October 2021

Mathematical model for osteosarcoma progression and treatments

Trang M. Le
University of Massachusetts Amherst

MATHEMATICAL MODEL FOR OSTEOSARCOMA PROGRESSION AND TREATMENTS

A Dissertation Presented

by

TRANG LE

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2021

Department of Mathematics and Statistics

MATHEMATICAL MODEL FOR OSTEOSARCOMA PROGRESSION AND TREATMENTS

A Dissertation Presented

by

TRANG LE

Approved as to style and content by:

Leili Shahriyari, Chair

Patrick Flaherty, Member

John Staudenmayer, Member

Zhengqing Ouyang, Member

Nathaniel Whitaker, Department Head
Mathematics and Statistics

Dedication

To my loving parents, Thuan Le and Hue Nguyen.

ACKNOWLEDGMENTS

I would like to thank my advisor, Leili Shahriyari, for her passion for cancer research, for her understanding and her thoughtful guidance during my studies at University of Massachusetts, Amherst.

I would like to express my appreciation to my committee members, Professors Patrick Flaherty, John Staudenmayer, Zhengqing Ouyang, for their generous support throughout my defense.

I would also like to thank my lab members, Sumeyye Su, Arkadz Kirshtein and Rachel Aronow, for their help and collaboration in my research, and for making my time at UMass so enjoyable.

I am thankful to all other faculties in the Department of Mathematics and Statistics for teaching me valuable skill sets and knowledge.

Lastly, I want to thank my family for always being very supportive, for believing in me and loving me.

ABSTRACT

MATHEMATICAL MODEL FOR OSTEOSARCOMA
PROGRESSION AND TREATMENTS
SEPTEMBER 2021
TRANG LE,
B.S., IOWA STATE UNIVERSITY
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST
Directed by: Professor Leili Shahriyari

Cancer is a complex disease where every tumor has its own characteristics, and thus different tumors may respond differently to the same treatments. Osteosarcoma, which is a rare type of cancer with poor prognosis, is especially characterized by its high heteogeneity. Therefore, it is important to study the progression of osteosarcoma tumors in different groups of patients with distinct characteristics. The immune system has been reported to play an important role in the development of various cancers with some immune cells having anti-tumor effects and others having pro-tumor effects. With recent advances in digital cytometry methods, which are techniques to estimate the fractions of various cell types from gene expression data of a bulk of cells, it became possible to obtain relative abundance of immune cells in tumors.

In this project, we review common digital cytometry methods, compare their performances and report the best method. We apply this best performing digital cytometry method to estimate abundance of immune cells in osteosarcoma tumors, and perform clustering using the estimated immune fractions to find groups of tumors with distinct immune compositions. We then model the growth of osteosarcoma tumors in each group while taking into account the interactions between immune cells and cancer cells. Lastly, we investigate the effects of adding chemotherapy on the progression of osteosarcoma, find the optimal chemotherapy dosages for tumors in each cluster, and compare the behaviors of immune and cancer cells under several conditions such as different treatment regimens and various treatment start times.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Cancer is the name for a group of diseases in which cells begin to grow uncontrollably. Every patient's tumor develops and progresses in its own unique way, and different tumors can have different sensitivities to cancer therapies [2]. The variations between tumors, called inter-tumor heterogeneity, have become an important task to be studied by researchers in order to provide more individualized and more effective treatments to cancer patients.

Several studies have shown that cancer cells and tumor infiltrating immune cells play a key role in tumor progression and the identification of malignant tumor types [3, 4, 5]. Innate immune cells contribute to tumor suppression in several ways, such as recognition and killing of cancer cells [6]. The immune response in the cancer microenvironment can be triggered by tumor antigen detection by immature dendritic cells, which then mature into dendritic cells [7]. Dendritic cells present these antigens to helper and cytotoxic T cells, resulting in their activation and the direct killing of cancer by cytotoxic cells [8, 9, 10]. Helper T cells and cytotoxic T cells also produce IFN-γ that inhibits tumor growth [10, 11, 12].

On the other hand, certain immune cells have promoting or dual effects on cancer progression. Regulatory T cells inhibit the differentiation and activities of helper and cytotoxic T cells, thus, indirectly promoting tumor by suppressing the

immune response [9, 13, 12, 14]. Macrophages, the most abundant immune cells in many cancers, have anti-tumor properties by activating helper and cytotoxic T cells through IL-12 and IL-23 production [15, 12, 16, 9], and also have pro-tumor properties through secreting IL-6, which supports cancer cells' proliferation [15, 17, 18, 19, 20].

Moreover, the immune system is one of the major players in the response to various cancer therapies [21, 22, 23], as it can improve or inhibit treatment effectiveness and tumor behaviors [24]. For example, the necrotic cell death of tumor cells caused by radiotherapy or chemotherapy triggers the production of high mobility group box 1 (HMGB1), which is a damage-associated molecular pattern (DAMP) molecule, and thus can induce immune responses [25, 26, 27, 28]. HMGB1 can promote dendritic cell maturation from naive dendritic cells [29, 30, 31, 32], leading to the activation of T cells and elimination of cancer cells. Meanwhile, most cancer therapies also kill immune cells, where immune cells have complex interactions with cancer cells, so the death of immune cells can have an indirect impact on the growth of tumor.

Since cancer is a heterogeneous disease with numerous components, such as immune cells, cancer cells, and lymphatic vessels [33], it is important take these various components of cancer into consideration while studying the tumor progression. However, in typical in vitro and in vivo studies, cancer mechanisms or components are usually analyzed one by one. While these experimental studies provide relevant insights about the mechanisms, none of them can provide the adequate required information to understand the complexity of cancer [34]. On the other hand, mathematical modeling allows us to investigate multiple components of a tumor at the same time. Thus, we can utilize mathematical models to combine our knowledge on the individual cancer components and study the heterogeneity

of cancer.

Mathematical models are commonly used to examine the growth of tumors, to identify the optimal combination of treatments, to improve responses to therapies, and to combat drug resistance in various types of cancer [35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45]. Among those, several studies model the interactions of immune cells and cancer cells; however, most of them only study one or two immune cells in their framework [46, 47, 48, 49, 50, 51, 52, 53]. A study by Wilkie et al. modeled the combination of all immune cells as one variable and analyzed its effects on tumor growth [54]. However, the impacts of the immune system on cancer are diverse with some immune cells having anti-tumor effects while others had pro-tumor effects, and thus modeling the whole immune system as one variable would fail to capture these important interactions. Only a few papers explore multiple immune cells [55, 56, 57], but even these models did not investigate the influence of macrophages, which have been shown to be the most abundant cell type in the tumor microenvironment of many cancers.

Malignant bone tumors are a rare type of cancers that arises in the bones. Osteosarcoma is the most common type of malignant bone tumors with about 1000 new cases diagnosed each year in the United States [58]. It can affect people of any age, but it is mostly occurred in the children aged 10 to 14 and in adults aged 65 and older [59, 60]. Osteosarcoma usually starts in the femur, the tibia and the humerus, and less commonly the skull, the jaw or the pelvis [15]. There are some factors such as gender, age, heritable syndromes and certain conditions such as Li-Fraumeni syndrome, hereditary retinoblastoma, and Bloom and Werner syndromes that affect the risk of osteosarcoma [61]. However, the cause of osteosarcomas is still not clear [62]. The types of standard treatments for osteosarcoma include surgery, chemotherapy, radiotherapy, and targeted therapy [63].

Although neoadjuvant chemotherapy has improved the outcomes of osteosarcoma, the overall survival of patients with metastatic tumors still remains in low rate in the last three decades [64, 65, 66]. Immunotherapy and targeted therapy have recently demonstrated significant results in the treatment of certain cancer types [67, 68]. Although these are popular alternative treatments for osteosarcoma, they are still ineffective for many patients [69]. Osteosarcoma tumors have also been reported to be resistant to the radiotherapy [70, 71]. Due to the poor prognosis of osteosarcoma, it would be beneficial to build mathematical models to investigate the progression of osteosarcoma, study the effects of common treatments on its growth, analyze the drug resistance, and explore new treatments' options.

There are a limited number of mathematical models that study cells in the bones or treatments for osteosarcoma. Some of these studies examine the growth of osteoblasts and osteoclasts in vitro, in healthy tissues or in bone metastases [72, 73, 74]. Osteoclasts are the cells that break down bones to initiate bone remodelling, while osteoblasts are the cells that synthesize new bone, and are also the cells of origin of osteosarcoma. Other studies focus on the pharmacokinetics of chemotherapy drugs for the treatment of osteosarcoma, [75, 76, 77]. However, to the best of our knowledge, there is currently no model that investigates the growth of primary tumors in osteosarcoma. And since the immune system plays an important role in cancer growth as mentioned above, we want to build a mathematical model for osteosarcoma that takes into account the interactions between cancer cells and immune cells.

The relationship between clinical outcome and immune abundance in osteosarcoma has been reported in many studies. Cytotoxic T cells, known as the primary receptor of immune response targeting ostersarcoma [10], have an important role in the immunological responses of osteosarcoma patients [78]. Also, a high number of

M1 macrophages in osteosaroma tumors has been associated with good prognosis in many studies [79, 85, 86], and it has been reported that low-risk patients have a high number of cytotoxic T cells and NK cells [80]. Moreover, certain chemotherapy drug such as Cisplatin can increase the cancer killing capacity of cytotoxic T cells [81, 82, 83, 84], so the effectiveness of the drug also depends on the number of cytotoxic T cells in the tumor microenvironment. Therefore, in this project, we will model osteosarcoma progression separately for patients with different immune compositions.

In order to distinguish patients with different immune compositions, we first need to obtain the relative immune abundance within the primary tumor of each patient. Many experimental approaches such as single-cell analysis tools, including immunohistochemistry, flow cytometry and mass cytometry, have been utilized to document tumor immune infiltrates. However, these methods are expensive and time-consuming [87], because they require laboratories, professionals and equipment. Obtaining the gene expression levels of a bulk of cells has become easier and cheaper thanks to new advances in high-throughput RNA-sequencing tools [88]. Therefore, several deconvolution methods (DMs) have been developed in recent years to estimate the relative abundance of each cell type in a bulk of cells, such as tumors, from their gene expression profiles. This process of digitally estimating the distribution of cell types from bulk gene expression data has also been referred to as digital cytometry [89].

We will use a digital cytometry method to estimate the relative number of each immune cell in primary tumors. We then divide osteosarcoma patients into groups based on their immune profiles and model osteosarcoma progression separately for each group. The estimated immune abundance will also be used as inputs to our model. The outline of this dissertation is as follows:

- Chapter 2: We review common digital cytometry methods and report the method with the best performance. The results of this chapter has been published in *Briefings in Bioinformatics* [90].

- Chapter 3: Using the best performing digital cytometry method to estimate immune abundance, we group osteosarcoma patients based on their immune compositions and compare survival outcomes among these groups. The results of this chapter has been published in *Mathematical Biosciences and Engineering* [91].

- Chapter 4: We build a mathematical model to study the progression of osteosarcoma tumors in each group and compare our findings to the results of Chapter 3. The results of this chapter has been published in *Cancers* [92].

- Chapter 5: We extend the model given in Chapter 4 to include common chemotherapy drugs for osteosarcoma and analyze the effects of chemotherapy on the growth of osteosarcoma tumors. The results of this chapter has been published in *Cells* [93].

CHAPTER 2

A REVIEW OF DIGITAL CYTOMETRY METHODS: ESTIMATING THE RELATIVE ABUNDANCE OF CELL TYPES IN A BULK OF CELLS

In order to build a mathematical model for osteosarcoma progression, we want to apply a digital cytometry method to estimate immune abundance in osteosarcoma tumors. The estimated abundance of immune cells in these tumors are then used to divide tumors into groups and as inputs to the mathematical model. As several digital cytometry methods have been recently developed, it is important to evaluate and compare their performance in estimating fractions of cells from the gene expression data of a bulk of cells, since the quality of the digital cytometry method we use will affect the results of our mathematical model.

There are two main categories of digital cytometry methods: linear models and ranked based models. In this chapter, we review five common digital cytometry methods [90], including three linear models: DeconRNASeq [94], CIBERSORT [95], and CIBERSORTx [89], and two rank-based models [96]: ssGSEA deconvolution method (DM) [97] and SingScore DM [98].

2.1 Methods

2.1.1 Linear models

DeconRNASeq

DeconRNASeq [94] stands for Deconvolution of RNA-Seq. This method treats the deconvolution task as a linear regression model with constraints on the model coefficients. This method assumes the total expression level of a gene in a sample is the sum of all the expression levels of the given gene in all cells in the sample.

DeconRNASeq takes as input the gene expression profile of a sample tissue, called mixture data, and a "signature matrix" where each column is a "typical" gene expression of a cell type. The method outputs the fractions of each cell type included in the signature matrix for the given sample. The general formula for this model is given as:

$$y = \mathbf{X}\beta. \tag{2.1}$$

Here, y denotes the observed gene expression level vector of a sample (mixture data), $\mathbf{X}$ denotes the signature matrix where each column is the gene expression level of a specific cell type, and β is the vector of estimated proportions of cell types.

DeconRNASeq finds the estimated proportions of cell types (β) by minimizing the following objective function:

$$\|y - \mathbf{X}\beta\|^2 \quad \text{s.t.} \quad \sum_i \beta_i = 1 \quad \text{and} \quad \beta_i \geq 0, \forall i, \tag{2.2}$$

where β_i is the estimated proportion of cell i in the sample. By minimizing this objective function, the linear regression model finds the coefficients that result in the smallest sum of squared difference between the observed and the predicted

expression levels in the sample. The constraints are designed to make sure that the cell proportions are positive and add up to 1. The optimization procedure is done using quadratic programming [99, 100, 101].

CIBERSORT

CIBERSORT [95] stands for Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts. Like DeconRNASeq, this method assumes the total expression level of a gene in a sample is the sum of expression levels of that gene in all the cells in that sample. CIBERSORT utilizes a machine learning technique called Support Vector Regression (SVR) for estimating cell proportions. Unlike linear regression, which tries to find the linear function that minimizes the sum of squared error, SVR tolerates a margin of error ϵ and only tries to minimize the sum of absolute error of data points that lie outside this margin of error. In particular, CIBERSORT uses the ν-SVR algorithm for this task. The general formula for CIBERSORT is the same as DeconRNASeq (Eq. (2.1)). Similar to DeconRNASeq, CIBERSORT takes a mixture data and a signature matrix as input and returns the model coefficient β as estimated fractions of each cell type in the sample. The only difference between these two methods is their optimization procedure; CIBERSORT finds β by minimizing the following objective function:

$$\frac{1}{2}\|\beta\|^2 + C\left(\nu\epsilon + \frac{1}{N}\sum_{i}^{N}(\xi_i + \xi_i^*)\right) \tag{2.3}$$

s.t. $y_i - \hat{y}_i \leq \epsilon + \xi_i$ and $\hat{y}_i - y_i \leq \epsilon + \xi_i^*$ and $\xi_i,\ \xi_i^* \geq 0$. Here, ϵ is the margin of error, and $(\xi_i + \xi_i^*)$ is the absolute error of data points that lie outside margin of error ϵ. ν is a model hyperparameter that gives an upper bound on the fraction of training error and a lower bound on the fraction of support vectors. Thus, the

value of ν is between 0 and 1.

Since y is a linear combination of $\mathbf{X}$ (Eq. (2.1)), CIBERSORT uses a linear kernel in ν-SVR. Unlike DeconRNASeq, CIBERSORT does not put any constraints on the model coefficient β during optimization, and there is no guarantee that elements of β will be non-negative and add up to 1. Thus, after the optimization process, CIBERSORT sets any negative coefficients to 0 and then normalizes the coefficients such that they sum to 1.

CIBERSORTx

Since the gene expression data sets can be collected through completely different experimental settings with the use of different experimentation plans, platforms, and methodologies, there are undesired batch effects in the gene expression values. These technical variations can in some cases be as large as the biological variations between different cell types [102]. It has been shown that the ComBat algorithm [103] can effectively remove these unwanted variations from bulk RNA-Seq data[102]. Newman et al. introduced CIBERSORTx [89], which extends CIBERSORT by adding batch correction using ComBat to address the possible cross-platform variations in gene expression data sets. CIBERSORTx introduces two strategies for batch correction: B-mode and S-mode.

CIBERSORTx B-mode

As in CIBERSORT, mixture data y is modeled as a linear combination of signature matrix $\mathbf{X}$ and cell fractions β.

$$y = \mathbf{X}\beta.$$

The algorithm of CIBERSORTx B-mode is as follows:

1. Use CIBERSORT to obtain estimated fractions $\hat{\beta}$ from mixture data y and signature matrix $\mathbf{X}$.

2. Create estimated mixture data $\hat{y}$, where $\hat{y} = \mathbf{X}\hat{\beta}$.

3. Use ComBat [103] to eliminate the cross-platform variation between y and $\hat{y}$, producing adjusted mixture data y_{adj}.

4. Use CIBERSORT to estimate final cell fractions β from y_{adj} and $\mathbf{X}$.

CIBERSORTx S-mode

When the technical variation between signature matrix and mixture data is more severe, Newman et al. [89] recommend to use S-mode, which adjusts the signature matrix instead of mixture data. As input, S-mode requires the mixture data y and the set of single-cell reference profiles R from which the signature matrix $\mathbf{X}$ was derived. R consists of single-cell transcriptomes from different cell types, typically multiple transcriptomes per cell type. The gene expression profile of each cell type in $\mathbf{X}$ is constructed by aggregating the corresponding single-cell transcriptomes in R. CIBERSORTx S-mode estimates cell fractions in the following way:

1. Let $\mu = [\mu_1, ..., \mu_c]$ be the fractions of each cell type from $\mathbf{X}$ in R and $\sigma = 2\mu$.

2. Generate artificial cell fractions β^* by drawing from normal distribution $Normal(\mu, \sigma)$.

3. Set negative values of β^* to 0 and normalize β^* so its components sum to 1.

4. Sample single-cell transcriptomes from R according to β^* and add them together to create artificial mixture data y^*.

5. Use ComBat [103] to eliminate the cross-platform variation between y and y^*, producing adjusted mixture data y_{adj} and y^*_{adj}.

6. Use non-negative least squares to find $\mathbf{X}_{adj}$ that minimizes $(y^*_{adj} - \mathbf{X}_{adj}\beta^*)^2$ such that $\mathbf{X}_{adj_{ij}} \geq 0$ for all i, j.

7. Use CIBERSORT to estimate final cell fractions β from original y and $\mathbf{X}_{adj}$.

2.1.2 Rank-based models

ssGSEA DM

All above-mentioned linear models rely on a signature matrix to deconvolve a bulk of cells using its gene expression profile. However, obtaining an accurate signature matrix is very challenging in practice, because factors such as variations in experimental settings and laboratory measurements can bias the signature matrix [104]. Fortunately, the most highly expressed genes for any given cell type are usually consistent across different laboratories and conditions. The deconvolution method based on single sample gene set enrichment analysis (ssGSEA) only uses these highly expressed genes of each cell type, here called cell signatures or up-regulated gene sets, instead of a signature matrix.

The ssGSEA method [104], which is a modification of gene set enrichment analysis (GSEA) [105], was developed in order to get an enrichment score for a single sample instead of two groups of samples. Here, we call the method developed by Senbabaoglu et al. [97], which utilizes the ssGSEA score specifically for the digital cytometry task, ssGSEA DM. This method takes mixture data and sets of highly expressed genes for each cell type as input, and returns the enrichment score for each cell type. The algorithm of ssGSEA DM is as follows:

1. Order mixture data by absolute expression from highest to lowest.

2. Replace gene expression values in mixture data by their ranks.

3. For each gene i, in the ordered rank data from step 2, compute the following:

$$P_G^w(G, y, i) = \frac{\sum_{r_j \in G, j \leq i} |r_j|^\alpha}{\sum_{r_j \in G} |r_j|^\alpha}, \tag{2.4}$$

$$P_{NG}(G, y, i) = \sum_{r_j \notin G, j \leq i} \frac{1}{(N - N_G)}, \tag{2.5}$$

where G is the given cell signature, containing N_G up-regulated genes, y is the mixture data, containing N genes, r_j is the rank of a gene j, and α is a parameter in $(0, 1]$.

4. The enrichment score for the sample y and cell signature G is given by:

$$ES(G, y) = \sum_{i=1}^{N} \left[P_G^w(G, y, i) - P_{NG}(G, y, i) \right]. \tag{2.6}$$

The enrichment score of the cell signature G tells us the relative fraction of the cell type with cell signature G. For example, assume T-cells and B-cells respectively have the cell signatures G_1 and G_2. If the enrichment score of the cell signature G_1 is higher than the enrichment score of the cell signature G_2, we conclude that the number of T-cells is higher than the number of B-cells in the mixture data.

