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

Supervisors: Professor Reiko Tanaka, Dr Tara Hameed
v2.4.2

Alexandre Yann Péré
CID: 01938104
January 16, 2024

Word Count: 4256

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 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 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 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 [7], especially the cytokine interferon-γ (IFNγ). IFNγ has four main effects:

1. First of all, it stimulates the production of tumour-infiltrating cytotoxic cells, mainly $CD8^+$ [8][9]. 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 [10].
2. Besides stimulating production of T-cells, it facilitates their proliferation (including $CD8^+$) by reducing negative regulatory pathways that lead to immunosuppression [11]. This involves the reduction of PD-1 concentration in $CD8^+$ cells [12], a membrane receptor that leads to apoptosis (programmed cell death).
3. It slows down tumour growth by reducing angiogenesis [13] and upregulating antigen-presenting pathways within tumour cells [14].
4. It was also discovered by Wang et al., 2000 [15], that IFNγ also in turn increases production of IL-12. This positive feedback loop is called IFNγ priming ([16][17]).

While these four pathways, illustrated in Fig 2.a, indicate that IL-12 has a very robust antitumour effect (though IFNγ), clinical studies demonstrated that systemic injection of IL-12 is exceedingly toxic as it triggers a large immune response throughout the whole body [18][19]. These severe treatment-related adverse effects dampened research about IL-12, waiting for a safer, more localised delivery method to be found.

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 a 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 an enhanced delivery method that can achieve high concentration of IL-12 specifically in cancerous microenvironments, which helps decreasing the toxicity of the treatment.

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 delivery method 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, 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%.

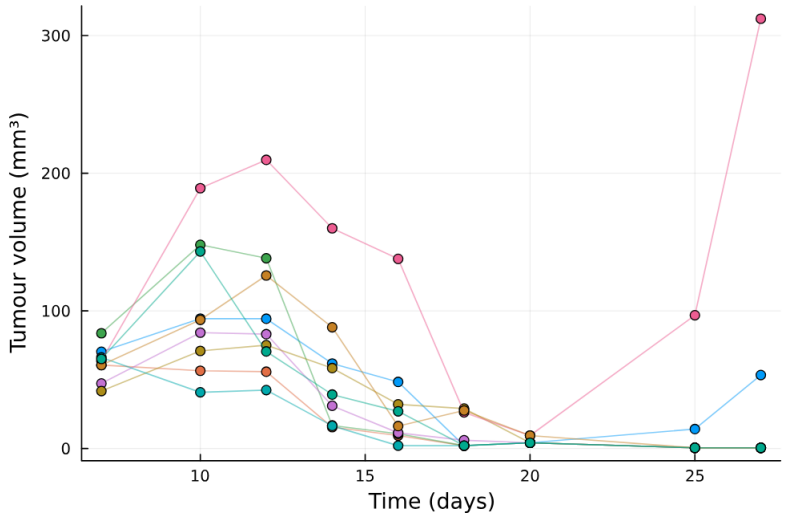


Figure 1: Evolution of tumour volume over time for a batch of mice shows that they present two distinct behavours: CR or non-CR (each trace is an individual mouse)

While these are promising results, 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 help differentiate the diverging outcomes.

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 [24], 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 without having to resort to extensive animal tests. He developped 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.

