ). IFNγ has four main effects:

1. It stimulates the production of tumour-infiltrating cytotoxic cells, mainly $CD8^+$ [9][10]. These are a type of T-lymphocytes whose main function is to carry out cytotoxic activity (i.e. killing the malignant cells) after detecting cancerous cells [11].
2. It facilitates proliferation of $CD8^+$ T-cells by inhibiting expression of the Programmed Death 1 (PD1) membrane receptor within these cells. When not inhibited, these receptors, responsible for cell death, are overexpressed in tumour microenvironment and thus induce an exhaustion of the cytotoxic activity [12].
3. It slows down tumour growth by reducing angiogenesis [13] and upregulating antigen-presenting pathways within tumour cells [14].
4. IFNγ also in turn increases production of IL-12 by dendritic cells [15]. This positive feedback loop is called IFNγ priming [16][17].

While these four functions, illustrated in Fig 2a, indicate that IL-12 has potent antitumour effects (through IFNγ), clinical studies for IL-12-based treatments demonstrated that systemic injection of IL-12 is too toxic to be approved, as it triggers a generalised immune response throughout the whole body [18][19]. The severe treatment-related adverse effects called for additional research focusing on safer, more localised delivery method for IL-12.

1.2 A novel immunotherapy: CBD-IL-12

Recent endeavours in this field of immunotherapies led to the development by Ishihara group at Imperial College in 2019 [20] of a new molecule, CBD-IL-12, that demonstrated promising results to treat melanoma. We first review how CBD-IL-12 works, and then we detail the experimental results obtained by the lab.

The CBD-IL-12 molecule consists of a collagen-binding protein (or collagen-binding domain, CBD) that is fused onto an IL-12 cytokine. The modified interleukin hence mainly accumulates in collagen-rich regions.

As collagen is the main component of cancerous microenvironment [21], this effectively results in localised delivery method that can achieve high concentration of IL-12 specifically in cancerous microenvironments compared to the rest of the body.

To measure treatment efficacy on cancer, a common metric is the complete response (CR) rate among patients [22]. CR corresponds to the disappearance of all known lesions [23] in the long run (steady-state). Since the optimal treatment regimen for CBD-IL-12 is not known, the authors measured the CR-rate in various settings: injection on day 7, or on day 9, in combination with other drugs or not, etc. For this project, we focus on the specific case of CBD-IL-12 monotherapy (no other drugs are being used) injected on day 7 to treat skin cancer (melanoma) since this is the setting for which we have the most comprehensive set of data. For this treatment protocol, nine mice were inoculated a skin tumour on day 0. Tumour volume was then recorded at specific time points until day 27. Fig. 1 plots these tumour volume evolution for each individual mouse. Tumour is successfully inhibited in all mice at different rates, up to day 20. At this point, two mice exhibited a resurgence of tumour growth, while for the remaining ones the tumour stayed inhibited (below 1mm^3) until the end of the experiments (day 27). We can conclude that, for this period of time at least, seven out of the nine mice experienced CR, resulting in a CR-rate of 77%.

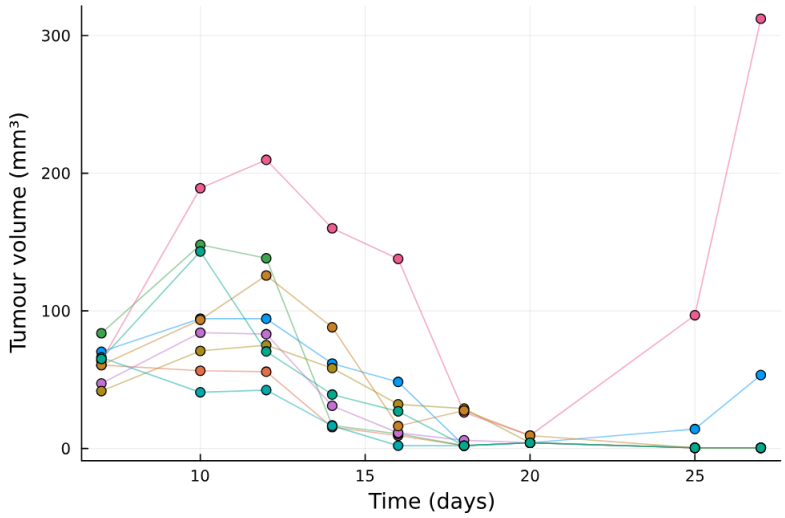


Figure 1: Evolution of tumour volume over time for a batch of mice shows that they present two distinct behaviours: CR or non-CR (each trace is an individual mouse)

While these are promising results compared to alternative treatments for melanoma [24], the heterogeneity of response (CR vs non-CR) could not be explained by the authors, and highlights the need to further analyse available data in order to identify the key biological parameters that can differentiate between CR and non-CR.