SingScore DM

SingScore [98] stands for single sample scoring of molecular phenotypes. Similar to ssGSEA DM, SingScore DM uses enriched gene sets instead of a signature matrix for performing digital cytometry. While ssGSEA DM only uses a set of up-regulated genes for each cell type, SingScore DM has the option to use both up-regulated and down-regulated gene sets for each cell type. Thus, SingScore DM takes as input a set of up-regulated genes and an optional set of down-regulated genes for each cell type along with the mixture data, and outputs a score for each cell type. The algorithm of SingScore DM is as follows:

1. Order mixture data by gene expression levels from highest to lowest.

2. Use the top half of genes in the sample as the up-set and the bottom half as the down-set. An important remark is that these up-set and down-set genes of the sample are different from the up-regulated and down-regulated gene sets of each cell type.

3. Rank these genes in ascending order for up-set and descending order for down-set.

4. For a given cell type and its up-regulated gene set G_{up} and down-regulated gene set G_{down}, calculate the following:

$$S_{dir,i} = \left(\frac{\sum_{g \in G_{dir}} R^{g}_{dir,i}}{N_{dir,i}} \right) \tag{2.7}$$

   where dir is the gene set direction (up-/down-regulated), $S_{dir,i}$ is the score for sample i against the directed gene set, $R^{g}_{dir,i}$ is the rank of gene g in the directed set from the sample (up-set or down-set), $N_{dir,i}$ is the number of genes in G_{dir} that are observed within the mixture data.

5. Calculate the normalized score:

$$\bar{S}_{dir,i} = \frac{S_{dir,i} - S_{min,i}}{S_{max,i} - S_{min,i}} \tag{2.8}$$

   where $S_{min,i} = \frac{N_{dir,i}+1}{2}$ and $S_{max,i} = \frac{2N_{total,i}-N_{dir,i}+1}{2}$

6. Calculate output score for sample i:

$$\bar{S}_{total,i} = \bar{S}_{up,i} + \bar{S}_{down,i} \tag{2.9}$$

Similar to ssGSEA DM, the output scores of SingScore DM algorithm are the relative levels of each cell type in the sample, rather than the actual fractions.

2.2 Approach

To compare the performance of the above-mentioned methods on the deconvolution task, we generate simulation data with known mixing fractions and signal to noise ratio (SNR) ranging from 100:5 to 100:50 (n = 100 samples for each SNR). We use two different signature matrices: LM22 [95] and LM6 [106] for DeconRNASeq, CIBERSORT, and CIBERSORTx. LM22 is derived from microarray data and consists of 547 gene expressions for twenty-two leukocytes, while LM6 is derived from RNA-Seq data and has 684 gene expressions for six leukocytes. These signature matrices are also used to derive the up-regulated gene sets used for ssGSEA DM and SingScore DM. We note that the single-cell reference profiles needed to run CIBERSORTx S-mode are not available for the LM6 signature matrix, hence this method is excluded from the LM6 results.

We construct the simulation data in the following manner: first, "known" mixing fractions for a sample are obtained by drawing random numbers from Uniform$(0, 1)$ and then normalized so that the fractions in a sample sum to 1. Mixture data is then formed by a linear combination of the LM22 source gene expression profile and the known mixing fractions, where LM22 source gene expression profile is the gene expression profile used to create LM22, before the gene selection step. Noise is induced to the simulation data by adding values drawn from Normal$(0, \frac{k}{100} \cdot \sigma)$, where σ is the global standard deviation of the original simulation data without noise, and k is an integer. This results in a SNR ratio of 100:k for a given value of k. We create ten sets of simulated data with k chosen from 5 to 50, in steady increments of 5, resulting in SNR ratios ranging from 100:5 to 100:50.

It is conceivable to encounter mixture data that consists of more cell types than

those available in the signature matrix. To test for each method's robustness to this phenomenon, we delete a few cell types in the signature matrix LM22 and run all five methods using the simulated mixture data with two signature matrices: reduced LM22 and the original LM6.

We also apply these methods on two experimental data sets: whole blood RNA-Seq data with ground truth cell fractions estimated by flow cytometry (n = 12) (available on Gene Expression Omnibus under the accession number GSE127813 [89]), and PBMC microarray data with ground truth fractions estimated by flow cytometry (n = 20) (available on Gene Expression Omnibus under the accession number GSE65133 [95]). Although these data sets come from different platforms, we cannot make any conclusive statements that these methods perform better on RNA-Seq vs microarray data due to the limited availability of data sets with ground truth fractions. For each experimental data set, we run all five methods with the original LM22 and the original LM6.

To facilitate a fair comparison of the performance of the five deconvolution methods, we want the signature matrix used in linear methods to come from the same database as the gene sets used in rank-based methods. Thus, we use the data sources of LM22 to create up-regulated gene sets for ssGSEA DM and SingScore DM by applying a method similar to [107]. First, we separate the samples in the single cell reference profiles of LM22 into groups according to their cell type. For each gene, we calculate the difference between the minimum expression in the group of interest and the highest mean expression of all other groups. If this difference is greater than a threshold, we select this gene as an up-regulated gene for the cell type of the analyzed group. We do this for every group to select the highly expressed genes in each cell type specific group. We apply the same technique to derive up-regulated gene sets from LM6. However, since we do not have access to

the reference profiles of LM6, we apply this technique on LM6 itself. These gene sets derived from LM22 and LM6 are available for download on our GitHub page (see section 2.5).

We ran the above digital cytometry methods using the following software:

- DeconRNASeq package in R for DeconRNASeq,
- CIBERSORT's R source code for CIBERSORT,
- CIBERSORTx's website application for CIBERSORTx,
- gsva package in R for ssGSEA DM,
- SingScore package in R for SingScore.

Each of these methods have their own form of normalization in their algorithm. In ssGSEA DM and SingScore DM, the normalization is not applied on the input mixture data, and since input gene sets are only lists of gene names, no normalization can be applied here either. The only normalization in rank-based methods is applied on output scores at the end of the algorithm, thus making output scores nicer for visualization without changing the correlation of the predicted values with the ground truth fractions. On the other hand, CIBERSORT and CIBERSORTx use z-score normalization on their inputs (the signature matrix and mixture data) as a mandatory initial step in their software, and DeconRNASeq software provides an option to first standardize the input mixture data. The results of these linear models with normalization will differ from those without. CIBERSORT and CIBERSORTx also have optional quantile normalization on the input mixture data. This quantile normalization is recommended for microarray data but not for RNA-Seq data. Since the PBMC microarray data set has been previously normalized with the limma package in R, using 'normexp' background correction

with negative controls [95], and the whole blood data set is RNA-Seq data, we use CIBERSORT and CIBERSORTx without quantile normalization in this study. For DeconRNASeq, ssGSEA DM and SingScore DM, we use normalization (which is the default setting for these methods).

We note that since rank-based methods output relative scores as opposed to frequencies, we cannot use traditional metrics such as mean square error to compare the performance of each method to the original data. Thus, we instead consider four different measures of correlation: Pearson correlation per sample, Pearson correlation per cell, Spearman correlation per sample, and Spearman correlation per cell. Correlation per sample between estimated and true fractions tells us how well a method estimates the relative frequency of all cell types in a given sample, while correlation per cell tells us how well a method estimates the relative frequency of a given cell type between all samples.

2.3 Results

2.3.1 Analysis of simulation data

We create simulation data as described in the Approach section, and apply each of the above-mentioned methods with both the reduced LM22 and the original LM6 signature matrices (or corresponding derived gene sets for rank-based methods). For all methods and signature matrices, the Pearson and Spearman correlation results are consistent with one another. With reduced LM22, the more noise is added to the data set, the lower the correlations observed between the ground truth fractions and the methods' predictions, with the following exceptions: the sample-level correlation of DeconRNASeq, ssGSEA DM, and SingScore DM does not vary

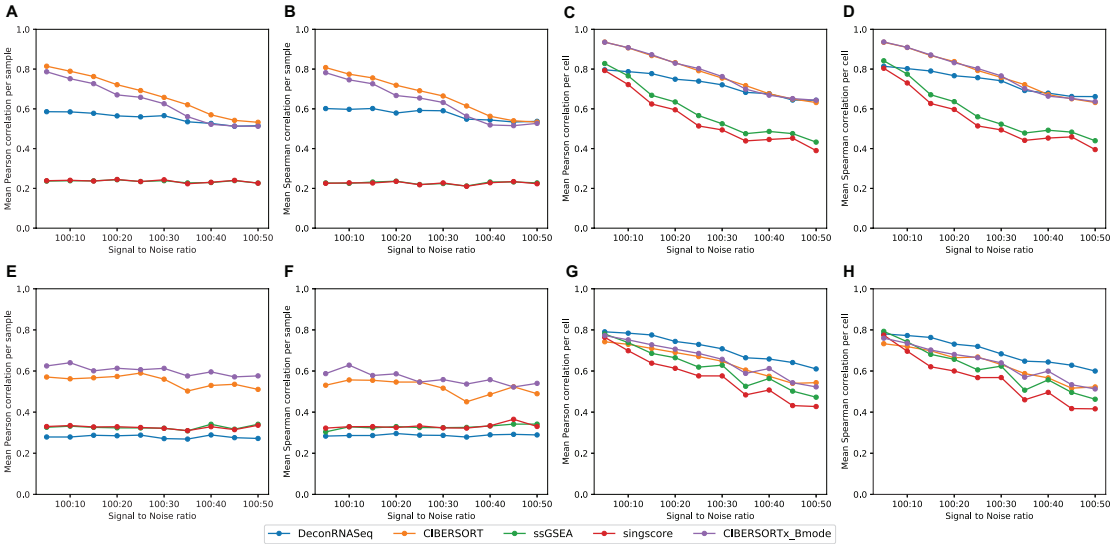

Figure 2.1. **Results on simulation data with different SNR from 100:5 to 100:50.** Note in A and B, the ssGSEA results align with SingScore. A-D: mean correlation between flow cytometry cell fractions and predicted cell fractions obtained from running methods with reduced LM22, across different noise levels. A: sample-level Pearson correlation, B: sample-level Spearman correlation, C: cell-level Pearson correlation, D: cell-level Spearman correlation. F-I: mean correlation between flow cytometry cell fractions and predicted cell fractions obtained from running methods with LM6, across different noise levels. F: sample-level Pearson correlation, G: sample-level Spearman correlation, H: cell-level Pearson correlation, I: cell-level Spearman correlation.

much when the signal to noise changes (Figure 2.1A,B). For LM22, CIBERSORT and CIBERSORTx B-mode perform best on data with high signal to noise. The batch correction in CIBERSORTx does not appear to improve the method for this simulated data, however we note that the simulated data and signature matrix are both derived from LM22, and therefore there should not be any cross platform variation to eliminate. Interestingly, DeconRNASeq with reduced LM22 performs just as well as CIBERSORT and CIBERSORTx when the noise level is high.

For all methods with LM6, cell-level correlations with true fractions again decrease as noise increases, but sample-level correlations stay roughly the same across all noise levels (Figure 2.1E-H). The insensitivity to the noise levels of sample-level results with LM6 could be due to the fact that simulation data was created using LM22 source gene expression profile instead of LM6. When LM6 is used, CIBERSORT and CIBERSORTx B-mode again perform the best per sample, although correlations are lower than with reduced LM22. Figure 2.1E,F shows that CIBERSORTx B-mode outperforms CIBERSORT with LM6, suggesting that CIBERSORTx B-mode is better than CIBERSORT when the signature matrix and mixture data are from different platforms. Interestingly, DeconRNASeq with LM6 performs worst per sample (Figure 2.1 E,F), but best per cell (Figure 2.1G,H). The poor performance per sample of DeconRNASeq with LM6, but strong performance with LM22, may indicate DeconRNASeq's lack of robustness when signature matrix comes from a different platform than mixture data. It is worth noting that the rank-based methods perform poorly across all noise levels with both signature matrices. In particular, rank-based methods produce very low sample-level correlations with ground truth fractions (Figure 2.1A-B,E-F).

In addition to examining the mean correlations, we created box plots of the 100 different sample-level correlations, and the sixteen (LM22) or six (LM6) dif-

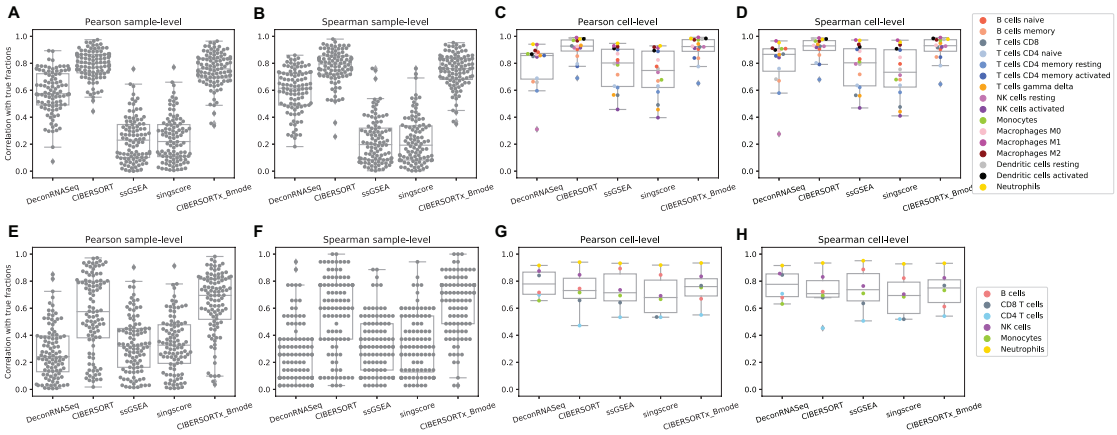

Figure 2.2. **Results on simulation data with SNR 100:10.** A-D: box plots of correlations between ground truth cell fractions and predicted cell fractions obtained from running methods with LM22. A: sample-level Pearson correlation, B: sample-level Spearman correlation, C: cell-level Pearson correlation, D: cell-level Spearman correlation. F-I: box plots of correlations between ground truth cell fractions and predicted cell fractions obtained from running methods with LM6. F: sample-level Pearson correlation, G: sample-level Spearman correlation, H: cell-level Pearson correlation, I: cell-level Spearman correlation.

ferent cell-level correlations, using the simulation data generated with SNR 100:10 (Figure 2.2). We observe in Figure 2.2A-B,E-F that the variances in sample-level correlations, particularly with LM6, are quite large for all methods. The cell-level correlation plots (Figure 2.2C-D) show that while CIBERSORT and CIBERSORTx with LM22 produce results that are highly correlated ($r > .75$) with the ground truth for all individual cell types besides CD4 memory resting T-cells, the other methods show a larger variation in performance among the different cell types. In particular, we note that DeconRNASeq does a poor job predicting the number of resting NK cells, bringing down the method's mean correlation per cell with LM22. However this observation does not hold when LM6 is used. This may help explain the discrepancy in DeconRNASeq's performance in mean cell-level correlation with LM22 and LM6. With LM6, all methods besides DeconRNASeq do a poor job predicting the relative number of CD4 T Cells, but do a great job predicting the relative level of Neutrophils.

2.3.2 Analysis of whole blood data

We obtain the ground truth fractions of neutrophils, lymphocytes, monocytes, T-cells, CD8+ T-cells, CD4+ T-cells, B-cells, and NK cells for whole blood data (GSE127813) from CIBERSORTx [89] and compare them with the corresponding estimated fractions obtained using the above-mentioned methods. Since LM22 has twenty-two leukocytes, we sum the estimated fractions of certain cell subtypes to match them with the ground truth cell types. For example, we sum the estimated fractions of CD4+ naive T-cells, CD4+ memory resting T-cells, and CD4+ memory activated T-cells to compare it with the ground truth fraction of CD4+ T-cells. However, there are still eight leukocytes that do not have a similar match to the categories of the ground truth cells, and so we exclude them from our

analysis. These cell types are macrophages (M0, M1, M2), dendritic cells (resting, activated), mast cells (resting, activated), and eosinophils. Hence we do not expect estimated percentages to sum precisely to 1. A similar procedure is applied to the results obtained with LM6, and we end up comparing estimated fractions and ground truth fractions in neutrophils, monocytes, T-cells, CD8+ T-cells, CD4+ T-cells, B-cells and NK cells.

In agreement with the simulation data results, the linear methods perform better in overall than the rank-based methods in terms of correlation per sample (Figure 2.3A-B,F-G). Of all the methods, CIBERSORT and CIBERSORTx B-mode again perform best per sample, both with LM22 and LM6, while CIBERSORTx S-mode performs worse than these two. DeconRNASeq performs much better in terms of Pearson correlation (and slightly better in terms of Spearman correlation) with the use of LM22 than with LM6. On the other hand, ssGSEA DM and SingScore DM perform very poorly with gene sets from LM22 (Figure 2.3A-B), but do a better job with gene sets from LM6 (Figure 2.3F-G). We should note that sample-level Pearson correlations of ssGSEA DM and SingScore DM are low for both gene sets, but since Pearson correlation is not a good measure for rank-based results, we should focus on Spearman correlation when analyzing the performance of these two methods.

With regards to correlation per cell, CIBERSORT and CIBERSORTx B-mode perform the best with LM22 (Figure 2.3C-D). In contrast to the simulation data results, rank-based methods perform best with LM6 (Figure 2.3H,I), though CIBERSORT and CIBERSORTx B-mode are not far behind. DeconRNASeq performs the worst out of all methods with LM6, but still achieves cell-level Pearson and Spearman correlations > 0.6. CIBERSORTx S-mode performs very well, but still worse than original CIBERSORT and CIBERSORTx B-mode. In particular, all

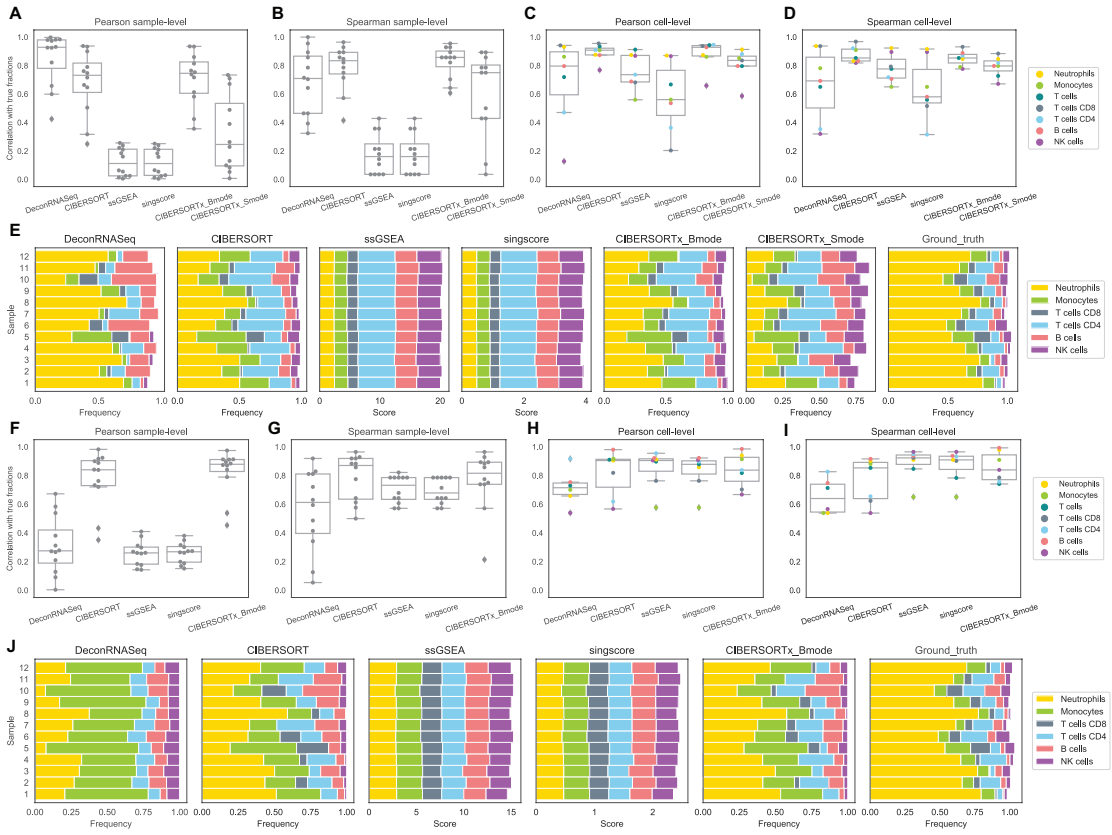

Figure 2.3. **Results on whole blood data.** A-D: box plots of correlations between flow cytometry cell fractions and predicted cell fractions obtained from running methods with LM22. A: sample-level Pearson correlation, B: sample-level Spearman correlation, C: cell-level Pearson correlation, D: cell-level Spearman correlation. E: stacked bar charts of predicted cell fractions by each method with LM22 and ground truth flow cytometry cell fractions. F-I: box plots of correlations between flow cytometry cell fractions and predicted cell fractions obtained from running methods with LM6. F: sample-level Pearson correlation, G: sample-level Spearman correlation, H: cell-level Pearson correlation, I: cell-level Spearman correlation. J: stacked bar charts of predicted cell fractions by each method with LM6 and ground truth flow cytometry cell fractions.

linear methods do a somewhat poor job predicting NK cells, compared to other cell types, and rank-based methods do a poor job at predicting monocytes.