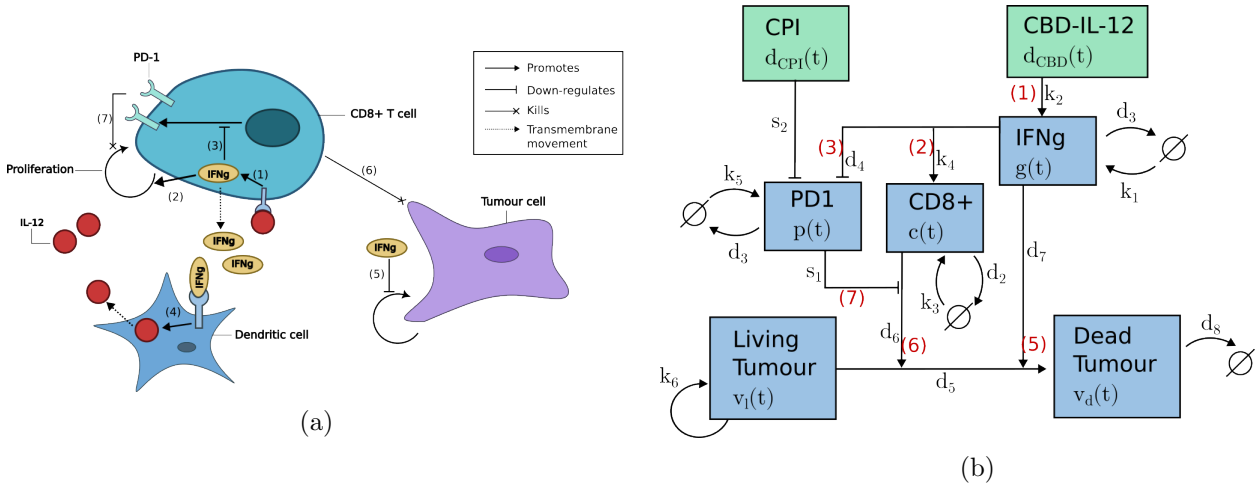


Figure 2: Mechanistic model of the IL-12-enhanced immune response. Numbers in the figures are to identify corresponding processes.
(a) Illustration of the action mechanisms of IL-12-based cytokine (such as CBD-IL-12). Processus included in the model are: (1) IL-12-induced production of IFNγ by T-cells, (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 input variables (drugs) and 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 equations [24]:

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

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, i.e. tumour volume. This process, called parameter estimation, was investigated and applied to the mechanistic model by Hines,

another previous member of the Tanaka group. However, naive parameter estimation can only work to extract parameter from one single time series, and hence cannot be directly applied to the collection of times series obtained by the lab. Hines' approach was to fit the mechanistic model to the average growth curve. This resulted in a set of parameters that can explain (i.e., reproduce) this average behaviour. However, Hines concluded in his report that this set of parameter was not sufficient to help differentiating the different drug responses, as it can only reproduce an average curve that in essence represents neither a CR outcome nor a non-CR outcome. Additionally, he also mentioned that a key biological process, the IFNγ-priming feedback loop, was missing from the mechanistic model, which could potentially impact its accuracy and reliability.

Hines' conclusion motivates us to use alternative parameter estimation methods that could help us identify the key biological factors that differentiate between CR and non-CR, by taking into account the whole dataset obtained in the lab. To this end, we first note that there is some shared information between the mice, since they are all sampled from the same population, as well as some variability between that explain the range of output obtained in the lab. Nonlinear mixed-effects models [25] are a standard method to analyse population dataset with both shared information (called fixed-effects) and individual variability (called random-effects). This is further detailed in Section 4.

2 Aims and Objectives

The aim of this project is verify whether Takuya's mechanistic model can be extended to identify the biological factors that can differentiate between CR and non-CR.

Objective 1: develop a computational mixed-effect 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 Univerity 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 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 relavance 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 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 $\boldsymbol{\theta}_j$ that can reproduce the observed data $\boldsymbol{y}_j$.

Many statistical approaches have been developed to perform parameter estimation on differential equation models from noisy data [26]. However, most of them cannot be applied to mixed-effect models. In the context of pharmacodynamics model , Donnet et al, 2013 [27] reviewed different techniques available to perform parameter estimation. For analysis of population data with observational noise, only two methods are available: Expectation-Maximisation (either Stochastic [28] or using First-Order Condition [29]) and Bayesian parameter estimation [30]. The authors concluded that Bayesian modelling in particular is the most flexible method, since it does not rely on assumptions and hence work for both individual 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 First-Order Conditon method was not proven to always 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.

4.2 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 [31]:

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

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.

4.3 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 [27][30]. For example, Rosenbaum et al., 2019 [30], 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 [31]:

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

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

4.3.1 Hierarchical Priors and Hyperpriors

In hierarchical Bayesian models, there are two types of parameters: hyperparameters $\boldsymbol{\phi}$ and individual parameters $\boldsymbol{\theta}_j$ [31]. 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 [30], 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:

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

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

4.3.2 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 [26][32] (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 [30].

4.4 Reduction of the Computational Burden

4.4.1 Sensitivity Analysis

As pharmacodynamics models can generally contain a high number of parameters, the Bayesian approach, which is computationally intensive [27], can result in intractable computations. To this end, Vasquez-Cruz et al, 2012 [33], 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. All the other parameters with a negligible influence were seet to a fixed value. Then, using a parameter fitting algorithm, they were able to show that this reduced model could still replicate the experimental data with minimal error.

4.4.2 Transforms

Another finding from the study of predator-prey systems [30] 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:

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

5 Risk Register

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

6 Evaluation

Objective 1: Bayesian Model

The first objective of the project is to design a hierarchical Bayesian model. To validate it and proceed to the next objective, we will follow the validation procedure outlined in [34]. It consists of three different 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.

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 [34]. 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 1.3.

- **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. 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). To evaluate their convergence, we use the potential scale factor reduction, usually termed $\hat{R}$, developed by Gelman et al., 1992 [35]. For a group of MCMC chains, a $\hat{R}$ below 1.05 indicates good convergence.
- **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 obtined in the lab.