1.3 Data Analysis with the Aid of Computational Modelling

To have a better understanding of the immune response to CBD-IL-12, Dr Miyano, a previous member of the Tanaka group [25], 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 while minimising the number of animal tests. He developed a

mechanistic model (Fig. 2b) based on our current model of the immune processes outlined in Section 1.1 (see Fig. 2a). This mechanistic model can predict tumour growth given a treatment, along with the concentration of IFNγ, CBD$^+$ and PD-1. The meaning of each model parameter is shown in Appendix A. It must be noted that this mechanistic model was developed to accommodate for different treatment regimen, including combination therapy (CBD-IL-12 with checkpoint inhibitor drugs, called CPI). This is indicated by the two green boxes on Fig. 2b, which indicates the presence of their respective drug. Since we are only concerned with monotherapy in this project, the CPI box can safely be ignored.

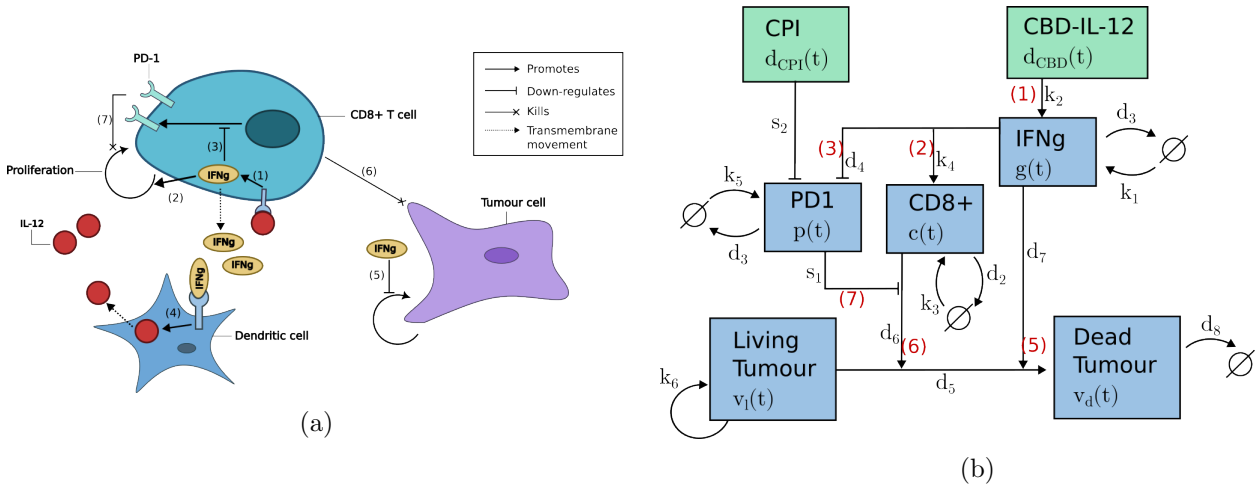


Figure 2: Mechanistic model of the IL-12-enhanced immune response. Numbers in the figures are to identify corresponding processes.
(a) Illustration of the action mechanisms of IL-12-based cytokine (such as CBD-IL-12). Processes included in the model are: (1) IL-12-induced production of IFNγ by T-cells, (2) upregulation of CD8$^+$ T-cells proliferation by IFNγ, (3) down-regulation of the PD-1 immunosuppressive pathway, (4) IFNγ-triggered production of IL-12 by dendritic cells, (5) reduced proliferation of tumour cells through suppressed angiogenesis and (6) CD8$^+$-mediated cytotoxic activity.
(b) Diagram of the molecular interactions in the tumour microenvironment. Green boxes represent the inputs (drugs), which are fixed since we consider a specific treatment regimen (see Section 1.2). Blue boxes represent the five state variables. All kinetic rates, corresponding to model parameters, are shown next to the corresponding process. Note that the IFNγ-priming process is missing, as the current mechanistic model does not include the feedback loop.

This model can be represented by the following set of Delay-Differential Equations (DDEs) [25] (see

Fig. 2b and Appendix A for the explanation of each symbol):

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

The parameters of this mechanistic model all represent biological factors that could potentially be responsible for the diverging treatment outcomes. However, these parameters are kinetic rates that cannot be measured experimentally: they can only be estimated from observable data (which is tumour volume in our project): this is called parameter estimation. However, when performing parameter estimation on multiple time series, one must choose the pooling type, which relates the different time series together. In the context of this project, where each time series correspond to an individual mouse, no-pooling means that each mouse is considered independently. Parameter estimation is thus performed on each mouse separately. On the other hand, complete-pooling assumes that the parameters are completely homogeneous across the experiments (i.e., all time series are characterised by the same parameter combination). It is performed by averaging all time series together, resulting in one unique average behaviour.