Overall, CIBERSORT and CIBERSORTx give the best results for the whole blood data set (Figure 2.3). Among CIBERSORT models, CIBERSORTx S-mode performs poorly, but still gives relatively good results in terms of Spearman correlation with ground truth fractions (Figure 2.3 B,D). The comparison between Figure 2.3A and Figure 2.3F shows that DeconRNASeq results are very different between LM22 and LM6, suggesting that DeconRNASeq is very sensitive to the signature matrix. In combination with the analysis of DeconRNASeq in simulation data, this result indicates that DeconRNASeq's performance is highly dependent on the compatibility between signature matrix and mixture data.

Lastly, we compare the cellular profiles generated by each method to the ground truth fractions in Figure 2.3E,J. Since ssGSEA DM and SingScore DM return enrichment scores instead of estimated fractions, the total sum of the output scores for each sample does not need to be less than or equal to 1. As mentioned earlier, since they are rank-based, we should not expect them to produce scores close to ground truth fractions, but rather hope to see their output scores consistent with the ranks of true fractions (i.e. if neutrophil have the highest number in the ground truth data, we would expect these methods to give neutrophils the highest score among all cell types). However, ssGSEA DM and SingScore DM estimate similar scores of cell types for different samples, even though ground truth fractions differ across samples (Figure 2.3E,J). Linear methods, on the other hand, are able to capture the difference in distribution of fractions across samples. Although fractions estimated by CIBERSORT and CIBERSORTx do not completely match the ground truth fractions, these methods do succeed in capturing important patterns such as the relative levels of neutrophils, T-cells, B-cells, and NK cells in samples. All

methods overestimate the fraction of B-cells, and all but DeconRNASeq with LM22 drastically underestimate the fraction of neutrophils (Figure 2.3E). The two rank-based methods, ssGSEA DM and SingScore DM, produce similar output scores to each other. CIBERSORT and CIBERSORTx are expected to give similar estimated fractions since they both use ν-SVR and the only difference is that CIBERSORTx uses batch-correction before applying ν-SVR. We observe a slight improvement in performance with CIBERSORTx compared to CIBERSORT for this data set.

2.3.3 Analysis of PBMC data

PBMC data, (GSE65133) [95], includes flow cytometry fractions for naive B-cells, memory B-cells, CD8+ T-cells, CD4+ naive T-cells, CD4+ memory resting T-cells, CD4+ memory activated T-cells, $\gamma\delta$ T-cells, NK cells, and monocytes. Following the same procedure as mentioned before, we compare the estimated fractions to flow cytometry fractions of naive B-cells, memory B-cells, CD8+ T-cells, CD4+ naive T-cells , CD4+ memory resting T-cells, CD4+ memory activated T-cells, $\gamma\delta$ T-cells, NK cells and Monocytes for LM22, and B-cells, CD8+ T-cells, CD4+ T-cells, NK cells, and Monocytes for LM6.

We repeat the correlation analysis with the PBMC data set, and the sample-level results are somewhat different from our findings from the whole blood data set. In general, the rank-based methods perform much better in terms of correlation per sample on the PBMC data than on the other two data sets, while DeconRNASeq performs considerably worse. With LM6, sample-level Spearman correlations show high variance across samples for all methods (Figure 2.4F-G). This implies that all methods have an inconsistent behaviour; i.e. for some samples they perform better than other samples. However, it is worth noting that there are only five cell types overlapping between LM6 cell types and ground truth cell types. The

sample-level correlation across only five cell types is susceptible to being low when only one or two cell types are poorly predicted, and as seen in Figure 2.4H-I, these methods do a poor job estimating the relative frequency of B-cells with LM6. In fact, we also observe high variance in LM6 sample-level results on the simulation data (Figure 2.2G), where the number of cell types (six) is small as well and the poor estimation of CD4 T-cells likely contributes to some samples having low correlation with true fractions with all methods. Overall, CIBERSORT, CIBERSORTx B-mode, ssGSEA DM and SingScore DM perform better per sample than DeconRNASeq and CIBERSORTx S-mode with LM22 (Figure 2.4A,B), and Figure 2.4 F,G indicates no significant differences in performance between linear models and rank-based methods with LM6.

With regards to cell-level correlation, CIBERSORT and CIBERSORTx B-mode again perform best with the LM22 signature matrix, while the rank-based methods perform considerably worse compared to the whole blood and simulation data (Figure 2.4C,D,H,I). However, when LM6 is used, the rank-based methods outperform the linear models (Figure 2.4H-I). All methods except DeconRNASeq with LM22 signature matrix do well on predicting the number of CD8+ T-cells and Monocytes. Additionally, all methods but DeconRNASeq with LM6, ssGSEA DM and SingScore DM with LM22 show high correlations with the ground truth for NK cells. CIBERSORT and CIBERSORTx B-mode with LM22 signature matrix are the only methods for which the majority of cell types have correlation coefficients $r > 0.5$. However, even these methods struggle to accurately predict the number of CD4+ memory resting T-cells and $\gamma\delta$ T-cells with LM22 (Figure 2.4C-D), as well as the number of B-cells when using LM6 (Figure 2.4H-I).

Similar to the whole blood data set, we also plot the predicted cellular profiles estimated by the methods in a stacked bar chart, along with the ground truth

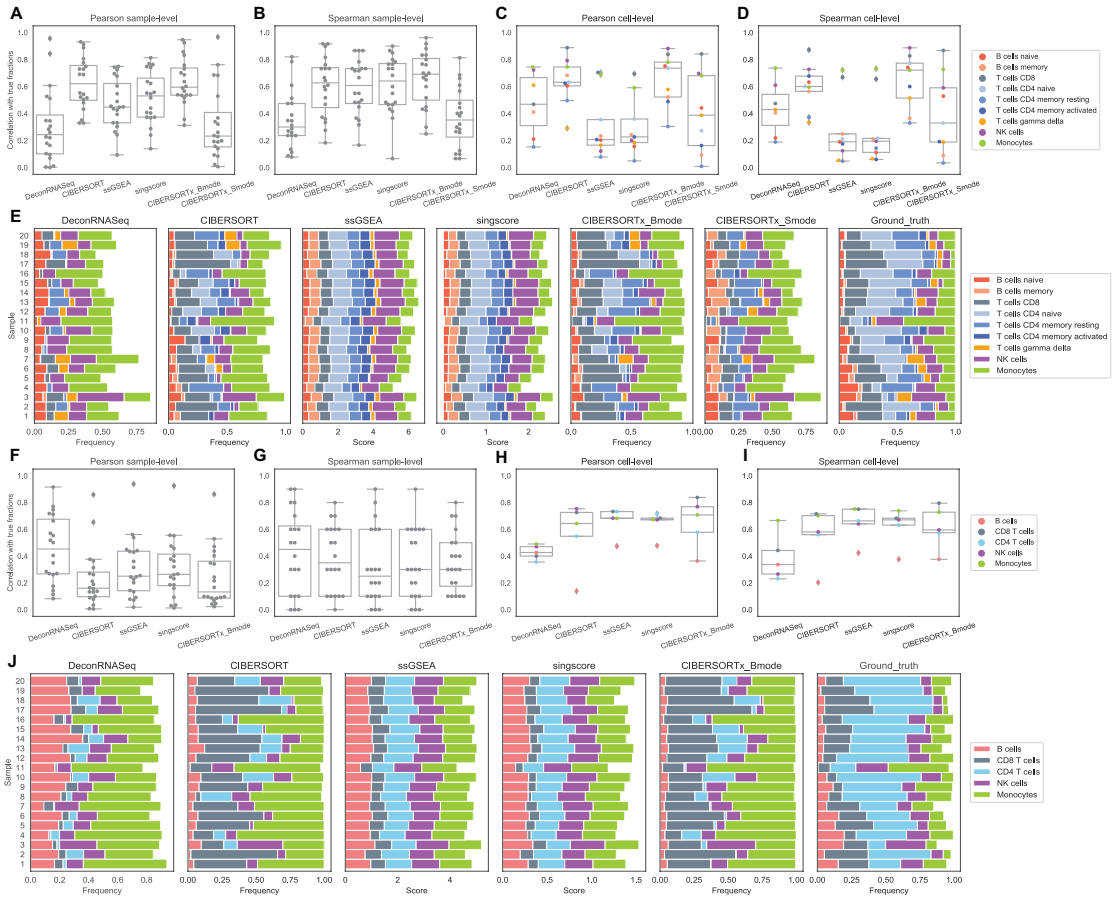

Figure 2.4. **Results on PBMC data.** A-D: box plots of correlations between flow cytometry cell fractions and predicted cell fractions obtained from running methods with LM22. A: sample-level Pearson correlation, B: sample-level Spearman correlation, C: cell-level Pearson correlation, D: cell-level Spearman correlation. E: stacked bar charts of predicted cell fractions by each method with LM22 and ground truth flow cytometry cell fractions. F-I: box plots of correlations between flow cytometry cell fractions and predicted cell fractions obtained from running methods with LM6. F: sample-level Pearson correlation, G: sample-level Spearman correlation, H: cell-level Pearson correlation, I: cell-level Spearman correlation. J: stacked barcharts of predicted cell fractions by each method with LM6 and ground truth flow cytometry cell fractions.

fractions (Figure 2.4E,J). We again note that we have omitted cell types that were present in the signature matrix but not in the PBMC data. Hence, we do not expect the frequencies to necessarily sum to 1. The rank-based methods do a slightly better job at capturing variations among samples in this data set as compared to the whole blood data. However, the CIBERSORT methods, particularly CIBERSORTx B-mode, again exhibit the best overall performance.

2.4 Discussion

As mentioned in the Approach section, the normalization in linear methods can affect their performance on the deconvolution of bulk gene expression data. Unlike CIBERSORT and CIBERSORTx, DeconRNASeq does not have clear guidance on whether to use its normalization and when to use it. We tried DeconRNASeq both with and without normalization on our data sets. DeconRNASeq with normalization gives overall better results in the whole blood and PBMC data sets, and gives similar results to without normalization in simulation data. However, we would like to note that just because DeconRNASeq with normalization works better on these specific data sets does not mean it would work better on other data sets as well. One positive aspect of the linear methods' normalization is that it helps these algorithms converge faster and easier, which reduces the run time significantly.

We provide a comparison of each of these methods' run time per sample in Table 2.4. Run time for each of the three data sets was calculated as the average of twenty runs, using a 2.5 GHz Intel Core i7 CPU with 16 GB of RAM, and then normalized by the number of samples in the data set. All methods besides CIBERSORTx were run in R. We note that the CIBERSORTx team provides a web portal to run their software, removing any dependencies on hardware or software.

Table 2.1: Method Runtimes per Sample (mean +/- standard deviation, in milliseconds)

Data set	**Method**	**Runtime with LM22**	**Runtime with LM6**
Wholeblood (58,581 genes/sample)	DeconRNASeq	53.11 ± 1.93	41.23 ± 0.81
	CIBERSORT	325.23 ± 15.11	181.84 ± 8.11
	ssGSEA	1296.12 ± 53.3	293.59 ± 13.85
	SingScore	13.93 ± 0.49	7.04 ± 0.35
PBMC (34,694 genes/sample)	DeconRNASeq	21.65 ± 0.54	18.86 ± 0.55
	CIBERSORT	354.68 ± 16.29	109.9 ± 5.67
	ssGSEA	115.7 ± 4.72	73.81 ± 2.91
	SingScore	6.36 ± 0.24	1.66 ± 0.11
Simulation (11,845 genes/sample)	DeconRNASeq	12.87 ± 0.14	11.97 ± 0.09
	CIBERSORT	158.79 ± 9.09	80.73 ± 4.66
	ssGSEA	64.24 ± 1.33	39.91 ± 0.66
	SingScore	1.39 ± 0.02	1.27 ± 0.01

Therefore it would not be a fair comparison to include this method in Table 2.4. While the original CIBERSORT also has a web portal to run the method, we used the CIBERSORT R source code to record time-to-compute, while making sure that the results from the R code are identical to those from the web portal. We note that CIBERSORTx takes longer than CIBERSORT, since the method runs batch correction before applying CIBERSORT. SingScore is the fastest of the five methods, followed by DeconRNASeq, while ssGSEA and CIBERSORT are significantly slower. Unsurprisingly, all methods run faster with the smaller

LM6 signature matrix than with LM22. In general, the run time per sample also decreases for data sets with a smaller number of genes per sample.

As discussed in the Results section, rank-based methods tend to estimate very similar scores across samples, while linear models are able to capture some variations in fractions among samples (Figure 2.3E,J and Figure 2.4E,J). Since rank-based methods use the rank of the genes instead of the actual expression value in the calculation of output score, as long as the genes have the same ranks across samples, ssGSEA DM and SingScore DM will output the same scores across samples (even when these genes have very different expression values across samples). Thus, rank-based methods can still successfully estimate the ranks of frequencies between cell types, but might fail to estimate the relative frequencies of a given cell type in samples. This is, in fact, the main disadvantage of rank-based methods, as mentioned in Table 2.4 where we discuss the advantages and disadvantages of each digital cytometry method. There has been a novel attempt, introduced by Aran D. et al [108], to transform enrichment scores to make them more comparable with cell fractions. These transformed enrichment scores are intended to be more on the same range with cell fractions, but are not designed to be used in place of cell fractions. Converting enrichment scores to cell fractions is generally a hard problem, since enrichment scores are derived using only ranks of the genes instead of gene expression values themselves.

Table 2.2: Advantages and disadvantages of methods

Methods	**Advantages**	**Disadvantages**

DeconRNASeq	• Outputs are cell fractions • Open source implementations available in Python and R • Quick run-time	• Requires a signature matrix as an input • Performance is highly dependent on the compatibility between signature matrix and mixture data
CIBERSORT	• Outputs are cell fractions • Open source implementations available in Python and R • Web portal available for running method • Good performance on digital cytometry task	• Requires a signature matrix as an input • Slow run-time
CIBERSORTx B-mode	• Outputs are cell fractions • Web portal available for running method • Good performance on digital cytometry task • Eliminates the batch effect between signature matrix and mixture data by adjusting mixture data	• Requires a signature matrix as an input

CIBERSORTx S-mode	• Outputs are cell fractions • Web portal available for running method • Eliminates the batch effect between signature matrix and mixture data by adjusting signature matrix	• Requires a signature matrix as an input • Does not perform as well as CIBERSORTx B-mode
ssGSEA DM	• Does not require a signature matrix; it only uses the up-regulated gene sets of each cell type • Open source implementations available in Python and R	• Outputs are scores for each cell type rather than cell fractions • Produces similar scores for samples with varying distributions of cell types • Slow run-time
SingScore DM	• Does not require a signature matrix; it can use both up-regulated and down-regulated gene sets of each cell type • Open source implementations available in Python and R • Quick run-time	• Outputs are scores for each cell type rather than cell fractions • Produces similar scores for samples with varying distributions of cell types

The rank-based methods generally perform better in terms of all four correlation metrics when the up-regulated gene sets came from the same platform as the mixture data. In particular, for the microarray PBMC and simulation data, ssGSEA DM and SingScore DM results have higher correlations with the ground truth when using LM22 (derived from microarray data) compared to LM6 (derived from RNA-Seq data). Similarly, for the RNASeq whole blood data, ssGSEA DM and SingScore DM results are more correlated with ground truth fractions when using LM6 compared to LM22.

We further note that the rank-based methods analyzed in this study were originally introduced for the task of analyzing the enrichment of a gene set in a single sample, and have recently been adopted for the digital cytometry task. Enrichment analysis refers to a group of methods for determining a set of enriched genes either in a sample or between two groups of samples. There are three generations of enrichment analysis methods: over-representation analysis, functional class sorting techniques, and pathway topology-based techniques [96]. Both ssGSEA and SingScore belong to the second generation, functional class sorting techniques. To the best of our knowledge, among all enrichment analysis methods, only these two single sample enrichment methods have been used for digital cytometry. It would be worth exploring whether adopting other single sample enrichment methods for digital cytometry would lead to better results.

In terms of linear methods, throughout this study, we have seen good performance from both CIBERSORT and CIBERSORTx B-mode. Since CIBERSORTx uses batch correction to account for cross-platform variation between the signature matrix and mixture data, we should expect CIBERSORTx to perform at least as well as CIBERSORT when signature matrix and mixture data come from different platforms. Indeed, we see marginal improvements of CIBERSORTx B-mode

over CIBERSORT in both PBMC and simulation data with LM6 (Figure 2.1E-H, 2.4F-I) where mixture data comes from microarray data and signature matrix comes from RNA-Seq data, and qualitatively similar performance between CIBERSORTx B-mode and CIBERSORT in whole blood data with LM22 (Figure 2.3A-D) where mixture data comes from RNA-Seq data and signature matrix comes from microarray data. In fact, CIBERSORTx B-mode slightly outperforms CIBERSORT in both PBMC and whole blood experimental data sets regardless of signature matrix used. However, CIBERSORTx B-mode underperforms CIBERSORT by a small margin in the simulation data with LM22, raising the possibility that batch correction may negatively affect the performance of CIBERSORT if signature matrix and mixture data come from the exact same platform.

Newman et al. [89] mention that CIBERSORTx B-mode should be used when signature matrix is derived from bulk sorted reference profile or when the technical variation between signature matrix and mixture data is moderate, while CIBERSORTx S-mode should be used when this variation is high. Figure 2.3A-D and Figure 2.4A-D show that CIBERSORTx S-mode performs worse than both original CIBERSORT and CIBERSORTx B-mode in both experimental data sets, suggesting that the technical variation between LM22 and these data sets is not high. These results also suggest that it is better to use CIBERSORTx B-mode than CIBERSORTx S-mode when the technical variation between signature matrix and mixture data is low. We would like to mention that many studies on tumor microenvironment have recently utilized digital cytometry methods, most commonly CIBERSORT [109, 110, 111, 112, 113] and ssGSEA [114, 115, 116, 117], and two separate studies on blood leukocyte and tumor infiltrating leukocytes enumeration indicate that iSort, a transcriptome deconvolution method based on CIBERSORTx, achieves highly accurate and robust results for both blood and tumor

samples [118, 119].

2.5 Data Availability

The PBMC data set and its flow cytometry fractions are available on the CIBERSORT website at `https://cibersort.stanford.edu` under the name "Fig 3a PBMCs Gene Expression" and "Fig 3a PBMCs Flow Cytometry", respectively. The whole blood data set is available on Gene Expression Omnibus with identifier GSE 127813, and its flow cytometry fractions are available on the CIBERSORTx website at `https://cibersortx.stanford.edu` under the name "Ground truth whole blood (txt)". The simulation data created for this study, as well as the up-regulated gene sets we derived from LM22 and LM6, can be found on our github page. The data portion of our github repository is located at `https://github.com/ShahriyariLab/TumorDecon/tree/master/TumorDecon/data`. The simulation data and gene sets can be found under the names "Simulation_data", "LM22_up_genes.csv", and "LM6_up_genes.csv", respectively.

CHAPTER 3

IMMUNE CLASSIFICATION OF OSTEOSARCOMA

As indicated in chapter 2, CIBERSORTx B-mode performs the best among popular digital cytometry methods, especially when the mixture data and signature matrix come from different platforms. In this chapter, we use CIBERSORTx B-mode to investigate the immune patterns of osteosarcoma tumors and analyze the relationship between immune composition and clinical features of osteosarcoma patients [91].

There have been many other studies that utilize a deconvolution method to study the tumor microenvironment of osteosarcoma. A number of them use the estimated immune infiltrations calculated using CIBERSORT and/or immune scores calculated using ESTIMATE to find immune-related genes that can predict the prognosis of osteosarcoma [85, 120, 121, 122]. Another set of studies find genes with prognostic values by applying Cox model on survival data or performing differentially expressed genes analysis between two groups of interest, and then investigate the relationship between these genes and estimated immune infiltrates [80, 86, 123, 124, 125]. Others study the association of immune abundance with clinical information directly [79, 121, 126, 127, 128]. Our work falls somewhat into the third category.