Objective 2: Responder Profile

For objective 2, through the validated bayesian model we will (1) identify the parameters that can differentiate responders from non-responders; (2) define the range of values of each of the parameters to differentiate CR from partial responders and non responders. To do so, we will compute the posterior distribution of the hypermean for two datasets: one with only CR data, and one with only non-CR data. If the mean of the posterior distributions are distinct (they are not within the 95% credible interval of each other), we can conclude that the Bayesian version of the mechanistic model can successfully help us differentiating between CR and non-CR by identifying the key biological factors. Potential future work would be to identify one or more biomarkers, e.g., PD-1, whose presence and/or its concentrations at given timepoints post intervention can reflect the responder profile parameters, hence has predictive value for future studies.

Objective 3: Impact of IFNγ priming

To evaluate the impact of the IFNγ priming feedback loop, we will first modify the Bayesian model to

include it. Once it has been validated following the same procedure as highlighted above, we will construct a new responder profile and compare it with the previously obtained profile (that does not include the feedback loop). If in each case different key parameters were identified, this means that IFNγ cannot be neglected and needs to be included in the model.

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 outcome: CR vs non-CR. The is a necessary feature whithout which we would not be enable 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. To avoid sampling in a 21-dimensional parameter space, which would be too computationally heavy, we used Hines' findings, according to which the DDE model is mostly sensitive to three parameters: k_6, s_2 and d_1 [36]. Fig. 3 shows the resuls of this analysis, where each axis of the cube corresponds to the value of the parameter 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.

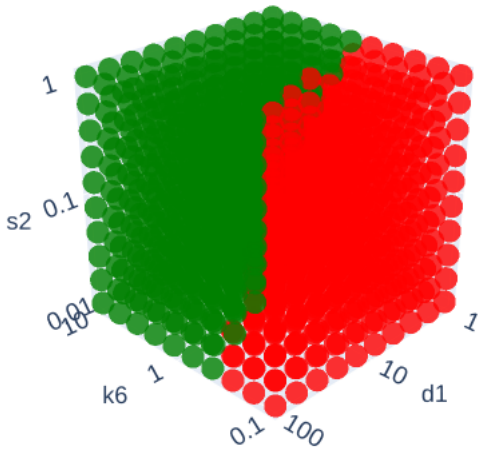


Figure 3: Stability analysis shows that there is a clear boundary in parameter space between CR and non-CR

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

In this section we show how the Bayesian model was validated, following the procedure highlighted in Section 6. 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 [36]. 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.

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

7.2.1 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, so we aim 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 that standard Gaussian distributions, allowing for values far from their center of mass. This is critical since we do not have much information about the true values of the parameters. The parameters are log-Cauchy distributed to ensure that they are strictly positive, a necessary condition since they are biologial parameters. Additional, 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.

In Fig. 4, 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 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 $\mathbb{R}$.

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. 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 successfuly perform parameter estimation or not.

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 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,

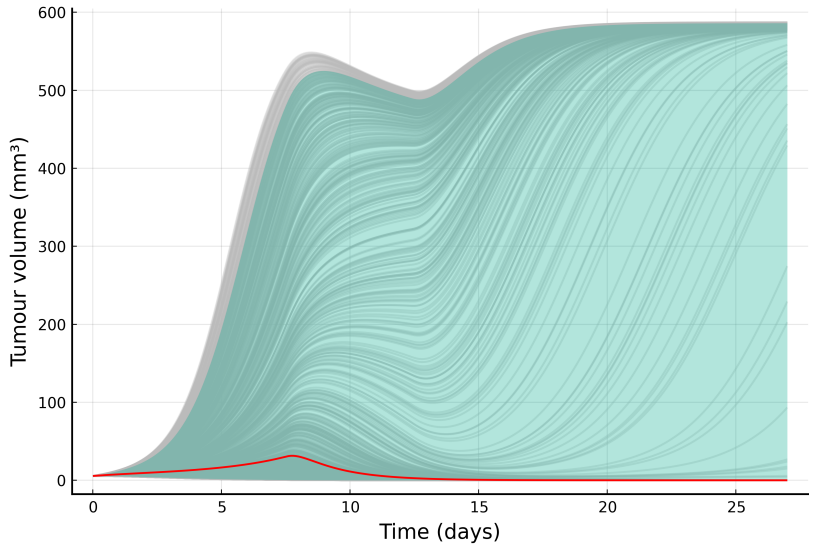


Figure 4: DDE solutions (grey lines) for 1,000 parameter vectors $\boldsymbol{\theta}$ sampled from the priors. The 95% credible interval for the tumour volume at each timestep (blue shade) shows that the model passes the prior predictive check: it contains our expected range of curves. The median grwoth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

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.