Hines, another previous member of the Tanaka group, applied parameter estimation on the mechanistic model by assuming complete-pooling. He obtained a set of parameters that can parameterise the mechanistic model so that it reproduces the average growth curve. However, the shortcoming of the complete-pooling approach is that it completely ignores the heterogeneity between individual and assumes that all individual behaves the same. It hence cannot to explain the heterogeneity of outcomes. This motivates us to use alternative pooling methods that could help us identify the key biological factors that differentiate between CR and non-CR.

Additionally, Hines noted that a key element that could explain why some mice exhibited continued immune response tens of days after the treatment was injected, thus achieving CR, could be a positive feedback loop that causes sustained production of IL-12, even days after injection. This is motivated by findings that positive feedback loops in cellular signalling systems can lead to switch-like behaviour [26]. In such systems, a short-lived input can have a long-lasting impact on the steady state, even when it is removed. While a positive feedback loop is not present in the current mechanistic model, Hines identified a biological process (IFNγ priming, see 1.1) corresponding to a positive feedback loop that could potentially be added to the model.

1.4 Proposed Approach: Hierarchical Bayesian Modelling

An alternative parameter estimation technique that we propose to use is Bayesian modelling, which comes with two advantages crucial to solve the issue outlined by Hines:

1. Unlike the parameter estimation used by Hines, Bayesian approach do not provide point estimate for each parameter, but rather probability distributions. Distributions, compared to point estimate,

contain additional information that could help us understand which mechanisms drive the immune response and the complete response rate. One possible way distributions can be useful is outlined in [27]: they allow us, in the case of parameter estimation, to quantify the sensitivity of the model to each parameter. For a given distribution associated to a parameter, a narrow credible interval (CI) means that the model is sensitive to this parameter. The opposite is true for a wide credible interval.

2. Bayesian modelling allow us to perform partial-pooling, which is a combination of no-pooling and complete-pooling. It formulates that, for a set of time series, each time series (and their corresponding estimated parameter set) are not completely independent nor homogeneous from each other, but rather contain some shared information and some individual variability. This is called hierarchical Bayesian modelling.

2 Aims and Objectives

The aim of the project is to identify the key biological parameters responsible for the different outcomes observed in the lab (CR vs non-CR)

Objective 1: develop a hierarchical Bayesian model based the mechanistic model that can reproduce the experimental data.
Objective 2: compute the probability distributions of the model parameters and verify whether they are bimodally distributed (to explain the two outcomes).
Objective 3: evaluate the impact of the IFNγ priming pathway on the treatment outcome.

3 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 University of Chicago (see Methods section of [20]).

The data derived from these experiments holds 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. As this concerns development of drugs for human use, it is necessary to develop a robust work ethics that avoid developing harmful therapies to human. Thus, to ensure that the results are reliable and can be reproduced by anyone, the full analysis along with the code will be published on GitHub.

4 Literature Review on Bayesian Parameter Estimation

As mentioned in Section 1, the project requires a parameter estimation method that can be applied to mixed-effects models. In this section, we first give a more rigorous definition of the parameter estimation along with an overview of the different methods available. We then review a specific method, called hierarchical Bayesian modelling, and highlight its particular relevance to the project.

4.1 Definition and Notation

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 treatment 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 $\boldsymbol{\theta}_j$ that can reproduce the observed data $\boldsymbol{y}_j$.

4.2 Bayesian Parameter Estimation

Bayesian Parameter Estimation is a method to estimate a parameter vector $\boldsymbol{\theta}$ given an observation vector $\boldsymbol{y}$. 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 [28]:

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

This formula is the direct application of Bayes' theorem. It is the product of the prior distribution $p(\boldsymbol{\theta})$, which represents our knowledge 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 in situations where data about different individuals was collected.

4.3 Hierarchical Modelling

In this project we seek to estimate a parameter vector $\boldsymbol{\theta}_j$ for each individual mouse j. However, as mentioned before, we model the data through partial-pooling: the parameter vectors $\boldsymbol{\theta}_j$ are connected in some way to each other, since they are all drawn from the same mice population. They way they are connected is as follows: in hierarchical Bayesian modelling, we can formulate that each $\boldsymbol{\theta}_j$ is sampled from a population-level distribution characterised by the (also unknown) hyperparameters $\boldsymbol{\phi}$. This common distribution corresponds to the shared information between each mouse, and the fact that each $\boldsymbol{\theta}_j$ is randomly sampled accounts for the individual variability. 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 [28]:

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

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

4.3.1 Hierarchical Priors and Hyperpriors

In hierarchical Bayesian models, there are two types of parameters: hyperparameters $\boldsymbol{\phi}$ and individual parameters $\boldsymbol{\theta}_j$ [28]. 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 convey 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 [29], 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:

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

4.3.2 Likelihood Function

The likelihood describes the probability density of observing the data as a function of the parameters. In this case, it not only captures the dynamics of the mechanistic model, but also the observational noise. Typically, the mean of the observations can be modelled as the solution of the DDE (parameterised by $\boldsymbol{\theta}_j$) and we can assume a white Gaussian observation error (with standard deviation σ_{err}) [30][31]:

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

where t is the time and $Y_j(t|\boldsymbol{\theta}_j)$ is the solution of the DDE model parameterised by $\boldsymbol{\theta}_j$ (usually obtained by numerical methods). Without additional information about the measurement methods, this is the approach suggested by Rosenbaum et al [29].