Among the studies that directly investigate the relationship between immune

infiltrations and clinical information, three of them use TARGET data set [79, 121, 127], one uses data from GSE21257 [126], and one uses data from GSE39058 [128]. Most of these studies use CIBERSORT [121, 127, 126, 128], while the other uses ssGSEA and ImmuCellAI along with expression of immune marker genes to get estimated abundance of immune cells [79]. Our study, on the other hand, utilizes the latest and best performing version of deconvolution methods called CIBERSORTx B-mode, and conducts our analyses on both TARGET data set and GSE21257 data set. We also perform K-means clustering using the estimated immune abundance to study the clinical characteristics of different immune patterns in osteosarcoma. Another study has used hierarchical clustering on immune abundance, but they estimate immune abundance using ssGSEA and do not focus on the clinical difference between clusters [129].

3.1 Materials and methods

3.1.1 Data collection and processing

The gene expression data sets in this study are obtained from 2 cohorts: TARGET (cohort 1) and GSE21257 (cohort 2). Cohort 1 includes FPKM normalized RNA-seq data of 88 osetosarcoma patients downloaded from the UCSC Xena web portal and their corresponding clinical data downloaded from the GDC data portal. Cohort 2 includes microarray data and corresponding clinical features of 53 osteosarcoma samples, downloaded from GEO website. Cohort 2's gene expression data were previously normalized with robust spline normalization before being downloaded.

3.1.2 Relative abundance of immune cells

We calculate the relative frequencies of 22 immune cell types by applying CIBERSORTx [89] B-mode algorithm with immune signature matrix LM22 on gene expression data from both cohort 1 and 2. The results in chapter 2 have shown that CIBERSORTx B-mode gives good estimates of immune abundance in both RNA-Seq and microarray data with the use of LM22, and in fact outperforms CIBERSORT and other tumor deconvolution methods. We obtain estimated immune fractions by running the algorithm on their website with 100 permutations. Similar to CIBERSORT, CIBERSORTx outputs a p-value for each deconvolved sample as an indicator of confidence of the results. We use samples with p-value less than 0.05 for analyses in this chapter.

3.1.3 Identification of immune patterns in osteosarcoma

To calculate the abundance of each cell type, which has several subtypes with a small abundance, we sum the proportions of their subtypes obtained from CIBERSORTx. The abundance of B cells is the summation of naive and memory B cells; NK cell is the summation of resting and activated NK cells; Mast cells is the summation of resting and activated Mast cells; Dendritic cells is the summation of resting and activated Dendritic cells; and CD4 T cells is the summation of follicular helper T cells, regulatory T cells, naive CD4 T cells, resting and activated memory CD4 T cells. We do not combine subtypes of Macrophages because M1 and M2 Macrophages have very different functions and Macrophages make up the majority of immune cells in osteosarcoma.

We then use K-means clustering to identify various immune patterns of osteosarcoma based on the estimated immune cells' abundance. The number of clusters in

K-means is determined using elbow method. A t-SNE visualization of the clusters is also included to see how well K-means algorithm distinguishes between samples with different immune patterns.

3.1.4 Immune scores of osteosarcoma tumors

Immune scores of all samples in cohort 1 and 2 are computed from ESTIMATE algorithm [130]. In order to do so, we run *estimate* package locally from R. We also divide all patients from both cohorts into low immune score and high immune score group using the median immune score as cut-off to study the relationship between immune score and survival outcome.

3.1.5 Statistical analysis

Chi-square test is used to analyze the relationship between categorical variables in this study. We employ Mann-Whitney-Wilcoxon test to detect any significant difference between groups of continuous variables, such as immune fractions, gene expression, age and immune score. Pearson correlation and corresponding p-value are used to study the correlation between different immune infiltrates.

To investigate the impact of immune infiltrates on survival, for each immune cell, we split all patients into high and low abundance group using the median value as cut-off and perform log-rank test to find significant difference in survival between groups. Kaplan-Meier curves are also plotted to visualize the differences between these groups.

All analyses in this study are conducted on all samples in both cohorts 1 and 2, except for metastasis-free survival analysis which is applied only on cohort 2 since cohort 1 does not include data on the time of metastasis development. All statistical analyses and visualizations are carried out in Python.

3.2 Results

3.2.1 The most abundant immune cells in osteosarcoma are Macrophages and CD4 T cells

Results of CIBERSORTx B-mode on gene expression profiles of 141 osteosarcoma patients (cohorts 1 and 2) demonstrates that M0 Macrophages is the most frequent immune cell in osteosarcoma tumors with an average of 40% of total immune cells, followed by M2 Macrophages and CD4 T cells (Figure 3.1B, C). Unsupervised hierarchical clustering of immune cell fractions shows that most abundant cells tend to be clustered together, as is shown in Figure 3.1A, where M0 and M2 Macrophages are clustered together and then grouped with CD4 T cells and other immune cells. In addition, the most frequent cells also have the highest variation in abundance across osteosarcoma tumors (Figure 3.1B).

3.2.2 Correlation between immune infiltrates in osteosarcoma

According to the CIBERSORTx B-mode results, abundance of CD8 T cells is negatively correlated with M0 Macrophages and positively correlated with M1 Macrophages with Pearson correlation coefficients of -0.62 and 0.55, with p-values of 4.8e$-$16 and 1.1e$-$12, respectively (Figure 3.1E). The proportion of $\gamma\delta$ T cells is also significantly negatively correlated with CD4 T cells and NK cells (Pearson coefficients of -0.6 and -0.62, with p-values of 4.3e-15 and 1.8e-16). Interestingly, frequencies of M0 and M1 Macrophages exhibit a negative correlation of -0.61 with p-value 1e$-$15.

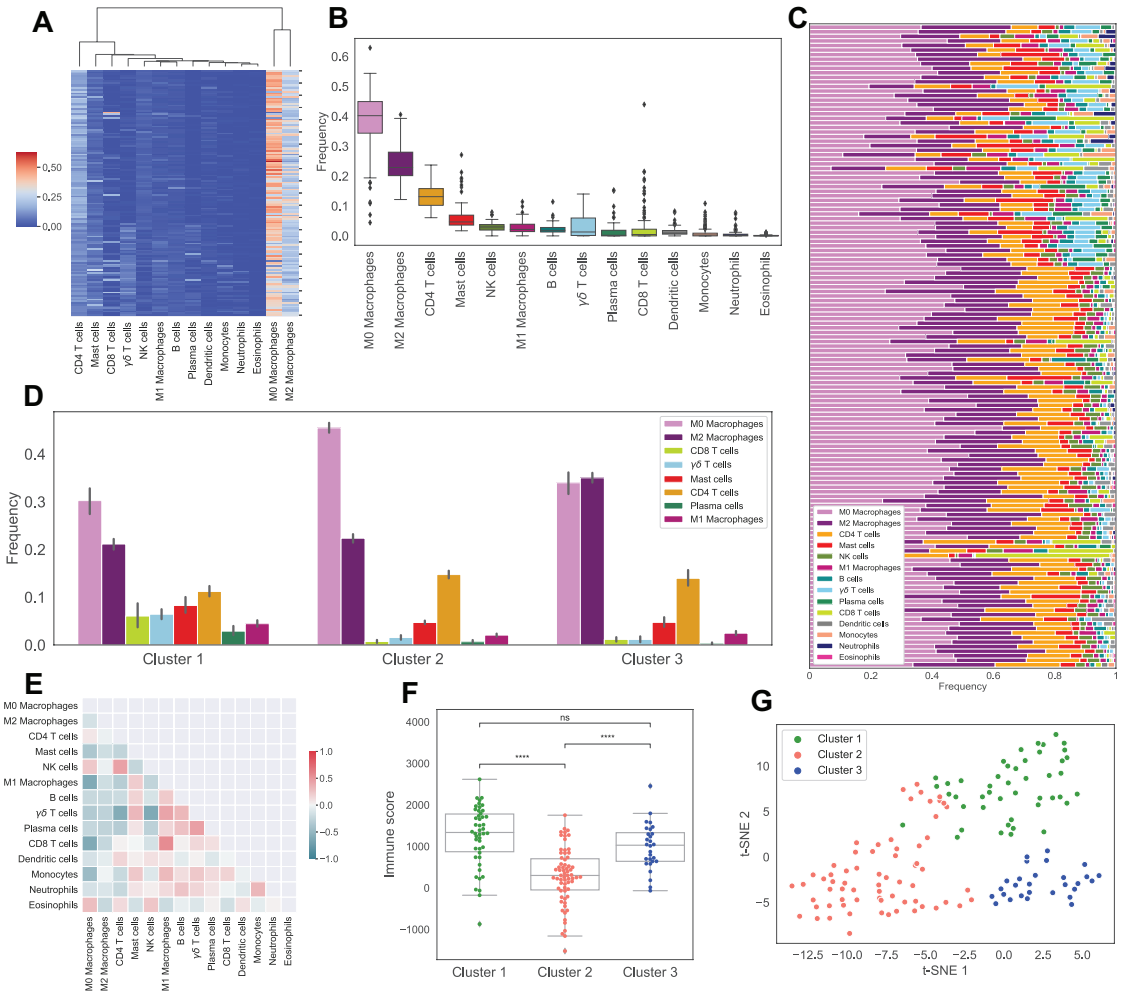

Figure 3.1. **Immune pattern of osteosarcoma.** Sub-figure A shows the hierarchical clustering of estimated immune cells' infiltration. Sub-figure B and C display the boxplot and stacked barchart of these immune cells' fractions. Sub-figures D shows the average frequencies of immune cells in 3 clusters obtained from K-means clustering. Sub-figure E indicates the correlation map of immune cell frequencies. Sub-figure F displays the boxplot of ESTIMATE immune scores in 3 clusters, with asterisks indicating significant difference from Mann-Whitney-Wilcoxon test (ns: no significance, *: $0.01 < p \leq 0.05$, **: $0.001 < p \leq 0.01$, ***: $0.0001 < p \leq 0.001$, ****: $p \leq 0.0001$). Sub-figure G shows t-SNE plot of estimated immune abundance, color coded by cluster.

3.2.3 There are 3 immune patterns of osteosarcoma

K-means clustering of immune cell proportions in osteosarcoma tumors indicates the existence of three distinct immune classes (Figure 3.1D), namely: Cluster 1, which has the highest proportions of CD8 T cells, $\gamma\delta$ T cells, M1 Macrophages, Mast cells and Plasma cells and the lowest proportion of M0 Macrophages; Cluster 2, in which the percentage of M0 Macrophages is the highest; and Cluster 3, which has the highest percentage of M2 Macrophages. A t-SNE plot of immune cell proportions suggests that K-means clustering algorithm successfully separates osteosarcoma patients with different immune patterns (Figure 3.1G).

3.2.4 Cluster 2 has the worst survival outcome among all clusters

While there is no significant difference in gender, age and proportion of metastasis at diagnosis between clusters (Figure 3.2A, B, F), we observe some differences in survival outcomes among clusters.

Kaplan-Meier curves indicate that cluster 2 has the worst survival probability throughout time out of all clusters (Figure 3.2G). In addition, cluster 2, along with cluster 1, has higher percentage of dead patients at the last time of follow up than cluster 3 (Figure 3.2C). Interestingly, cluster 2 also has the lowest immune scores compared to other clusters (Figure 3.1F).

Cluster 3 appears to have the best outcome among clusters. It has the lowest percentage of dead patients at the last time of follow up among all clusters (Figure 3.2C), and better survival rate than cluster 2 over time (Figure 3.2G) with p-value 0.07 from the log-rank test.

Cluster 1, which has the highest amount of CD8 T cells, $\gamma\delta$ T cells, M1 Macrophages and Mast cells, has slightly better overall survival time than clus-

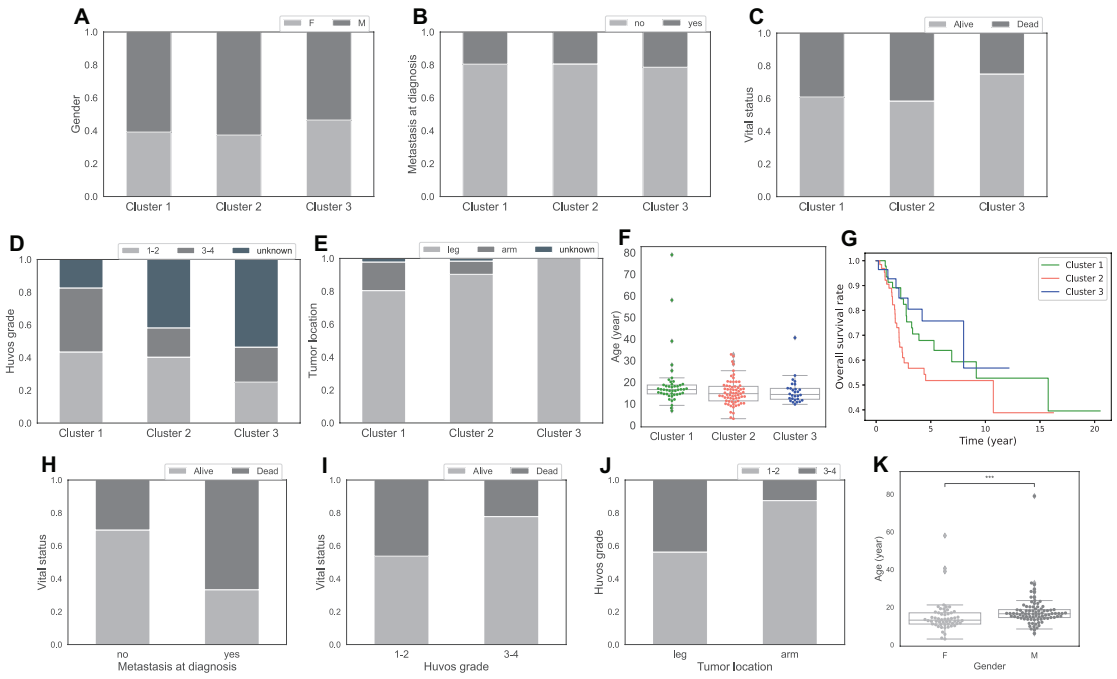

Figure 3.2. **Clinical characteristics of clusters and relationship between clinical features of osteosarcoma.** Sub-figure A-E show percentage of patients by gender (A), metastasis at diagnosis (B), vital status at the last time of follow-up (C), Huvos grade (D), primary tumor location (E), in the 3 clusters. Sub-figure F shows a boxplot of patients' age at diagnosis in each cluster. Sub-figure G displays Kaplan-Meier curves of overall survival across 3 clusters. Sub-figures H-K shows the association between clinical features, H: percentage of alive and dead patients by metastasis at diagnosis, I: percentage of alive and dead patients by Huvos grade, J: percentage of low and high Huvos grade by tumor location, K: boxplot of age at diagnosis by gender, with asterisks indicating significant difference from Mann-Whitney-Wilcoxon test (ns: no significance, *: $0.01 < p \leq 0.05$, **: $0.001 < p \leq 0.01$, ***: $0.0001 < p \leq 0.001$, ****: $p \leq 0.0001$).

ter 2 (p = 0.16, Figure 3.2G). However, cluster 1 seems to have worse outcome than cluster 3 due to its higher percentage of dead patients at the last time of follow up (Figure 3.2C). It is worth noting that there is no significant difference in the survival rate between cluster 1 and 3 according to the log-rank test (p = 0.5, Figure 3.2G).

3.2.5 There is a relationship between certain clinical features of osteosarcoma

The Chi-square test and Mann-Whitney-Wilcoxon test show a relationship between vital status and metastasis at diagnosis (p = 0.001, Figure 3.2H), vital status and Huvos grade (p = 0.036, Figure 3.2I), which is a grading system to evaluate a patient's response to chemotherapy based on the percentage of necrosis in the tumor after treatment, and between gender and age (p = 2.3e−4, Figure 3.2K) where male patients are older on average. We observe that patients with metastasis at diagnosis have much higher percentage of being dead at the last time of follow up than patients without metastasis (Figure 3.2H). This makes perfect sense since metastases have been known to associate with late stages of tumor and poor prognosis in many cancers. The other clinical feature with a relation to vital status is Huvos grade. Higher percentage of patients with high Huvos grade (3-4) are alive at the last time of follow up than patients with low Huvos grade (1-2) (Figure 3.2I), which is reasonable since a high Huvos grade means good response to chemotherapy. Figure 3.2J suggests that primary osteosarcoma tumors in the arm respond more poorly to chemotherapy than leg tumors, as illustrated by the high proportion of Huvos grade 1-2 in arm tumors. However, it is important to note that primary osteosarcoma tumors happen more often in the leg than in the arm (Figure 3.2E), and the observation in Figure 3.2J is based on a small number

of arm tumor samples (n = 8), thus the relationship between tumor location and Huvos grade is not considered significant by the Chi-square test (p = 0.18).

3.2.6 Immune score does not relate to vital status directly, but does relate to survival probability over time

Figure 3.3D indicates no clear difference in immune score between alive and dead patients at the last time of follow up. However, Kaplan-Meier curves of high and low immune score with a median cut-off reveal that the high immune score group has a better outcome (Figure 3.3E). The log rank test supports this observation with a p-value of 0.03. Thus, higher immune score is associated with better survival probability throughout time. This is consistent with the observation of outcome in the clusters. Cluster 2 has significantly lower immune score than cluster 1 and 3 (Figure 3.1F), with p-values of 5e−9 and 2.1e−5 from Mann-Whitney-Wilcoxon test, and accordingly worse overall survival probability over time than cluster 1 and 3 (Figure 3.2G). Cluster 1, with the highest average immune score among clusters (Figure 3.1F), even though has about the same proportion of dead patients at the last time of follow up as cluster 2 (Figure 3.2C), shows better survival time than cluster 2 (Figure 3.2G).

3.2.7 Relationship between immune infiltrates and survival outcome in osteosarcoma

The Mann-Whitney-Wilcoxon test shows that there is a significant difference in the level of $\gamma\delta$ T cells and Mast cells between alive and dead patients at the last time of follow up, with p-values of 0.045 and 0.022, where dead patients are associated with higher percentages of $\gamma\delta$ T cells and Mast cells than alive patients (Figure 3.3A). Dendritic cells, NK cells and CD8 T cells also seem to associate with

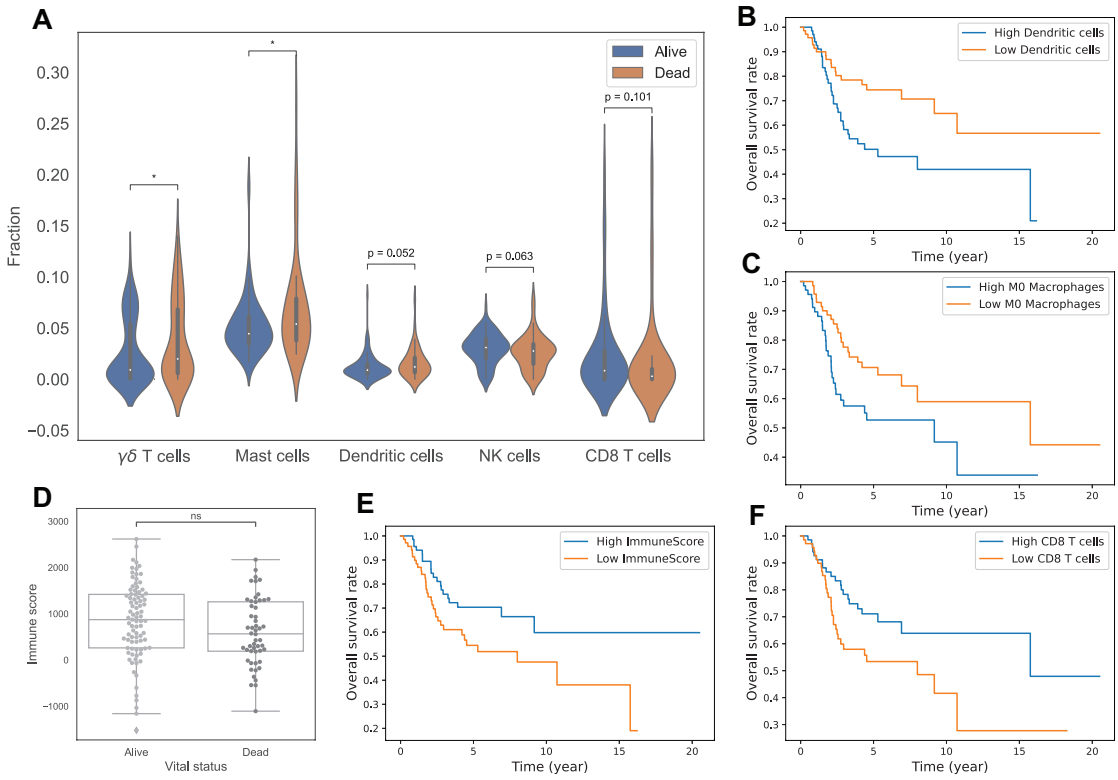

Figure 3.3. **Relationship of immune infiltrations with survival in osteosarcoma.** Sub-figure A displays violin plots of fractions of $\gamma\delta$ T cells, Mast cells, Dendritic cells, NK cells, CD8 T cells between alive and dead patients. Sub-figures B, C, E, F show Kaplan-Meier curves of overall survival between 2 groups, B: high vs low Dendritic cells, C: high vs low M0 Macrophages, E: high vs low ESTIMATE immune score, F: high vs low CD8 T cells. Sub-figure D shows a boxplot of ESTIMATE immune scores between alive and dead patients. Note: asterisks indicate significant difference from Mann-Whitney-Wilcoxon test (ns: no significance, *: $0.01 < p \leq 0.05$, **: $0.001 < p \leq 0.01$, ***: $0.0001 < p \leq 0.001$, ****: $p \leq 0.0001$).

survival status in osteosarcoma. We observe higher level of NK cells ($p = 0.063$), CD8 T cells ($p = 0.1$) and lower level of Dendritic cells ($p = 0.052$) in alive patients than in dead patients (Figure 3.3A).