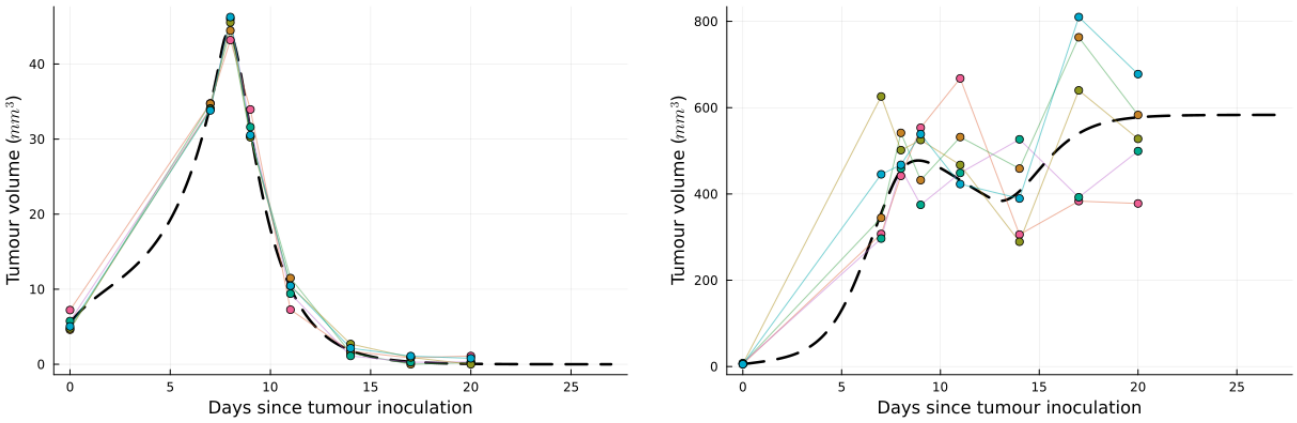
Table 3: Summary of the generation process for the two datasets A and B

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

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 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



(a) Dataset A (b) Dataset B

Figure 5: Fake data sets (scatter plot and colored 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.

Table 4: Assessement 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

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 centered on the true parameter values.

Table 5: Assessement 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 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 B, showing that a log-normal transformation is key to make exploration of the posterior easier to perform.

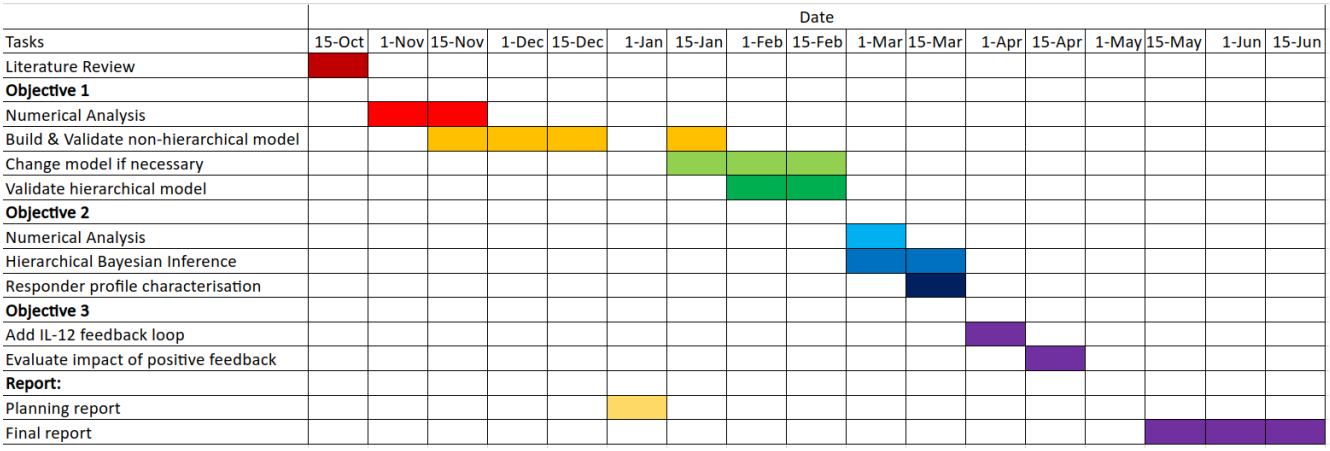
Table 6: Assessement of convergence for the MCMC chains for informative Normal priors

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 timiring 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 pf $CD8^+$ 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] 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] 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.

[9] 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.

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

[11] 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.

[12] G. Ding, T. Shen, C. Yan, et al. IFNg down-regulates the PD-1 expression and assist nivolumab in PD-1-blockade effect on CD8+ T-lymphocytes in pancreatic cancer. *BMC Cancer*, 19(1053), 2019.

[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] 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.

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

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

[26] 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.

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

[28] 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.

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

[30] 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.

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

[32] 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.

[33] 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.

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

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

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