It was also shown that, when the parameters to be estimated differ in value by several order of magnitude, fitting them on a logarithmic scale makes the inference more robust: faster convergence and more accurate posteriors [29]. This changes the likelihood function to the following:

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

4.4 Computational Methods

While the previous section lays the theoretical background necessary to understand hierarchical Bayesian modelling, several issues can arise during practical implementation. In this section, we discuss two potential issues along with common solutions found in the literature.

4.4.1 Numerical Approximation of Posterior Distributions

The posterior distribution $p(\boldsymbol{\theta}_j, \boldsymbol{\phi}|\boldsymbol{y}_j)$ represents the updated joint probability distribution for each parameter and hyperparameter, obtained by updating the priors with the likelihood function through the

Bayes theorem. Ideally, the likelihood function can be written in closed-form, so that the posterior can be evaluated analytically as the normalised product between the priors and the likelihood (see Sec. 4.3).

However, in many cases the likelihood can only be evaluated for given a set of parameters, but its analytical form cannot be written. This can arise, as in this project, when it relies on a set of differential equations that can only be solved numerically. In these cases, a common approach is to use Markov Chain Monte-Carlo (MCMC) [32]. MCMC is a method used to estimate a probability distribution when we only know how to calculate the probability density function for a given sample of the parameter space. In Bayesian modelling, it approximates the posterior by iteratively sampling random sets of parameters and calculating their corresponding posterior probability. This set of samples is called a MCMC chain. As its length grows to infinity, the distribution approximated through MCMC converges toward to true distribution.

Several metric exist to measure convergence of the MCMC chains, a key characteristic as it reflects the quality of the approximation of the distribution [33] (in this case, of the posterior distribution). The most popular method to assess convergence is the potential scale factor reduction, usually termed $\hat{R}$, developed by Gelman et al., 1992 [34]. 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.

4.4.2 Reduction of the Parameter Space through Sensitivity Analysis

As Bayesian parameter estimation is computationally intensive [35], this approach can benefit, in terms of computation time, from reducing the number of parameters when possible. Varella et al, 2012 [36], proposed a method to reduce the number of parameters in ODE models through sensitivity analysis. By taking the example of a crop growth model (the STICS model) with 13 parameters, the authors used an eFAST sensitivity analysis to determine the total sensitivity index ST for each parameter. This index indicates the fraction of the variance of model output variable that is imputable to a given parameter and its interaction with other parameters. For a parameter vector $\boldsymbol{\theta}$, the ST index of the i-th parameter can be computed as follows:

$$ST_i = \frac{V(Y) - \sum_{j\neq i} V_j}{V(Y)},$$

where Y is the model output variable, $V(Y)$ is its total variance and V_j is the variance of Y when only $\boldsymbol{\theta}_j$ is varied. The authors proposed to use this metric to classify parameters into two categories: if ST>0.1, the corresponding parameter should be retained and considered as a free parameter. Otherwise, the parameter can be fixed to a nominal value, found in the literature for example.

Vazquez-Cruz et al, 2014 [37], applied this method to another crop model called TOMGRO. After reducing the number of parameters from 17 to 8 (using the screening technique explained above), they showed that

the reduced model was still able to reproduce experimental data with level of error comparable to the full model.

5 Risk Register

The risks associated with the project along with a mitigation plan are described in Table 1

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 an 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 [38]. 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 mentioned in Section 1.3.

6 Evaluation

Objective 1: Bayesian Model

The first objective of the project is to design a hierarchical Bayesian model. Once it is built, we must first verify its validity before proceeding to the next objective. We will follow the validation procedure outlined in [38]. It consists of three tests that must each be passed:

- **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, otherwise the test is not passed.
- **Fake Data Check:** to perform this test, we first need to generate an artificial dataset using known values of parameters, and then fit the Bayesian model to these fake datasets. If the 95% credible interval of the posterior distributions each include the corresponding true parameter value, then the model passes the test. Additionally, to ensure that the posterior distributions are reliable, we must verify that the MCMC chains converge (otherwise, it means that we cannot make inference from the posteriors). Their corresponding $\hat{R}$ must be below 1.05 (see Sec. 4.4.1)
- **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 criterion for the model to pass the test is that the 95% credible interval should contain the experimental time series obtained in the lab.