Kaplan-Meier curves (Figure 3.3B, C, F) and the log rank test indicate an association between survival outcomes and levels of Dendritic cells, M0 Macrophages and CD8 T cells, with a p-value of 0.01, 0.04 and 0.04, respectively. Low Dendritic cells, low M0 Macrophages and high CD8 T cells are associated with better survival probability over time in osteosarcoma patients. This is again consistent with the outcome of the clusters where cluster 2, with the highest level of M0 Macrophages and lowest CD8 T cells, has the worst overall survival.

Overall, we found that $\gamma\delta$ T cells, Mast cells, Dendritic cells, M0 Macrophages, NK cells and CD8 T cells have a relationship with the survival of osteosarcoma patients.

3.2.8 Association of immune infiltrates with other clinical features

We see no significant relationship between age or metastasis at diagnosis and the frequencies of immune cells. However, we notice an association of M1 Macrophages and CD8 T cells' frequencies to metastasis-free survival. High levels of M1 Macrophages and CD8 T cells are associated with better metastasis-free survival probability across time in osteosarcoma (Figure 3.4A, B), with p-values of 0.05 and 0.08 from the log-rank test, respectively. This means that patients with more M1 Macrophages and CD8 T cells are less likely to develop metastasis or die at any given time than patients with low percentage of these cells.

A relationship between some immune infiltrates and other clinical features of osteosarcoma has also been observed. Higher level of NK cells is associated with good response to chemotherapy ($p = 0.035$, Figure 3.4C). Patients with arm tumors

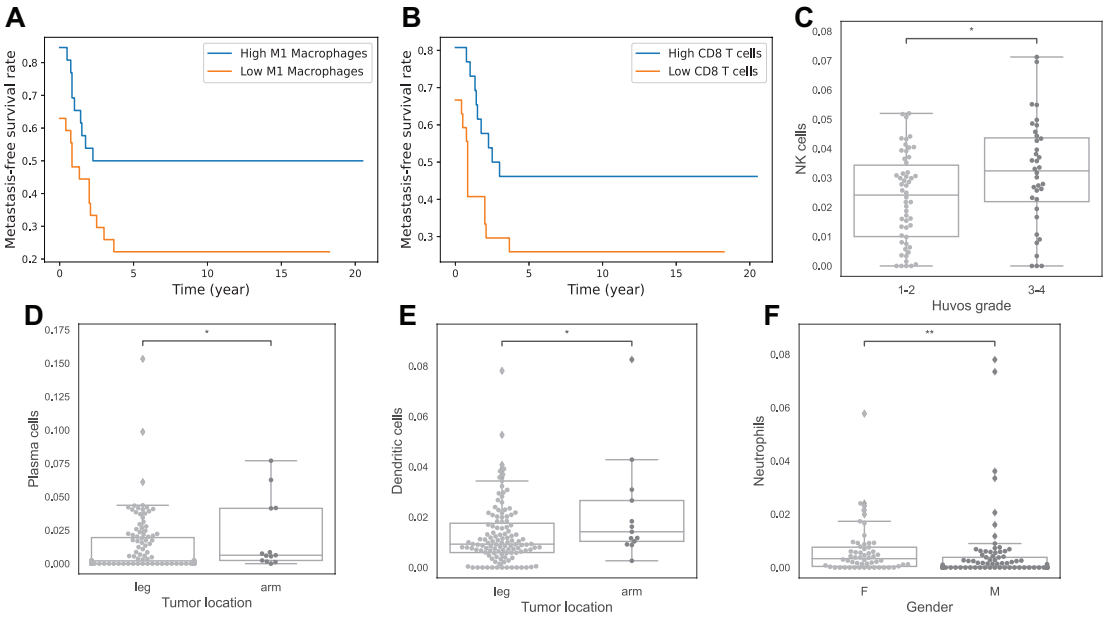

Figure 3.4. **Association of immune infiltrations and other clinical features.** Sub-figures A and B show Kaplan-Meier curves of metastasis-free-survival in cohort 1 between 2 groups, A: high vs low M1 Macrophages, B: high vs low CD8 T cells. Sub-figures C-F are boxplots to indicate relationship of immune infiltrates with Huvos grade (C), primary tumor location (D and E), and gender (F). Note: asterisks indicate significant difference from Mann-Whitney-Wilcoxon test (ns: no significance, *: $0.01 < p \leq 0.05$, **: $0.001 < p \leq 0.01$, ***: $0.0001 < p \leq 0.001$, ****: $p \leq 0.0001$).

have higher percentage of Plasma cells and Dendritic cells than patients with leg tumors (Figure 3.4D, E), with p-values of 0.046 and 0.033. Lastly, female patients are shown to have higher frequency of Neutrophils (Figure 3.4F), with p-value 0.0016.

3.2.9 Expression level of genes encoding PD-1, INF-γ, CTLA4, TNF, IL1-β, IGF1, IL-6 and RUNX2 are significantly different for some clusters

We use the gene expression values for some important proteins, and we analyze gene expression value of the proteins separately for cohort 1 and 2 because they have different data types: RNA-Seq and microarray, respectively.

Programmed cell death protein 1 (PD-1) is a type of protein on T cells and cancer cells use it to bind with PD-1 ligand (PD-L1) and PD-2 ligand (PD-L2) to escape cell death by immune cells. There is a high correlation between PDCD1 gene, which encodes PD-1 protein, and CD8 T cells in both data sets with correlation coefficient of 0.70 and 0.77, respectively, and p-values less than 0.05 (Figure 3.6). As a result of this correlation, PDCD1 is the highest in cluster 1 (Figure 3.5A1, B1). Cluster 1 also has the highest expression of CTLA4 gene (Figure 3.5A3, B3) that encodes Cytotoxic T-Lymphocyte Associated Protein 4 (CTLA4), which is a member of immunoglobulin superfamily and has been found to significantly associate with the risk of osteosarcoma [131, 132]. Moreover, we see that gene expression value of CTLA4 is significantly correlated with CD8 T cells in osteosarcoma tumors (Figure 3.6).

Interferon γ (INF-γ), encoded by IFNG gene, has antiviral, immunoregulatory, and anti-tumor properties in the immune system and is secreted by mostly T cells

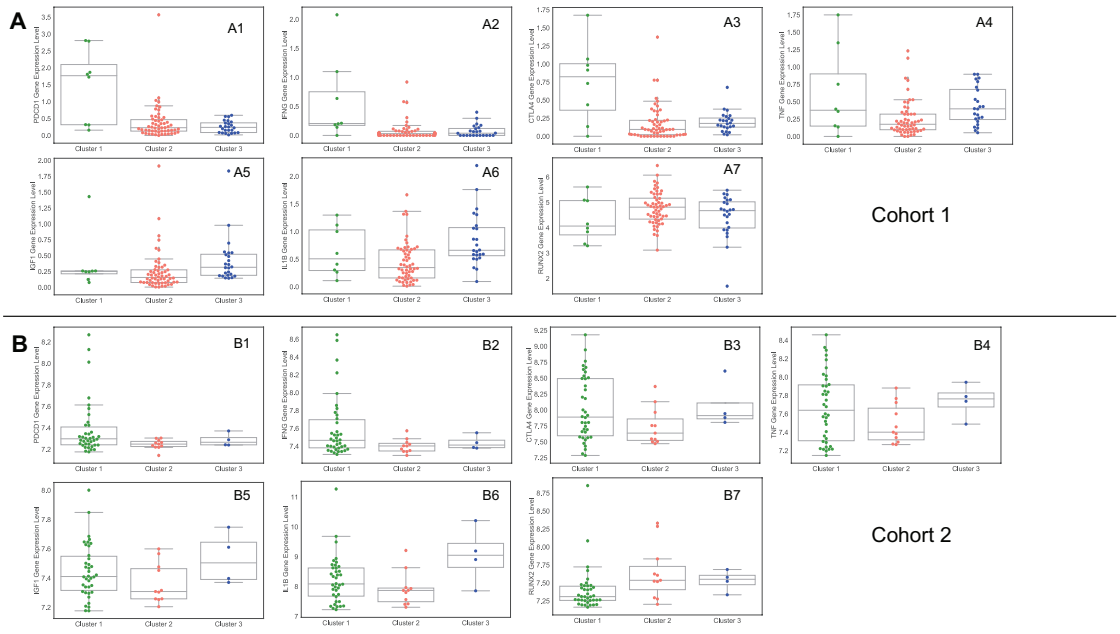

Figure 3.5. **Gene expression values of important proteins in the clusters.** Sub-figures (A1–A7) and (B1–B7) show the gene expression values that come from cohort 1 and cohort 2 data sets, respectively.

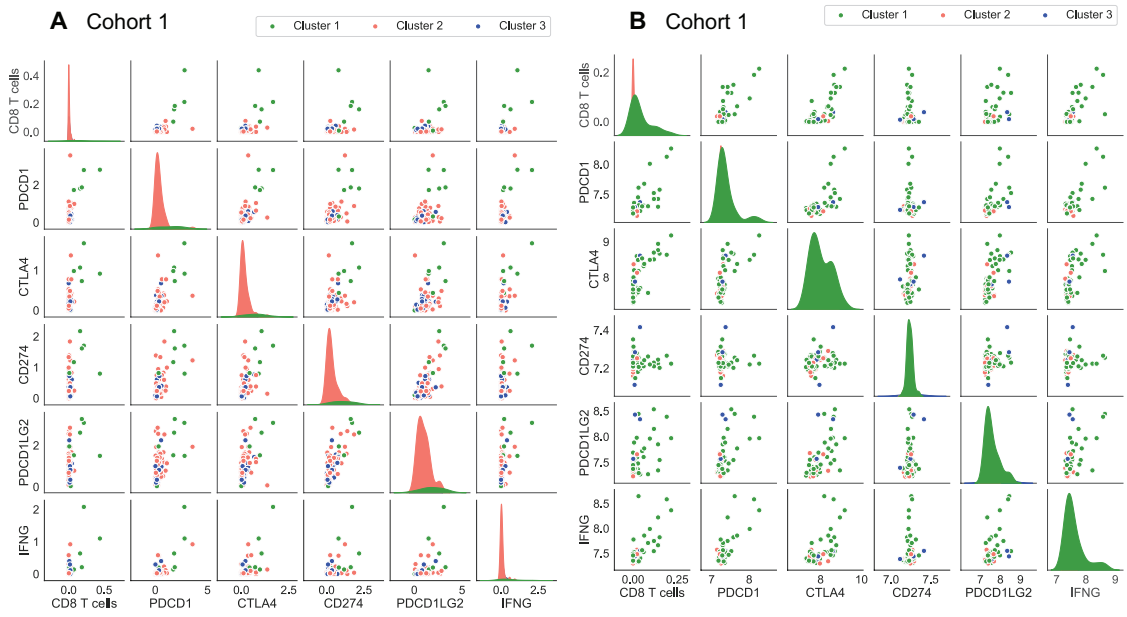

Figure 3.6. **Correlation and distribution of important proteins in the clusters.** Sub-figures A and B represent the cohort 1 and cohort 2 data sets respectively.

and NK cells [133]. Importantly, CD8 T cells frequency and PDCD1 gene expression are significantly correlated with IFNG gene expression so that cluster 1 has the highest level of IFNG compared the other clusters (Figure 3.5A2, B2). Beside these, we do not see any significant correlation between expression levels of CD274 and PDCD1LG2 genes, that encodes PD-L1 and PD-L2 respectively, with the expression levels of PDCD1 and IFNG, and the percentage of CD8 T cells in osteosarcoma tumors (Figure 3.6).

Tumor necrosis factor (TNF) is a cytokine that is mainly produced by Macrophages and has crucial roles in tumor development and tumor progression inducing apoptosis, necrosis, angiogenesis, immune cell activation, differentiation, and cell migration [134]. We notice that cluster 2 has the lowest TNF gene expression among clusters, while cluster 1 and 3 have roughly similar average expression of this gene

(Figure 3.5 A4, B4). In addition, we analyze gene IL1B that encodes cytokin protein Interleukin-1 beta (IL-1β), which is produced by activated Macrophages [135], and see that IL1B gene expression is the lowest in cluster 2 and the highest in cluster 3 (Figure 3.5 A6, B6).

Beside these, Insulin-like growth factor 1 (IGF-1) is a hormone that has important function in the development and function of many tissues and it has been used as a diagnostic marker for osteosarcoma [136, 137]. Similar to TNF and IL1B genes expression, cluster 2 has the lowest amount of IGF1 among other clusters (Figure 3.5 A5, B5). Furthermore, we examine RUNX2 oncogene that is associated with amplifications and it has been found to correlate to poor response to chemotherapy in osteosarcoma [138, 139]. In our analysis, cluster 1 has the lowest amount of RUNX2 gene and cluster 2 and 3 show almost similar expression of RUNX2 gene (Figure 3.5 A7, B7).

3.3 Discussion

The findings from analyses using estimated immune infiltrations in osteosarcoma have varied among studies, perhaps due to the small number of osteosarcoma tumors with available gene expression data in the literature. In this study, we find that infiltration of CD8 T cells, NK cells and M1 Macrophages have a positive association with prognosis, while infiltration of $\gamma\delta$ T cells, Mast cells, M0 Macrophages and Dendritic cells have a negative association with prognosis. Yu et al [121] also illustrates that high level of CD8 T cells is a good prognosis in their survival analysis, and results from [86] and [80] indirectly suggest the positive prognostic value of CD8 T cells. Tang et al [86] reports that CD8 T cells infiltration has a positive correlation with CXCR3 expression which is related to good prognosis. Khader et

al [80] shows that low-risk patients have high level of CD8 T cells and NK cells, which supports with our conclusion on these cells. In agreement with our findings on prognostic value of M1 Macrophages, Song et al [79] demonstrates that high level of M1 Macrophages is associated with good prognosis, while Zhang et al [85] and Tang et al [86] imply the same from their results. Our conclusion about M0 Macrophages aligns with the results from [85] and [86], but contradicts with the finding from [128] that abundance level of M0 Macrophages is positively correlated with survival.

The observable difference in outcomes between clusters are likely due to the relationship between immune infiltrates and prognosis in osteosarcoma, because we cluster the patients based on their immune composition. Our results indicate that cluster 2 has the worst outcome, while cluster 3 seems to have the best outcome among clusters. The main difference in immune composition between these two clusters is that cluster 2 has much higher percentage of M0 Macrophages, and lower percentage of CD8 T cells and M1 Macrophages than cluster 3. In general, we found that high levels of CD8 T cells and M1 Macrophages are associated with good prognosis, while a high level of M0 Macrophages correlates with poor prognosis in osteosarcoma. These results make sense because CD8 T cells are known to kill cancer cells directly [9, 12] and M1 Macrophages exhibit anti-tumor effects by producing cytokines that inhibit osteosarcoma growth [140]. These facts could also explain the observed differences between the outcomes of patients in clusters 2 and 3.

Meanwhile, cluster 1 has worse outcome than cluster 3, but better outcome than cluster 2. This could be due to the fact that cluster 1 has both high level of immune cells associated with good prognosis such as CD8 T cells and M1 Macrophages, and high level of immune cells associated with poor prognosis according to our results

such as $\gamma\delta$ T cells and Mast cells. High infiltration of mast cells have been associated with poor prognosis, low survival and increased metastasis in many cancers [141], while $\gamma\delta$ T cells show dual effects on cancer growth [142]. Both mast cells and $\gamma\delta$ T cells promote tumor development by supporting angiogenesis through angiogenic factors production [141, 142]. Mast cells also produce proteases, which lead to extracellular matrix degradation and tissue remodeling, and thus promote tumor growth [141]. $\gamma\delta$ T cells have been reported to secrete TGF-β [142], which is a pro-tumor cytokine in osteosarcoma [17, 20, 143]. Overall, the clinical outcomes of the clusters are consistent with our findings and biological knowledge on prognostic values of immune cells in osteosarcoma.

On the other hand, we did not observe any difference in age or metastasis at diagnosis between clusters. This can be explained by the lack of correlation between immune infiltrates and these clinical variables, which suggests that the immune composition of the primary tumor has no effect on age or metastasis status at diagnosis.

Our results indicate that cluster 2 has the worst outcome, while cluster 3 seems to have the best outcome among clusters. It is interesting to note that the immune patterns in cluster 2 and 3 are fairly similar (Figure 3.1D), with the main difference being cluster 2 has much higher percentage of M0 Macrophages and lower percentage of M2 Macrophages than cluster 3. Since we found that a high level of M0 Macrophages correlates with poor prognosis, this could be an explanation for the difference in outcome between clusters 2 and 3. Cluster 1, which has very different immune patterns from clusters 2 and 3, has worse outcome than cluster 3 but better outcome than cluster 2. This could be due to the fact that cluster 1 has both high level of immune cells with good prognosis such as CD8 T cells and M1 Macrophages, and high level of immune cells with poor prognosis such as $\gamma\delta$ T cells

and Mast cells. Overall, the clinical characteristics of the clusters are consistent with our findings about prognostic values of immune cells in osteosarcoma.

Immune checkpoints have an important role in the immune system to prevent autoimmune diseases, but unfortunately they can allow immune tolerance against tumors. PD-1 and CTLA-4 are the main checkpoints that tumor cells use to block immune system [144, 145, 146]. Blocking PD-1 pathway has improved oncological survival of several patients with metastatic cancers, including melanoma, renal cell carcinoma, and colon cancer [147, 148]. Also, targeting CTLA-4 in patients with metastatic melanomas demonstrates significant development about overall survival [149].

It has been reported that osteosarcoma patients treated with an anti PD-1 drug, Pembrolizumab, show some improvement in disease progression [150]. Combination of PD-1 and CTLA-4 blockade therapy in bone sarcoma have shown better response compared to single checkpoint inhibitor therapy [151]. Note, patients in cluster 1 have the highest expression levels of IFNG, PDCD1 and CTLA-4 that are significantly correlated with CD8 T cells (Figures 3.5 and 3.6), and it has been shown that INF-γ increases the CD8 T cells expansion [152]. Thus, patients in cluster 1 might respond well to combination of PD-1 and CTLA-4 blockade therapies.

It has been suggested in several studies that bacteria are able to activate anti-tumor immune responses [153, 154]. In a study with combination of Bacillus Calmette-Guerin(BCG) injection and tumor vaccine, 18% of the patients remained alive and disease-free and it has been reported that bacterial vaccine caused increased level of immunoregulatory cytokines such as TNF-α, IFN-γ, and IL1-β that might be involved in inducing tumor regression [155]. As a result, bacterial vaccine and inactivated tumor cells injection can be thought of as a treatment that activates anti-tumor immune responses [9]. In our results, cluster 2 has the lowest

amount of gene expression level of immunoregulatory cytokines TNF-α, IFN-γ, and IL1-β (Figure 3.5) so that tumors in this cluster might be treated with bacterial vaccine. Otherwise, targeting RUNX2 oncogene with chemotherapy is suggested as a new therapeutic approach to osteosarcoma patients in recent studies [156, 139] and cluster 2 has the highest amount of RUNX2 gene expression values compared to other clusters (Figure 3.5A7-B7) so with the help of further studies, tumors similar to those in cluster 2 also might be good candidates to treat with targeting RUNX2 in conjunction to standard chemotherapy.

Targeting tumor associated macrophages (TAM) is another alternative treatment method for osteosarcoma tumors and treatments that suppress M2 Macrophages phenotype or block the polarization of M1 Macrophages to M2 Macrophages have shown positive results in several studies [10, 157, 158, 159]. Thus, tumors in cluster 3, which has the highest amount of M2 macrophages (Figure 3.1D) might respond well to treatments that target TAMs.