Objective 2: Bimodal Distributions for the Parameters

Once the Bayesian model has been validated, we can use it to compute the posterior distributions for each parameter. To identify the key parameters that can differentiate between CR and non-CR, along with their threshold value that separate each outcome, we propose the following method: (1) We label each experimental time series as either CR or non-CR (depending on whether tumour volume converged to 0 or diverged at the end of the experiment). We can then construct two distinct datasets: one containing only CR time series, and one with only non-CR time series. (2) We fit the model to each dataset separately. We hence obtained two sets of posterior distributions, each corresponding to a population that either achieved CR or non-CR, respectively. For any parameter, if the CR-population has a statistically different hypermean from the non-CR-population, we can conclude that this parameter is differently distributed between mice that achieved CR and those who did not. (3) By identify all parameters that are differently distributed between the two populations, we can construct a set of biological parameters that help us differentiate between CR and non-CR (we call this set the responder profile). If this set is non-empty, it means that the project was successful, and subsequent animal experiments can potentially be undertaken to verify these results.

Objective 3: Impact of IFNγ priming

To evaluate the impact of the IFNγ priming positive feedback loop, we will first modify the Bayesian to include it. Once it has been validated following the same procedure as highlighted above (see Sec. 6 for Objective 1), we will construct a new responder profile and compare it with the previously obtained profile (that does not include the feedback loop).

7 Preliminary Results

7.1 Verifying the Dynamics of the Mechanistic Model

Before extending the mechanistic model to a Bayesian model (objective 1), the very first aspect of the mechanistic model that we wanted to verify was its ability to capture two specific treatment outcomes: CR vs non-CR. This is a necessary feature without which we would not be able to differentiate between the different outcomes observed in the lab. 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). 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. The choice of this method was motivated by the fact that this method is fast to implement compared to finding the fixed points numerically. Its shortcoming is that it can only find an approximation of the bifurcation boundary, but this does not matter for this step since the goal was to simply to check the existence of a clear boundary rather that its exact location. To avoid sampling in a 21-dimensional parameter space, which would be too computationally heavy, we used Hines' findings, according to which the solution of the DDE model (i.e. the simulated tumour growth) is mostly sensitive to three parameters: k_6, s_2 and d_1 [39]. Fig. 3 shows the results of this grid-search analysis, where each axis of the cube corresponds to the value of the parameter k_6, s_2 or d_1 respectively. Green dots represent a parameter combination which lead to CR (tumour volume below 10mm^3 at day 27), dots were coloured in red otherwise. As can be seen, there seem to 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 knowing on which side of the boundary he is.

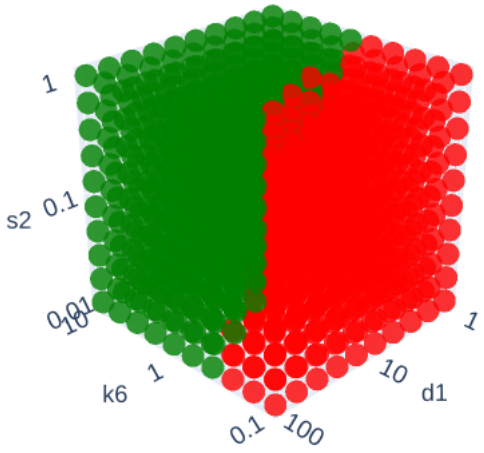


Figure 3: Grid-search stability analysis. Each dot represents a combination of parameters k_6, d_1 and s_2. Green indicates a combination that led to CR, red indicates non-CR. This shows that there is a clear boundary in parameter space between CR and non-CR

7.2 Construction and Validation of the Bayesian Model (Objective 1)

In this section we show how our initial Bayesian model was tested, following the procedure highlighted in Section 6. We focus on a reduced model with only three free parameters, k_6, d_1, s_2, as justified above (see Sec. 7.1). All other parameters of the model were fixed to their respective value estimated by C. Hines through complete-pooling parameter estimation. This ensures that they do not bias the model towards an extreme outcome (whether CR or non-CR), since they characterise an "average" growth curve.

7.2.1 Prior Predictive Check

Note: to represent a distribution truncated at 0, we use either a "+" (positive half) or "-" (negative half) superscript. For example: $X \sim \text{Cauchy}^- \Rightarrow -\infty < X < 0$.

In this case, we do not have much data on the typical values of the parameters since it is impossible to measure it, so we aimed to design an uninformative prior. Fig. 4 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 distributions:

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

We chose Cauchy distributions since they are less informative than standard Gaussian distributions, allowing for values far from their centre of mass. This is critical since we do not have much information

about the true values of the parameters. In our priors, we assume that the parameters are log-Cauchy distributed to ensure that they are strictly positive, a necessary condition since they are biological parameters. Additionally, the Cauchy distributions were truncated to ensure that the parameters are either below 1 (for k_6 and s_2) or above 1 (for d_1). This is to avoid numerical instability. [The threshold of 1 was chosen experimentally and needs to be refined more rigorously.]

Fig. 4, is the result of the prior predictive check. Each grey line represents a solution of the DDE model characterised by a random parameter combination sampled from the priors. The blue shade represents the 95% credible interval, and the red line is the median growth curve. As we can see, the 95% credible interval spans our expected range of curves (any curve between no growth and a continuous growth up to 600mm^3), meaning that they are relatively uninformative priors. The median curve has the shape of the typical growth curve, as observed in the labs. Since the priors do not bias the model towards a CR or non-CR, but rather cover the full range of growth curve we might expect, we conclude that the prior predictive check is passed.

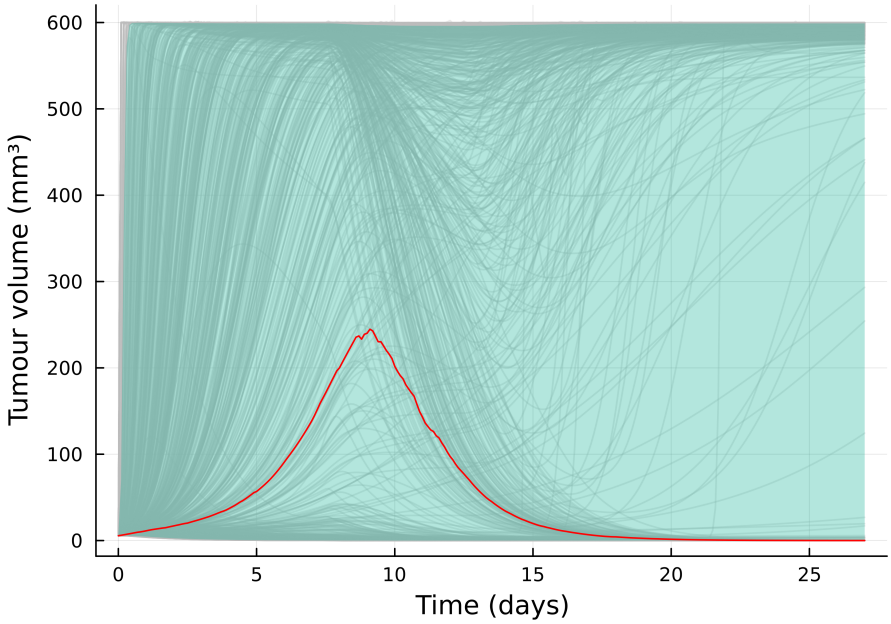


Figure 4: DDE solutions (grey lines) for 1,000 parameter vectors $\boldsymbol{\theta}$ sampled from the priors. The 95% credible interval for the tumour volume at each time step (blue shade) shows that the model passes the prior predictive check: it contains our expected range of curves. The median growth curve (red line) confirms this, as it is very similarly to the growth curve observed in the lab.

7.2.2 Fake Data Check

For the fake data check, we first generate artificial datasets using known parameter values, and then fit the Bayesian model to these fake datasets. Comparing the estimated parameter value in the form of the posterior distribution to the true values can help us make informed decisions about modifying the model before proceeding to fit it to real data. While a successful fake data check (we can recover the true parameter values using the posteriors) does not mean the inference on real data will be successful or

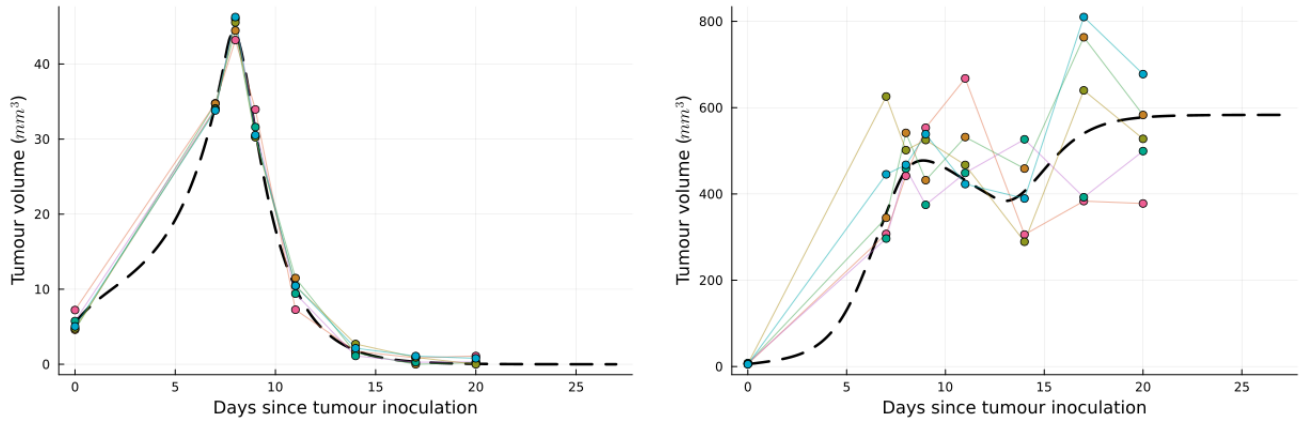
biologically meaningful, a failed fake data check implies that the model will never be able to fit on real data. In this case, the way the model fails can guide us when improving the model.