CHAPTER 4

DATA-DRIVEN MATHEMATICAL MODEL OF OSTEOSARCOMA

As mentioned in chapter 1, even though several mathematical models have been developed to study the initiation and progression of many cancer types, there are currently no mathematical models for the progression of osteosarcoma, to the best of our knowledge. In chapter 3, we found that there were three distinct groups of immune patterns of osteosarcoma primary tumors. In this chapter, we develop a data-driven mathematical model of osteosarcoma based on the interactions between various components of the tumor microenvironment such as cancer cells, necrotic cells and immune cells, and use a system of ordinary differential equations (ODEs) to represent these interactions [92].

We then investigate the differences in the tumor growth of patients belonging in three distinct groups of immune patterns, which are obtained from chapter 3. We calculate the group-specific parameters from data in each group and use the previously estimated immune abundances as inputs in the mathematical model. Lastly, we analyze the dynamics of tumors in each group to find relationships that could potentially explain the effects of the tumor microenvironment on the progression of osteosarcoma tumors.

4.1 Materials and Methods

We built a kinetic model based on the key interactions between the immune system and osteosarcoma cells. In particular, we utilized a system of ordinary differential equations to study the changes in population of the various components of tumor microenvironment throughout time in units of days. For biochemical processes $A + B \rightarrow C$, we apply the mass action law $\frac{dC}{dt} = \lambda AB$, where λ is the production rate of C from A and B. For all the equations in our model, the symbol λ denotes proliferation, activation, or production rates, and the symbol δ denotes inhibition, decay, or death rates. The variables in the model are given in Table 4.1 and their interactions are illustrated in Figure 4.1.

4.1.1 Cytokines

We modeled the dynamics of cytokines through the rate at which they are produced and their natural decay. We assumed that cytokine production rates are proportional to the population of cells that produce them, similar to [160], and that cytokine decay rates are proportional to their own population, which is a common approach [55, 56, 57, 160, 161]. In order to simplify the system of equations, we combine some cytokines with similar functions and use the quasi-steady state assumption on other cytokines.

We combine TGF-β, IL-4, IL-10, and IL-13 as μ_1. TGF-β and IL-10 are secreted by helper T cells, M2 macrophages, and cancer cells [15, 16, 18, 140, 162, 163]. IL-4 and IL-13 are secreted by helper T cells and M2 macrophages [15, 140, 164]. Thus, we model the dynamics of μ_1 as:

$$\frac{d[\mu_1]}{dt} = \lambda_{\mu_1 T_h}[T_h] + \lambda_{\mu_1 M}[M] + \lambda_{\mu_1 C}[C] - \delta_{\mu_1}[\mu_1] \tag{4.1}$$

μ_2 consists of IL-6 and IL-17, where IL-6 is produced by M1 macrophages,

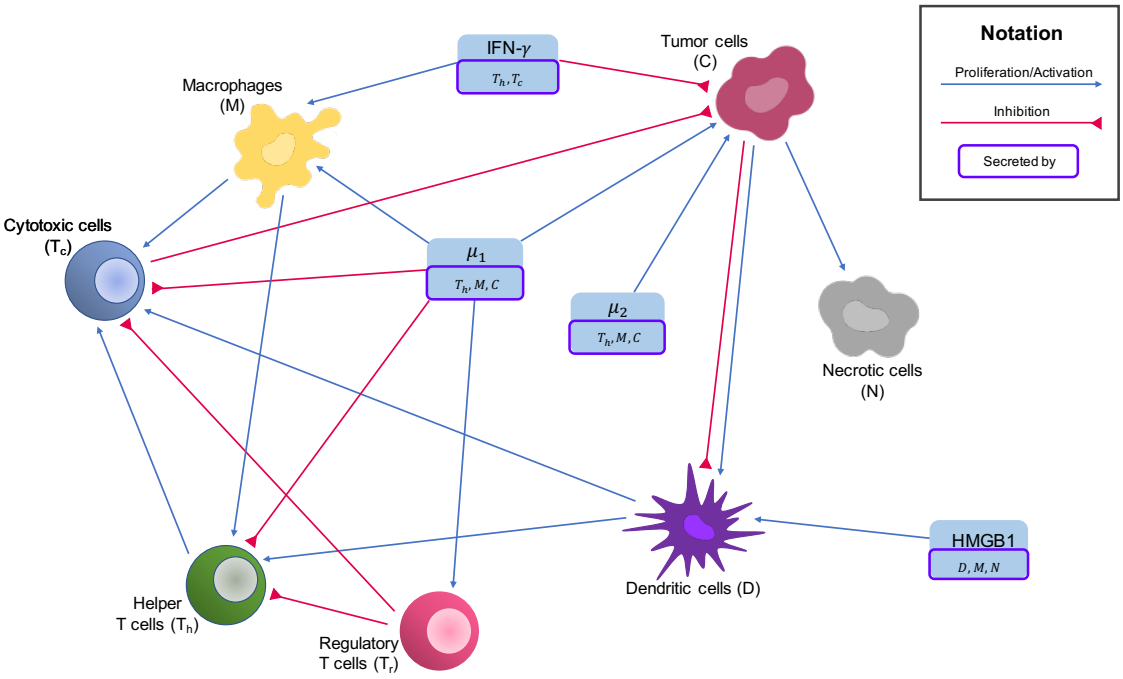

Figure 4.1. **Interaction network of the tumor microenvironment in osteosarcoma**. Activations and proliferations are shown by blue arrows, and inhibitions are indicated by red arrows.

Table 4.1: Model Variables. Names and descriptions of the variables used in the model.

Variable	Name	Description
T_N	Naive T-cells	
T_h	Helper T-cells	
T_C	Cytotoxic cells	includes CD8+ T-cells and NK cells
T_r	Regulatory T-cells	
D_N	Naive dendritic cells	
D	Activated dendritic cells	antigen presenting cells
M_N	Naive macrophages	includes naive macrophages and monocytes
M	Macrophages	includes M1 macrophages and M2 macrophages
C	Cancer cells	
N	Nectrotic cells	
H	HMGB1	
μ_1	Cytokines group μ_1	includes effects of TGF-β, IL-4, IL-10 and IL-13
μ_2	Cytokines group μ_2	includes effects of IL-6 and IL-17
I_γ	IFN-γ	

helper T cells, and cancer cells [16, 19, 140, 163, 165], and IL-17 is produced by helper T cells [15]. The corresponding equation for μ_2 is:

$$\frac{d[\mu_2]}{dt} = \lambda_{\mu_2 T_h}[T_h] + \lambda_{\mu_2 M}[M] + \lambda_{\mu_2 C}[C] - \delta_{\mu_2}[\mu_2] \tag{4.2}$$

IFN-γ is secreted by helper T cells, cytotoxic T cells, and natural killer cells [12, 15, 166]. As a result, the equation for IFN-γ is written as:

$$\frac{d[I_\gamma]}{dt} = \lambda_{I_\gamma T_h}[T_h] + \lambda_{I_\gamma T_c}[T_c] - \delta_{I_\gamma}[I_\gamma] \tag{4.3}$$

HMGB1 is passively released by necrotic cells [8, 30, 167, 168] and actively

released by macrophages and dendritic cells [29, 30, 31, 32, 167, 169], leading to the following equation:

$$\frac{d[H]}{dt} = \lambda_{HM}[M] + \lambda_{HD}[D] + \lambda_{HN}[N] - \delta_H[H] \tag{4.4}$$

We use the quasi-equilibrium state assumption on the other cytokines and estimate them to be proportional to the number of cells that produce them. IL-12 and IL-23 are both secreted by M1 macrophages and dendritic cells [12, 15, 140, 162, 163, 170]; therefore, we model the concentration of these cytokines as:

$$[\text{IL-12}] \approx c_1 \times [M] + c_2 \times [D] \tag{4.5}$$

$$[\text{IL-23}] \approx c_3 \times [M] + c_4 \times [D] \tag{4.6}$$

where c_1, c_2, c_3, and c_4 are constants.

4.1.2 Cells in the Tumor Microenvironment

Since mature immune cells are differentiated from naive immune cells, we model the population of each mature immune cell to be proportional to its respective naive immune cell, where the proportion is determined by the cells/cytokines that activate the naive cells. Similar to the cytokine equations, for each mature immune cell, we also include a natural death rate δ_{cell}.

Macrophages

Since macrophages have many phenotypes and are constantly changing their phenotype, we model all macrophages together as one variable to avoid overly great complexity. M1 and M2 macrophages are differentiated from naive macrophages or monocytes. M1 macrophages are activated by IFN-γ [18, 162, 163], while M2 macrophages are activated by IL-4, IL-10, and IL-13 [16, 162, 163, 171], where

IL-4, IL-10, and IL-13 belong to μ_1. Therefore, we can write the dynamics of macrophages as:

$$\frac{d[M]}{dt} = \left(\lambda_{MI_\gamma}[I_\gamma] + \lambda_{M\mu_1}[\mu_1]\right)[M_N] - \delta_M[M] \tag{4.7}$$

By taking into account the activations from Equation (4.7) and introducing the independent naive macrophage/monocyte production parameter A_{M_N}, we have the equation for naive macrophages/monocytes:

$$\frac{d[M_N]}{dt} = A_{M_N} - \left(\lambda_{MI_\gamma}[I_\gamma] + \lambda_{M\mu_1}[\mu_1]\right)[M_N] - \delta_{M_N}[M_N] \tag{4.8}$$

T Cells and NK Cells

We model the following subtypes of T cells: helper T cells, regulatory T cells, and cytotoxic cells, where cytotoxic cells include cytotoxic T cells and natural killer cells.

Helper T cells are activated by dendritic cells, IL-12, and IL-23 [9, 12, 15, 172], and are inhibited by regulatory T cells, IL-10, and TGF-β [12, 14, 173, 174], resulting in the equation:

$$\frac{d[T_h]}{dt} = \left(\lambda_{T_hM}[M] + \lambda_{T_hD}[D]\right)[T_N] - \left(\delta_{T_hT_r}[T_r] + \delta_{T_h\mu_1}[\mu_1] + \delta_{T_h}\right)[T_h] \tag{4.9}$$

Regulatory T cells are activated by IL-10 and TGF-β [12, 175], hence their dynamics are modeled by:

$$\frac{d[T_r]}{dt} = \lambda_{T_r\mu_1}[\mu_1][T_N] - \delta_{T_r}[T_r] \tag{4.10}$$

Cytotoxic cells (cytotoxic T cells and NK cells) are activated by helper T cells, dendritic cells and IL-12 [8, 9, 10, 12, 16, 176, 177] and are inhibited by regulatory T cells, IL-10, and TGF-β [9, 13, 163, 175]. The corresponding equation is:

$$\frac{d[T_c]}{dt} = \left(\lambda_{T_cT_h}[T_h] + \lambda_{T_cM}[M] + \lambda_{T_cD}[D]\right)[T_N] - \left(\delta_{T_cT_r}[T_r] + \delta_{T_c\mu_1}[\mu_1] + \delta_{T_c}\right)[T_c] \tag{4.11}$$

Combining all the activations from Equations (4.9)–(4.11) as well as adding parameter A_{T_N} for the independent production rate of naive T cells, we obtain the equation for naive T cells:

$$\begin{aligned}\frac{d[T_N]}{dt} =& A_{T_N} - \left(\lambda_{T_hM}[M] + \lambda_{T_hD}[D]\right)[T_N] - \lambda_{T_r\mu_1}[\mu_1][T_N] \\ & - \left(\lambda_{T_cT_h}[T_h] + \lambda_{T_cM}[M] + \lambda_{T_cD}[D]\right)[T_N] - \delta_{T_N}[T_N]\end{aligned} \tag{4.12}$$

Dendritic Cells

Dendritic cells are activated by cancer cells and HMGB1 [9, 29, 30, 31, 32]. However, cancer cells can also promote apoptosis in dendritic cells through many tumor-derived factors, such as gangliosides, neuropeptides, etc. [170]. By introducing the independent production rate of naive dendritic cells A_{D_N}, we can describe the dynamics of naive and mature dendritic cells with the following system:

$$\frac{d[D]}{dt} = \left(\lambda_{DC}[C] + \lambda_{DH}[H]\right)[D_N] - \left(\delta_{DC}[C] + \delta_D\right)[D] \tag{4.13}$$

$$\frac{d[D_N]}{dt} = A_{D_N} - \left(\lambda_{DC}[C] + \lambda_{DH}[H]\right)[D_N] - \delta_{D_N}[D_N] \tag{4.14}$$

Cancer Cells

Osteosarcoma cells are typically of osteoblastic origin and are characterized by abnormally high proliferation and low apoptosis. We denote the high proliferation rate of cancer cells as λ_C.

Osteosarcoma growth is promoted by IL-6, IL-17, and TGF-β [12, 17, 18, 19, 20, 143, 165, 178]. Tumor cells are killed by cytotoxic cells [9, 179, 180], while their growth is inhibited by IFN-γ [9, 10, 166]. In the mathematical modeling of cancer, it is common to estimate the growth to be proportional to $[C]\left(1 - \frac{[C]}{C_0}\right)$, where C_0 is the carrying capacity [181, 182]. As a result, we have the following equation for

cancer cells:

$$\begin{aligned}\frac{d[C]}{dt} = &\left(\lambda_C + \lambda_{C\mu_1}[\mu_1] + \lambda_{C\mu_2}[\mu_2]\right)[C]\left(1 - \frac{[C]}{C_0}\right) \\ &- \left(\delta_{CT_c}[T_c] + \delta_{CI_\gamma}[I_\gamma] + \delta_C\right)[C]\end{aligned} \tag{4.15}$$

Necrotic Cells

Necrotic cells, which are cells that go through the process of necrotic cell death, are promoted by cancer cells since, when cancer cells are killed by cytotoxic cells, a proportion of them become necrotic cells. In particular, the "production" rate of necrotic cells can be modeled as a fraction of the dying cancer cells, resulting in the following dynamics:

$$\frac{d[N]}{dt} = \alpha_{NC}\left(\delta_{CT_c}[T_c] + \delta_{CI_\gamma}[I_\gamma] + \delta_C\right)[C] - \delta_N[N] \tag{4.16}$$

4.1.3 Data of the Model

In the last chapter, we applied CIBERSORTx B-mode on the gene expression data sets from two cohorts, TARGET and GSE21257, to estimate the immune cell abundances within osteosarcoma tumors. Then, K-means clustering was applied on the estimated immune cell fractions. As a result, we found that there were three distinct immune patterns of osteosarcoma tumors. In this chapter, we use the same cluster assignment for the TARGET data with 88 samples and use our mathematical model to study the dynamics of the tumor microenvironment of each cluster from the initial time of diagnosis until reaching their steady state. The general workflow of this study is described in Figure 4.2, and the average immune fractions of various cell types in each cluster are shown in Figure 4.3, where the vertical bars denote the 95% confidence intervals.

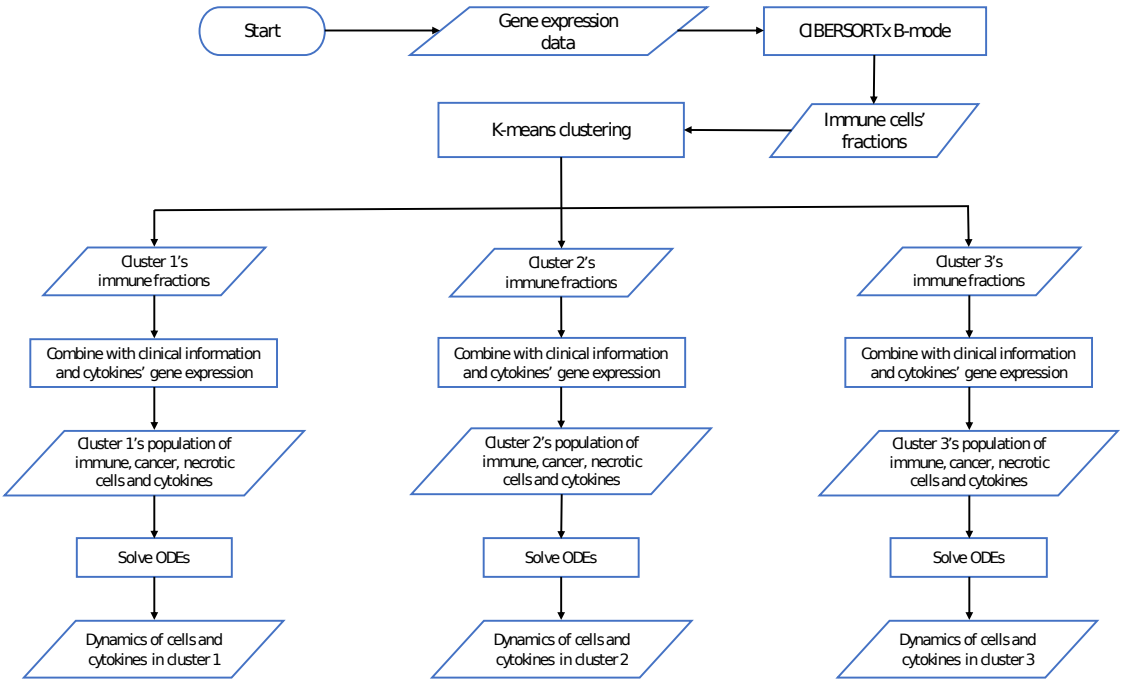

Figure 4.2. **The general workflow of this study.** Given the gene expression data of tumors, immune cell fractions were estimated using CIBERSORTx B-mode. Then, K-means clustering was applied to find three clusters with distinct immune compositions. For each cluster, the populations of immune, cancer, and necrotic cells were derived from immune fractions and clinical information. These cell populations and cytokine expression levels were used as input (either as the initial conditions or steady states) in the system of ODEs to find the dynamics of the components of the tumor microenvironment in each cluster.

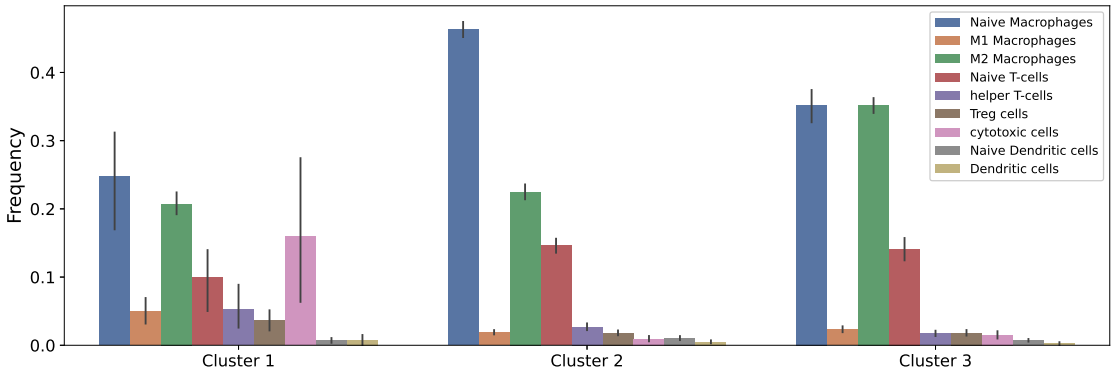

Figure 4.3. **The immune cell fractions used in the model.** Clusters were derived based on differences in 22 immune cell types of osteosarcoma tumors.

The outputs of CIBERSORTx only provide the fractions of each immune cell within the tumor tissue; however, we need the number of immune cells along with the number of cancer and necrotic cells as inputs to our model. Thus, we download the supplementary data of the TARGET project, which has information on the percentage of normal, stroma, tumor, and necrotic cells of each sample. We use the percentage of normal cells to represent the percentage of total immune cells in the sample.

First, we convert the immune cell fractions to the immune cell population by multiplying the fractions with a scaling factor α_{dim}. Then, knowing the percentage of total immune cells, cancer cells, and necrotic cells, we derive the population of cancer and necrotic cells from the population of total immune cells. For example, given the total immune population I, the cancer and necrotic cell abundance can be calculated as

$$C = I \times \frac{\%\text{ of cancer cells}}{\%\text{ of total immune cells}} \tag{4.17}$$

$$N = I \times \frac{\%\text{ of necrotic cells}}{\%\text{ of total immune cells}} \tag{4.18}$$

where C and N are the cancer and necrotic cell population, respectively.

To choose a reasonable value for α_{dim}, we first estimate the average osteosarcoma tumor volume. We find the mean volume of Ewing sarcomas to be 275 mL based on the tumor volumes given in [183], and Ewing sarcoma has been reported to have a similar volume to osteosarcoma [184]. Thus, we estimate the average osteosarcoma's volume to be 275 mL.

Osteoblasts, which are the cells of origin of osteosarcoma, have a diameter of 20–50 m [185]; therefore, we approximate osteosarcoma cells to have an average diameter of 35 m, resulting in an average of 6.4×10^9 osteosarcoma cells in osteosarcoma tumors. We then choose $\alpha_{dim} = 1.765 \times 10^8$ to match the average number of cancer cells among all patients in our data to 6.4×10^9 cells. However, it is important to note that α_{dim} is simply a scaling factor and does not have any effects on the dynamics of cells or on the relative cell abundance between clusters.