Data Generation

Each fake growth curve was generated by sampling a value of $\boldsymbol{\theta}$ from the prior distributions (shown in the previous subsection), 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 expect from the labs. This was done in two different ways, resulting in two distinct datasets. For dataset A, we simply added a white Gaussian noise with variance $\sigma_{\text{err}} = 1$ to the simulation (additive noise). The variance was chosen arbitrarily, the only condition being that it is one order of magnitude smaller than tumour volume. For dataset B, we 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 dataset A. The generation process is summarised 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. 5 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 5: Fake data sets (scatter plot and coloured lines) generated from a baseline growth curve (dashed line), either with **(a)** additive noise or **(b)** multiplicative noise. Baseline curves were obtained by solving the DDE model parameterised by a vector $\boldsymbol{\theta}$ sampled from the prior distributions.

Fitting the model to the fake datasets

Before checking if the estimated values match the true ones, we first assess convergence of the MCMC

chains. The $\hat{R}$ values (see Section 6) are reported in Table. 4 as the average $\hat{R}$ value across 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 the "Number of Chains" column only repots the number of chains used to compute $\hat{R}$. If all chains are excluded, meaning that none of them converged, we simply report N/A.

Table 4: Assessment of convergence for the MCMC chains for uninformative priors

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

Looking at Table 4, we can conclude that the chains did not converge, meaning that the model cannot make inference with $\boldsymbol{\theta} \in \mathbb{R}^3$ and uninformative priors. To further diagnose the model, we performed 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, used to generate the fake datasets. 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 centred on the true parameter values.

Table 5: Assessment of convergence for the MCMC chains for informative Cauchy priors

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 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. We can observe that chains converged or were close to convergence only for dataset B, showing that a log-normal transformation is key to make exploration of the posterior easier to perform.

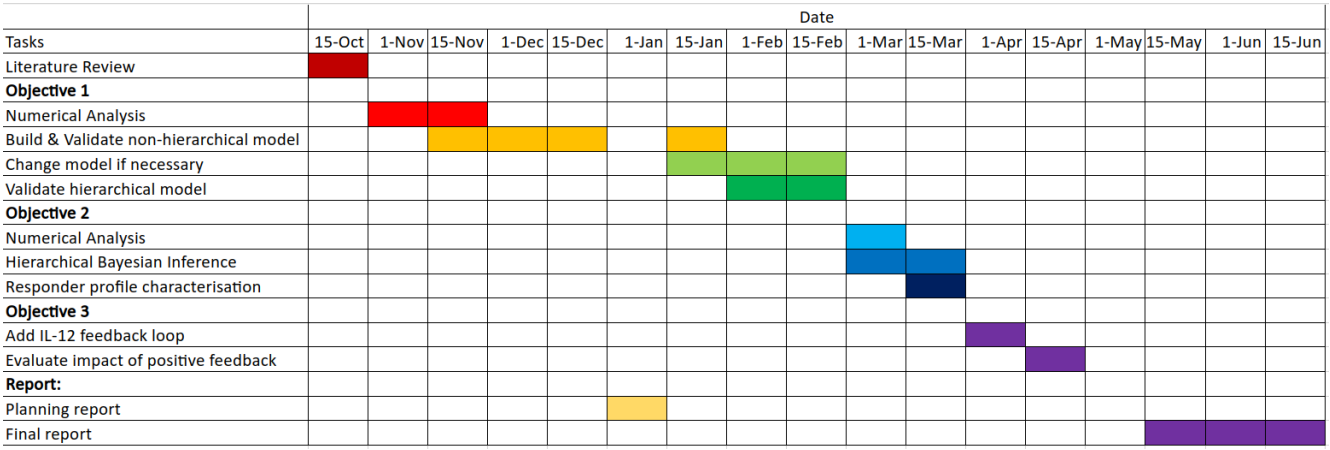
Table 6: Assessment of convergence for the MCMC chains for informative Normal priors

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

Conclusion

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. These two approaches will enable us to further diagnose the model to make appropriate corrections.