4.1.4 Parameter Estimation

Some parameters of our model, such as the decay/death rates of immune cells and cytokines, are taken from available research (more details in Appendix A.2.1), while others are estimated. We follow the common approach from mathematical biological models to use assumptions on the steady state values of the system to derive those unknown parameters [186, 187]. In particular, we make the assumption that after a tumor reaches a very large size, the immune variation within the tumor microenvironment is minuscule, and we denote this state as the steady state of our system.

Different immune patterns of tumors, such as high or low levels of helper and cytotoxic T cells in one group versus another group, indicate that the activation rates of different T cell sub-types from naive T cells vary from one group of tumors to another group. Hence, many parameters of the model, such as the activation

rates of T cell sub-types, depend on the tumor immune profile, and therefore we estimate the parameters separately for each cluster.

We assume the samples with a large number of cancer cells are at the steady state. For each cluster, we use the 85th percentile of cancer abundance as the cutoff, and calculate the steady state values for the cluster by averaging the values from samples that have more cancer cells than this cutoff. Table 4.2 shows the steady state values of every cluster.

Table 4.2: Steady-state abundance of cells and cytokines.

Cluster	$\mathbf{M_N^\infty}$	$\mathbf{M^\infty}$	$\mathbf{T_N^\infty}$	$\mathbf{T_h^\infty}$	$\mathbf{T_r^\infty}$
1	6.236×10^6	1.977×10^7	4.926×10^6	7.092×10^6	3.675×10^6
2	3.248×10^7	1.842×10^7	1.047×10^7	1.973×10^6	8.673×10^5
3	1.944×10^7	2.698×10^7	1.368×10^7	1.205×10^6	1.405×10^6
	$\mathbf{T_c^\infty}$	$\mathbf{D_N^\infty}$	$\mathbf{D^\infty}$	$\mathbf{C^\infty}$	$\mathbf{N^\infty}$
1	2.292×10^7	4.826×10^5	9.865×10^5	1.343×10^{10}	3.764×10^8
2	3.155×10^5	8.927×10^5	7.135×10^5	1.604×10^{10}	4.257×10^8
3	1.802×10^6	4.591×10^5	3.732×10^5	1.340×10^{10}	1.544×10^9
	$\mathbf{I_\gamma^\infty}$	$\boldsymbol{\mu_1^\infty}$	$\boldsymbol{\mu_2^\infty}$	$\mathbf{H^\infty}$	
1	0.868	21.510	2.067	5.076	
2	0.049	20.714	1.611	4.948	
3	0.263	23.663	1.371	4.453	

Our assumption above asserts that the rate of change of our model's variables is 0 at the steady state, or equivalently $\frac{dX}{dt} = 0$ at the steady state. With the additional assumptions in Appendix A.2.1, as well as knowing the steady state values of our model's variables, we can derive parameter values for each cluster using the *fsolve* function from the SciPy package in Python. The parameter values

for each cluster are given in Table A.1.

4.1.5 Non-Dimensionalization

To remove the scale dependence and obtain additional numerical stability, we apply non-dimensionalization on all equations of our system. For a model variable X converging to the steady state value X^{∞}, we create a non-dimensional variable $\overline{X}$ such that $\overline{X} = \frac{X}{X^{\infty}}$. Then, $\overline{X}$ satisfies the equation $\frac{d\overline{X}}{dt} = F(\overline{X}, \overline{\theta}, t)$, where $\overline{\theta}$ is the vector of non-dimensional parameters. The full system of non-dimensionalized equations are given in Appendix A.3.

To solve the non-dimensional dynamical system for each cluster, we apply the *odeint* function from the SciPy package [188], with the initial conditions from a data point of interest from the TARGET data set.

4.1.6 Sensitivity Analysis

To evaluate the quality of our parameters through how they affect the dynamics of the system, we perform a global gradient-based sensitivity analysis on all parameters of our system.

For the non-dimensional system $\frac{d\overline{X}}{dt} = F(\overline{X}, \theta, t)$ with N parameters $\theta = \theta_1, \ldots, \theta_N$, the (first order) sensitivity s_i of parameter θ_i was defined as the gradient of the model output with respect to the parameter [189]:

$$s_i = \frac{\partial \overline{X}}{\partial \theta_i} \tag{4.19}$$

We calculate the sensitivity s_i for each parameter at the steady state of the equation for two quantities of interest: cancer cell abundance and total cell abundance. Consider the general steady state system as $F(X^*, \theta) = 0$, with X^* being the equilibrium values of our model's variables. The sensitivity vector s can be

obtained analytically by differentiating the steady-state equation with respect to parameter vector θ, that is,

$$\nabla F(X^*, \theta)\frac{\partial X^*}{\partial \theta} + \frac{\partial F(X^*, \theta)}{\partial \theta} = 0 \tag{4.20}$$

where $\nabla F(X^*, \theta)$ is the Jacobian matrix of $F(X^*, \theta)$ with respect to X. Then, to compute sensitivity vector s at equilibrium, or equivalently $\frac{\partial X^*}{\partial \theta}$, we simply need to numerically invert $\nabla F(X^*, \theta)$.

Generally, s_i varies for different values of the parameter set; thus, we define the local sensitivity S_i of parameter θ_i for a chosen neighborhood $\Omega(\theta)$ of the given parameter set as

$$S_i = \int_\Omega s_i(\theta) d\theta \tag{4.21}$$

where the integral is evaluated numerically using sparse grid points [190, 191].

Since we made many assumptions to derive the parameter values for our model and different assumptions can lead to different parameter values, we vary these assumptions by a scaling factor of 0.01 to 100 for K times and obtain the local sensitivity S_i^k, with $k = 1, \ldots, K$, for parameter θ_i derived from the k^{th} set of new assumptions. Then, the global sensitivity $\mathbf{S}_i$ of parameter θ_i is a weighted average of the local sensitivities S_i^k for $k = 1, \ldots, K$:

$$\mathbf{S}_i = \sum_{k=1}^{K} w_k S_i^k \tag{4.22}$$

where w_k is chosen so that the parameter values that are closer to the original parameter set have larger weights and the parameter values that are very different from the original parameter set have smaller weights. This method of choosing w_k is based on the idea of the weighted average of local sensitivities in [189].

4.2 Results

We obtain the dynamics of the components in the tumor microenvironment by solving the above mentioned system of ODEs with parameters derived from the cancer patient data using the steady state assumption as mentioned in Section 4.1.4. Given non-negative initial conditions and non-negative parameters, the solution of the systems remains non-negative and globally bounded (Appendices A.1.2 and A.1.3).

4.2.1 Dynamics of the Tumor Microenvironment

We are interested in exploring the dynamics of different components of the osteosarcoma microenvironment as well as the difference in cancer progression between clusters. Hence, we want to model the dynamics with similar initial cancer populations among clusters. We first choose the sample with the smallest cancer population in cluster 1, and then choose a sample from cluster 2 and 3 that has the most similar cancer population to the chosen sample in cluster 1. We use these samples as the initial conditions for their corresponding cluster. Table 4.3 shows the dimensionless initial condition values of each cluster.

We observe that, as the cancer population grows, helper T cells, dendritic cells, cytotoxic cells, and IFN-γ populations first increase and then decrease over time. This makes sense biologically since, in the early stage of cancer, naive dendritic cells come in contact with tumor antigens, inducing the activation and increase in the number of dendritic cells [7, 9]. Dendritic cells present tumor antigens to helper T cells and cytotoxic cells and activate them [192], resulting in an increase of these cells. Helper T cells and cytotoxic cells then produce IFN-γ [12, 15, 166], leading to this cytokine's increased abundance.

Table 4.3: The non-dimensional initial conditions for each cluster.

Cluster	$\mathbf{M_N/M_N^\infty}$	$\mathbf{M/M^\infty}$	$\mathbf{T_N/T_N^\infty}$	$\mathbf{T_h/T_h^\infty}$	$\mathbf{T_r/T_r^\infty}$	$\mathbf{T_c/T_c^\infty}$	$\mathbf{D_N/D_N^\infty}$
1	2.367	1.005	0.019	0.794	0.764	0.828	1.122
2	0.954	0.753	1.299	1.451	2.313	0.062	0.071
3	0.866	1.104	0.572	0.340	0.484	0	1.643
	$\mathbf{D/D^\infty}$	$\mathbf{C/C^\infty}$	$\mathbf{N/N^\infty}$	$\mathbf{I_\gamma/I_\gamma^\infty}$	$\boldsymbol{\mu_1/\mu_1^\infty}$	$\boldsymbol{\mu_2/\mu_2^\infty}$	$\mathbf{H/H^\infty}$
1	0	0.020	0.160	2.394	1.104	1.806	1.059
2	0.693	0.005	0.018	0.859	1.307	3.259	0.988
3	0	0.014	0.0008	0.276	1.030	1.296	1.284

The switch in dynamics from increasing to decreasing in dendritic cells, helper T cells, cytotoxic cells, and IFN-γ occurs around the same time that cancer cells start growing fast. Contrastingly, the number of regulatory T cells decreases when these cells increase and increases when these cells decrease. Hence, regulatory T cells start increasing in density when the tumor is at its peak of growing. Regulatory T cells have the role of modulating the immune system and consequently promote tumor growth; therefore, we can expect the opposite dynamics to anti-tumor immune cells and cytokines, such as dendritic cells, helper T cells, cytotoxic cells, and IFN-γ. In general, it is important to study this switch in the dynamics since it can be used as the predictor of the highest growth of cancer cells during tumor development.

On the other hand, the macrophage population first decreases and then increases during osteosarcoma progression, while necrotic cells, HMGB1, along with the cytokine groups μ_1 and μ_2 increase in population as cancer cells grow. As both μ_1 and μ_2 support tumor growth, their population growth over time could

contribute to the fast progression of osteosarcoma. Necrotic cells are mainly cancer cells that were killed by cytotoxic cells or IFN-γ; thus, it is reasonable to see their population grow over time. As a result, HMGB1, which is largely produced by necrotic cells, increases in abundance as the tumor progresses.

Cluster 2's cancer cells begin by growing more slowly than cluster 1; however, at around 500 days, they start growing very fast and end up having the highest cancer population at the steady state out of all clusters. The results of chapter 3 based on the clinical information of the TARGET dataset also indicate that patients in cluster 2 have the worst survival outcomes among the three clusters.

Figure 4.4 shows that cluster 2 has the lowest number of cytotoxic cells, macrophages, and IFN-γ and the highest number of naive macrophages during tumor progression. A high population of cytotoxic cells and IFN-γ are generally associated with a good prognosis because they directly kill cancer cells, while a high level of naive macrophages have been found in chapter 3 to associate with poor prognosis. Cluster 2 also has the slowest growth rate of necrotic cells. A high number of necrotic cells means many cancer cells have been killed by the immune system and is an indication of a good prognosis. Thus, cluster 2, with a slow growth rate of necrotic cells, high growth rate of cancer, and the highest cancer population at the steady state, has a poor prognosis based on our model's dynamics.

Cluster 3 has the slowest cancer growth rate among all clusters and a smaller cancer population at the steady state compared with cluster 2. Cluster 3's necrotic cells have the fastest growth rate and the highest population at the steady state out of the 3 clusters. Hence, the dynamics of cluster 3 appear to be the most favorable. This is in agreement with the findings on the survival outcomes of cluster 3 in the previous chapter.

Cluster 3 has the smallest amount and the slowest growth rate of the cytokine

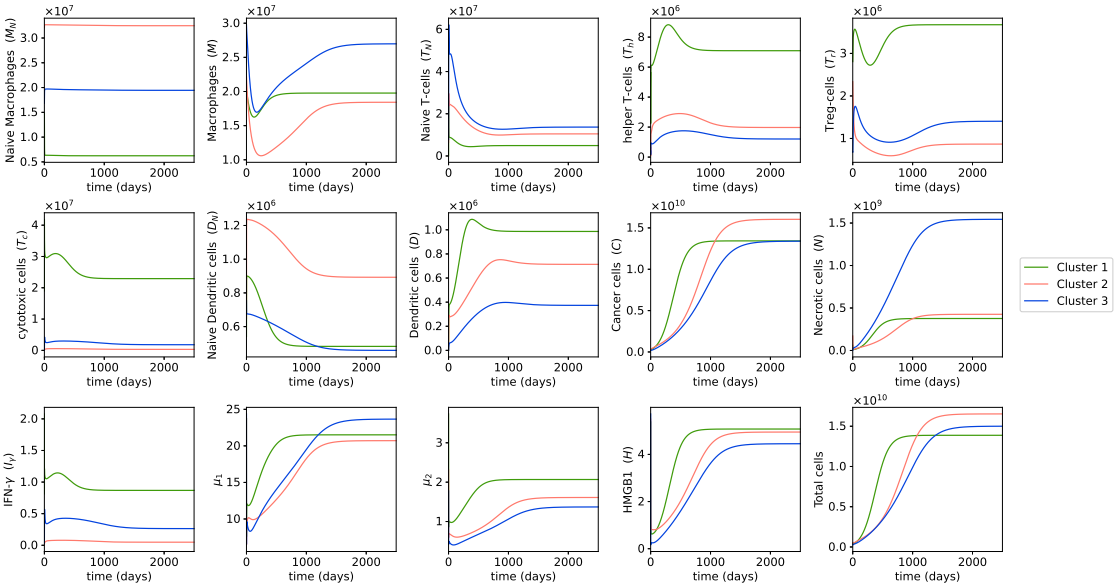

Figure 4.4. **Dynamics of cells and cytokines in osteosarcoma tumors**. Evolution of the cells and cytokine population in the model is plotted over the time in units of days. This figure shows the dynamics of the variables of the model starting from the time of the first diagnosis of small tumors in each cluster until reaching their steady state values, i.e., the average values of the largest tumors in the same cluster. The different color lines describe the dynamics of different clusters.

group μ_2, which has tumor-promoting effects, both initially and at the steady state (Figure 4.4). Interestingly, cluster 3 also has the lowest population of helper T cells and dendritic cells over time. These two cells are known to correlate with good prognoses. If we were to simply look at the immune composition of the patients in cluster 3, we might make the wrong prediction on their prognosis due to the low abundance of certain immune cells with good prognostic values. Therefore, it is important to take into consideration the interaction between immune cells and cancer cells, and investigate the dynamics of cancer in addition to studying the immune composition.

Cluster 1 has a high cancer growth rate from the beginning and thus its cancer population reaches the steady state faster than the other clusters. However, its cancer cells do not reach as high population at the steady state as the cancer cells in cluster 2. Cluster 1 has the highest levels of both immune cells and cytokines with good prognoses, including cytotoxic cells, helper T cells, dendritic cells, and IFN-γ, and those with poor prognoses, such as regulatory T cells and μ_2 during tumor progression. Thus, it is again necessary to look at the interactions within the tumor microenvironment for such clusters.

We observe that μ_1 and μ_2 grow fast and reach the steady states very quickly in cluster 1. Since both μ_1 and μ_2 promote tumor proliferation, this could be the reason why cancer cells quickly reach the steady state in cluster 1. Overall, since cluster 1 has a lower cancer population at the steady state compared with cluster 2 but a higher cancer growth rate than cluster 3, its cancer dynamics are worse than cluster 3 but better than cluster 2, which aligns with the results of chapter 3.

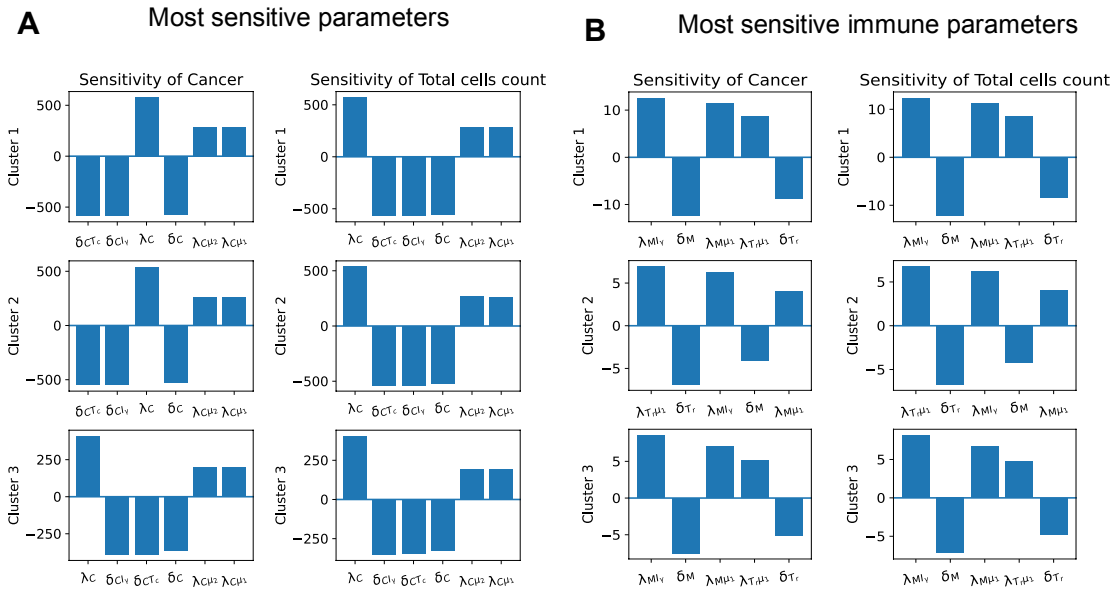

Figure 4.5. **Sensitivity analysis.** (A) The sensitivity level of the most sensitive parameters for cancer and total cell population at the steady state. (B) The most sensitive parameters associated with immune cells. The most sensitive parameters for each cluster are shown in each row of plots.

4.2.2 Sensitivity Analysis

We perform global sensitivity analysis with parameters derived from patient data with the steady state assumption in each cluster. The sensitivity analysis is performed on the dimensionless system, and evaluated at the steady states. We are interested in finding which parameters in our system strongly affect the growth of tumors, and thus we use the cancer population and total cell population as variables of interest in the sensitivity analysis.

Figure 5.3A presents the six most sensitive parameters in every cluster. Since we also want to study the effects of the immune system on cancer progression, we look at the five most sensitive parameters from the immune cells equations as well. Therefore, we plot the top five most sensitive parameters excluding the parameters from the cancer cell Equation (4.15) and necrotic cell Equation (4.16) (Figure 5.3B).

The most sensitive parameters across the three clusters are the cancer proliferation and inhibition parameters in the cancer Equation (4.15). As expected, an increase in any of the cancer proliferation parameters ($\lambda_C, \lambda_{C\mu_1}, \lambda_{C\mu_2}$) results in an increase in the number of cancer cells, and an increase in any cancer inhibition parameters ($\delta_{CT_c}, \delta_{CI_\gamma}, \delta_C$) results in a decrease in the number of cancer cells. It is worth noting that all sensitive parameters presented in Figure 5.3 have similar effects on cancer populations as on total cell populations.

The most sensitive immune parameters are activation and the decay rates of macrophages and regulatory T cells for all clusters. An increase in any activation rates of macrophages and regulatory T cells leads to higher cancer and total cell numbers, while an increase in their decay rates causes a decrease in these quantities of interest. This implies that both macrophages and regulatory T cells have tumor-promoting effects.

Since regulatory T cells inhibit helper T cells and cytotoxic cells, they hinder IFN-γ production and, thus, down-regulate cytotoxic cells and IFN-γ's ability to kill cancer cells. Macrophages, on the other hand, have both anti-tumor phenotype (M1 macrophages) and pro-tumor phenotype (M2 macrophages). However, the predominant portion of macrophages in the patient data across all three clusters is M2 macrophages (Figure 4.3), which can cause the main effect of macrophages in our model to be pro-tumor.

4.2.3 Dynamics with Varying Assumptions

Since we made some assumptions in order to derive the parameter values for each cluster, we want to see how the dynamics of cancer population would change when we vary these assumptions. Based on the results of the global sensitivity analysis, we determine that the parameters in the equations of cancer cells, macrophages, and regulatory T cells are the most sensitive parameters. We vary each assumption relating to these sensitive parameters (Equations (A.15)–(A.19)) by five times in both directions (scale five-times bigger or five-times smaller) and observe how the progression of cancer changes with the new assumptions (Figure 4.6). For example, since λ_C and $\lambda_{C\mu_1}$ are sensitive parameters, we vary the assumption $\lambda_C = 40\lambda_{C\mu_1}\mu_1^{\text{mean}}$ (Equation (A.16)) by five times, resulting in the following new assumptions:

$$\lambda_C = 200\lambda_{C\mu_1}\mu_1^{\text{mean}}, \quad \lambda_C = 8\lambda_{C\mu_1}\mu_1^{\text{mean}}, \tag{4.23}$$

where the cancer dynamics with the original assumption (Equation (A.16)) is the left plot in Figure 4.6A (scale = 1), and the cancer dynamics with the new assumptions (Equation (4.23)) are the middle and right plots in Figure 4.6A (scale = 1/5 and scale = 5).