8 Implementation Plan



A Parameters of the Computational Model

Table 7: Parameters of Miyano's model along with their description

Parameter	Description
t_{delay}	time delay from injection timing to deliver CBD-IL-12 to tumour
t_{last}	time duration of drug effects of CBD-IL-12
$t_{delay12}$	time delay from injection timing to deliver IL-12 to tumour
t_{last12}	time duration of drug effects of IL-12
t_d	time delay for producing $CD8^+$ via IFNγ
k_1	production rate of IFNγ via unspecified pathways
k_2	production rate of IFNγ via CBD-IL-12
k_3	production rate of $CD8^+$ via unspecified pathways
k_4	production rate of $CD8^+$ via IFNγ
k_5	production rate of PD-1 via unspecified pathways
k_6	proliferation rate of tumour
d_1	elimination rate of IFNγ via turnover
d_2	elimination rate of $CD8^+$ via turnover
d_3	elimination rate of PD-1 via turnover
d_4	elimination rate of PD-1 via turnover IFNγ
d_5	elimination rate of living tumour via turnover
d_6	elimination rate of living tumour via $CD8^+$
d_7	elimination rate of living tumour via IFNγ
d_8	elimination rate of dead tumour via turnover
s_1	inhibition strength of PD-1 on anti-tumour effects of CDB^+
s_2	inhibition strength of tumour volume on antitumour effects of IFNγ and $CD8^+$ cells

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] A. Sosa, Cadena E. Lopez, Olive C. Simon, Karachaliou N., and Rosell R. Clinical assessment of immune-related adverse events. *Ther Adv Med Oncol*, 10, 2018.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] Jason K. Whitmire, Joyce T. Tan, and J. Lindsay Whitton. Interferong acts directly on CD8+ T cells to increase their abundance during virus infection . *Journal of Experimental Medicine*, 201(7):1053–1059, 04 2005.

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

[12] Y. iang, Y. Li, and B. Zhu. T-cell exhaustion in the tumor microenvironment. *Cell Death Dis*, 6, 2015.

[13] 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.

[14] 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.

[15] I.M. Wang, C. Contursi, A. Masumi, X. Ma, G. Trinchieri, and Ozato K. An IFN-gamma-inducible transcription factor, ifn consensus sequence binding protein (ICSBP), stimulates IL-12 p40 expression in macrophages. *J Immunol*, 165:271–279, 2000.

[16] J. Liu, S. Cao, LM. Herman, and X Ma. Differential regulation of interleukin (IL)-12 p35 and p40 gene expression and interferon (IFN)-gamma-primed IL-12 production by IFN regulatory factor 1. *J Exp Med*, 198:1265–1276, 2003.

[17] X. Ma, W. Yan, Q. Zheng, H. nd Du, L. Zhang, Y. Ban, et al. Regulation of IL-10 and IL-12 production and function in macrophages and dendritic cells. *F1000Res*, 4, 2015.

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

[19] Z. Jia, D. Ragoonanan, KM. Mahadeo, J. Gill, R. Gorlick, E. Shpal, and S Li. IL12 immune therapy clinical trial review: Novel strategies for avoiding CRS-associated cytokines. *Front Immunol.*, 13, 2022.

[20] 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.

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

[22] N.F. Aykan and T. Özatlı. Objective response rate assessment in oncology: Current situation and future expectations. *World J Clin Oncol*, 11(2), 2020.

[23] L.C Villaruz and M.A. Socinski. The clinical viewpoint: definitions, limitations of recist, practical considerations of measurement. *Clin Cancer Res*, 19(10), 2013.

[24] A. Knight, L. Karapetyan, and J.M. Kirkwood. Immunotherapy in melanoma: Recent advances and future directions. *Cancers*, 15(4), 2023.

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

[26] Ferrell JE Jr. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. *Curr Opin Cell Biol.*, 14(2), 2002.

[27] Tina Toni. *Approximate Bayesian computation for parameter inference and model selection in systems biology*. PhD thesis, Imperial College London, 2010.

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

[29] 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.

[30] 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.

[31] 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.

[32] Don van Ravenzwaaij, Pete Cassey, and Scott D. Brown. A simple introduction to Markov Chain Monte–Carlo sampling. *Psychon Bull Rev*, 25:134–154, 2016.

[33] Théo Moins, Julyan Arbel, Anne Dutfoy, and Stéphane Girard. On the use of a local $\hat{R}$ to improve mcmc convergence diagnostic, 2023.

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

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

[36] Hubert Varella, Samuel Buis, Marie Launay, and Martine Guérif. Global sensitivity analysis for choosing the main soil parameters of a crop model to be determined. *Agricultural Sciences*, 3(7), 2012.

[37] 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.

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

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