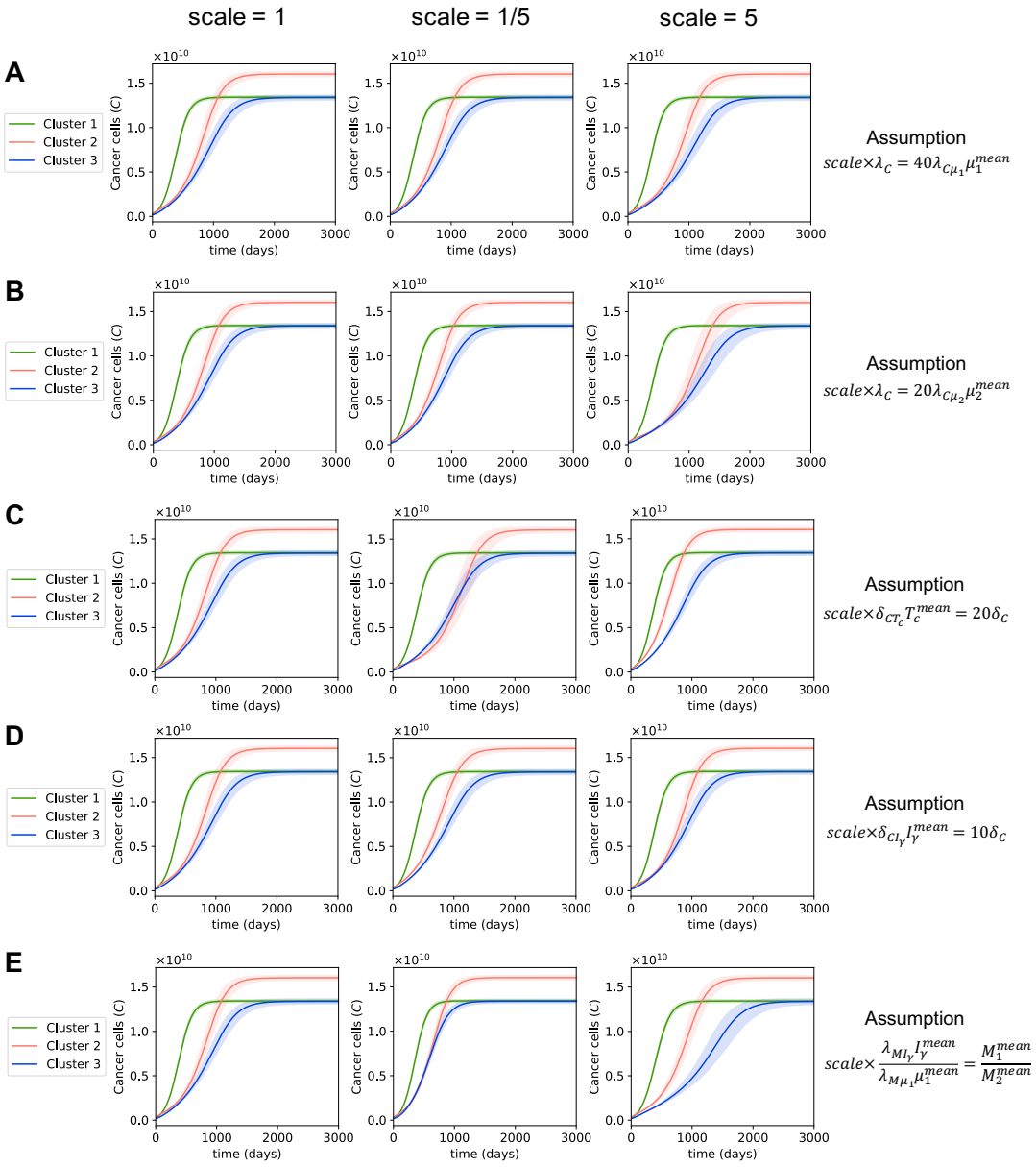

Figure 4.6. **The dynamics of cancer when the assumptions of the sensitive parameters are varied.** (A-E) The cancer growth of all three clusters for each assumption of sensitive parameters. The left plot in every sub-figure is the original cancer dynamics, the middle and right plots are the cancer dynamics obtained when the given assumption is scaled by 1/5 and by 5, respectively.

We notice that when we vary the assumptions of the most sensitive parameters, the time for the cancer population to reach the steady state changes by a relatively small amount; however, the overall observation of the cancer dynamics between clusters does not change (Figure 4.6). That is, these different assumptions lead to the same observations: cluster 1's cancer population reaches a steady state the fastest among all clusters, cluster 2's tumors grow slower than cluster 1's at first but then begin growing fast and result in the highest steady state population, and cluster 3 has the most favorable cancer progression with the slowest growth of cancer cells and one of the lowest steady state cancer populations.

The largest changes in the dynamics of cancer are due to the assumptions for the activation rates of macrophages (Figure 4.6E):

$$\frac{\lambda_{MI_\gamma} I_\gamma^{\text{mean}}}{\lambda_{M\mu_1} \mu_1^{\text{mean}}} = \frac{M_1^{\text{mean}}}{M_2^{\text{mean}}}.$$

This assumption is based on the fact that M1 and M2 macrophages are activated by IFN-γ and μ_1, respectively, and thus the ratio of macrophages activated by IFN-γ to macrophages activated by μ_1 should be approximately equal to the ratio of M1 to M2 macrophages. This is a reasonable assumption that uses patient data to derive the activation rates of macrophages. We expect to see the ground truth ratio of macrophages activated by IFN-γ to macrophages activated by μ_1 to be close to our assumption, rather than to differ by five times. Therefore, it is very unlikely to observe cancer dynamics, such as in the middle and right plots in Figure 4.6E with our data. On the other hand, the assumptions for the death rate of cancer by IFN-γ and the apoptosis rate of cancer, $\delta_{CI_\gamma} I_\gamma^{\text{mean}} = 10\delta_C$, appear to have a negligible to no impact on cancer progression (Figure 4.6D).

The shaded regions in Figure 4.6 denote the changes in dynamics when we varied the most sensitive parameters ($\lambda_C, \lambda_{C\mu_1}, \lambda_{C\mu_2}, \delta_{CT_c}, \delta_{CI_\gamma}, \delta_C, \lambda_{MI_\gamma}, \lambda_{M\mu_1}, \delta_M, \lambda_{T_r\mu_1}$,

and δ_{T_r}) by 10% in negative and positive directions. We observe that varying the most sensitive parameters by 10% does not create large changes to the cancer dynamics. Overall, Figure 4.6 shows that, when we change the assumptions of the most sensitive parameters or vary the sensitive parameters themselves, the observations we made about cancer development between clusters in Section 4.2.1 are not affected. Furthermore, even though several assumptions were made to estimate the parameters, the dynamics of cancer do not greatly depend on these assumptions.

4.2.4 Dynamics with Different Initial Conditions

For each cluster, we also look at the dynamics with different initial conditions from the different samples within that cluster (Figure 4.7). We observe that different initial conditions in a cluster lead to similar growth patterns of cancer. This makes sense since the dynamics are determined by the parameters of the ODE system, which are derived from the patient data through the steady state assumption in each cluster. As a result, the cancer growth rates and patterns are similar among patients within the same cluster but different among patients in different clusters. Thus, if we know which cluster a patient belongs to, we can predict their cancer growth more accurately than by using the same cancer progression model for all patients.

To verify that the parameters in each cluster are what drives the dynamics of the cluster, we examine the dynamics of each cluster with the initial conditions from other clusters (Figure 4.8). In particular, we plot dynamics of cluster 1 with the initial conditions in Table 4.3 from clusters 2 and 3. These cross-cluster initial conditions quickly converge to the same dynamics, confirming that the dynamics in each cluster are more influenced by the parameters rather than by the initial

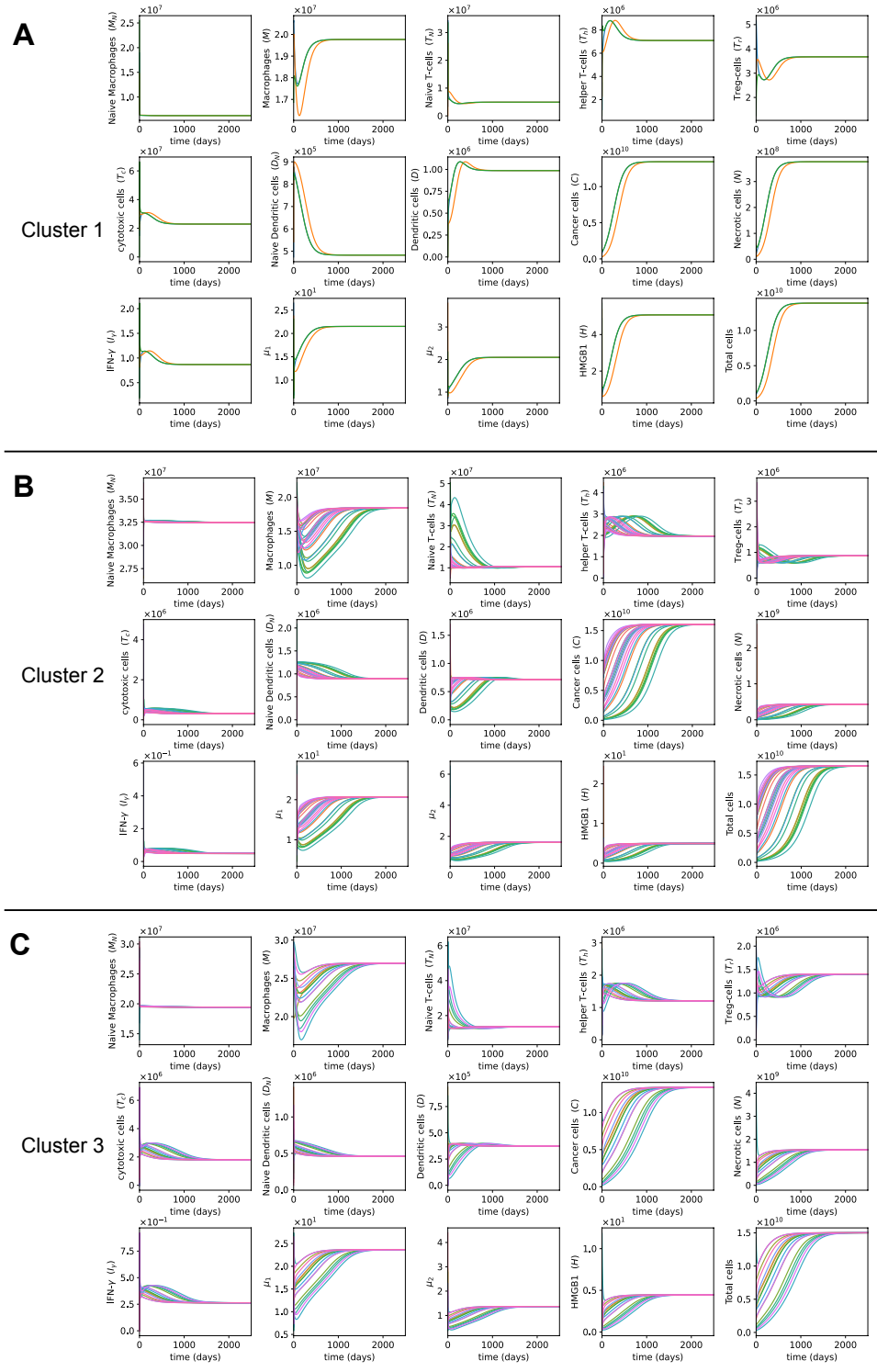

Figure 4.7. **The dynamics with varying initial conditions.** (A-C) The dynamics of cells and cytokines with initial conditions from different patients in clusters 1, 2, and 3, respectively.

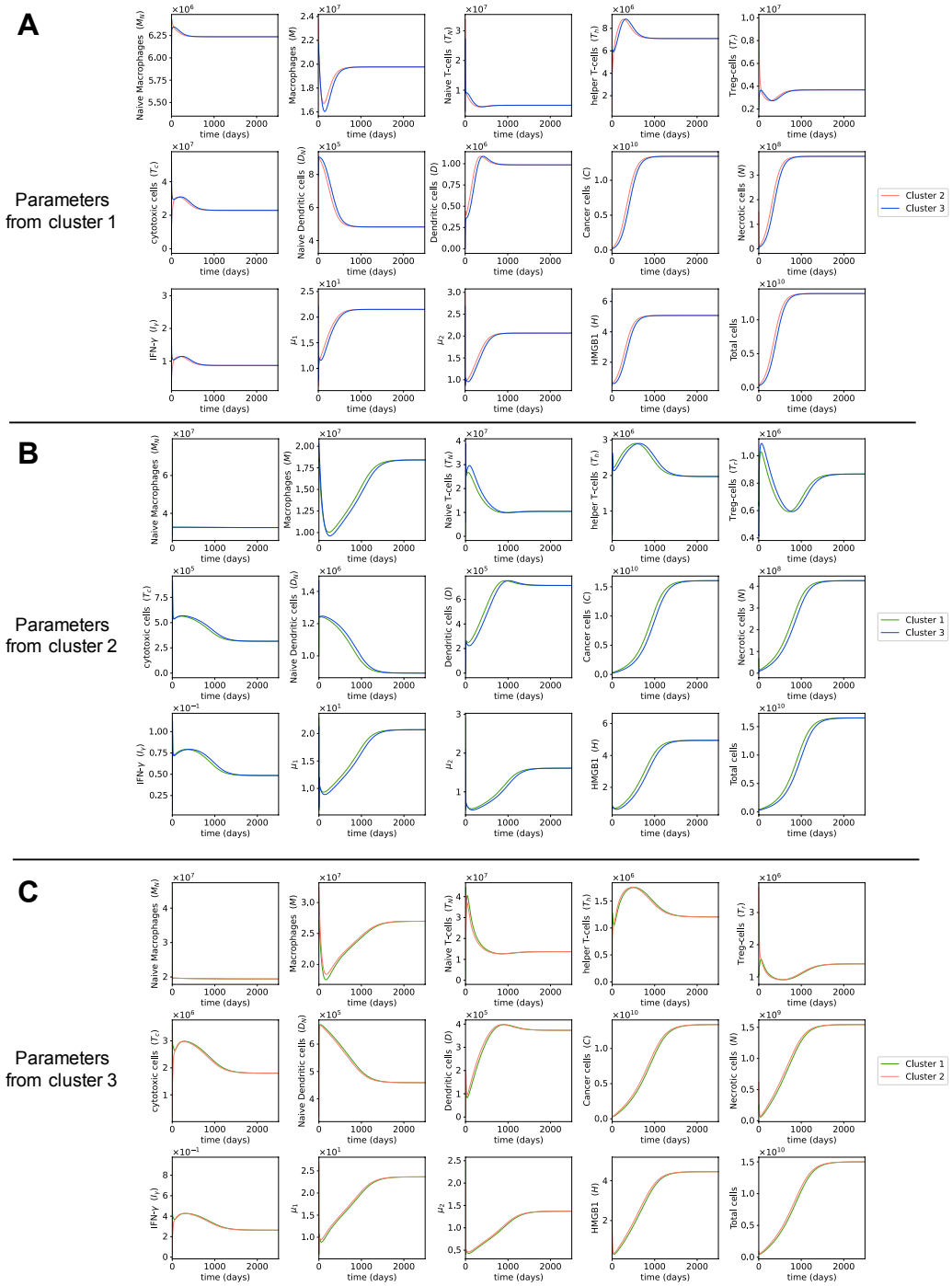

Figure 4.8. **Dynamics with cross-cluster initial conditions.** (A) The dynamics of cells and cytokines with parameters from cluster 1 and initial conditions from clusters 2 and 3. (B) The dynamics of cells and cytokines with parameters from cluster 2 and initial conditions from clusters 1 and 3. (C) The dynamics of cells and cytokines with parameters from cluster 3 and initial conditions from clusters 1 and 2.

conditions.

4.3 Discussion

Our results show that, as cancer cells grow in number, the helper T cell, dendritic cell, cytotoxic cell, and IFN-γ populations increase at first and then decrease with time, while regulatory T cells first decrease in population and then increase. This switch in the dynamics of immune cells happens around the time that cancer cells have the fastest growth. Notably, we also find that, in order to make reasonable predictions regarding the prognosis of cancer patients, it is necessary to study the interactions between immune cells rather than to simply look at the abundance of a certain immune cell type. This observation can be supported by [194], which states that the immune response following from activation of T cells is dependent on the presence of other immune protagonists, such as macrophages, implying that the interactions between immune cells can affect the immune response.

Our results indicate that cluster 3 has the slowest cancer growth and a relatively low population of cancer cells at the steady state. Meanwhile, cluster 2 has one of the fastest cancer growth rates and, more importantly, the highest number of cancer cells at the steady state. Thus, cluster 3 has the most favorable cancer progression, and cluster 2 has the least favorable cancer progression. These results are in agreement with the findings from clinical data in chapter 3 that cluster 3 has the best outcomes and cluster 2 has the worst outcomes.

Our global sensitivity analysis shows that the rate at which cytotoxic cells kill cancer cells has a large impact on the growth of osteosarcoma. Therefore, it is probable that treatments that attempt to increase this rate of cytotoxic cells attacking tumor cells, such as PD-1 or CTLA-4 inhibitors, would work well for osteosarcoma.

In fact, a phase 2 trial reported that some improvement in cancer progression was observed in osteosarcoma patients treated with the anti PD-1 drug, Pembrolizumab [195]. The combined treatment of PD-1 and CTLA-4 blockade therapy has shown even better responses compared with single checkpoint inhibitors in bone sarcoma [196].

In the mathematical modeling of cancers, one of the main challenges is the large number of unknown parameters and a limited availability of data sets to derive parameters from. To combat this challenge, many mathematical models adopt one or a couple of the following approaches: assuming biologically feasible values for some parameters, using estimated parameters from other diseases or rodent studies, calibrating parameters to fit the biological behaviors from an experimental data set, and varying the parameter values within a reasonable range to study the impact of those parameters on the results. In our work, we acquire parameter values from experimental studies in the literature and estimate the others using the steady state assumption with the steady state values derived from patient gene expression data. Importantly, we also perform global sensitivity analysis on the estimated parameters.

All mathematical models thus far use the same parameters for all patients, while our model estimates parameters separately for each cluster of patients with distinctive immune compositions. Since patients with different immune compositions have shown different prognoses and different responses to treatment [197, 198, 199, 200, 201], estimating the parameters for each cluster separately helps us model the effects of immune cells on cancer growth and their responses to treatment more accurately.

To avoid adding complexity to an already complex network, our study does not model the healthy cells in the tumor microenvironment. While several math-

ematical models for tumor growth study the competition between healthy cells and cancerous cells [160, 202, 203, 204], these models typically only investigate a small subset of immune cells, unlike our model, which focuses on many important components of the immune system. Moreover, as the cancer self-proliferation rate (λ_C) in our model is taken from osteosarcoma growth data in humans, which is naturally the growth of tumors with the presence of healthy cells, this parameter already encodes the inhibition of cancer growth due to competition with healthy cells. Therefore, even though we do not explicitly model healthy cells, the impact of healthy cells on cancer growth is incorporated implicitly through λ_C.

While it would be ideal to use time course data to derive the parameters in each cluster, the availability of such data is currently limited, and so instead we use the large tumors in each cluster as the steady state values to estimate these parameters. Despite this limitation due to the lack of time course data, our model still provides valuable insights on the progression of osteosarcoma and the impact of the immune system on its growth, and many studies can build upon this one. Ways to improve this model include utilizing partial differential equations to study the growth of osteosarcoma tumors, both in space and in time, or applying different parameter fitting algorithms [205, 206, 207, 208] to better match the dynamics of the system to real patient data.

CHAPTER 5

INVESTIGATING OPTIMAL CHEMOTHERAPY OPTIONS FOR OSTEOSARCOMA PATIENTS THROUGH A MATHEMATICAL MODEL

In the last chapter, we have developed a data driven mathematical model for the interaction network between key immune cells and cancer cells to investigate tumors' growth behaviors of three distinct groups of osteosarcoma tumors, grouped based on their immune compositions, and group-specific parameters have been calculated to discover differences in tumor growth between groups. In this chapter, we extend our previous model by adding the interactions between the most common chemotherapy drugs for osteosarcoma and important cell types in the tumor microenvironment in order to examine the effects of these drugs on osteosarcoma tumors in each group.

Most chemotherapy treatments for osteosarcoma include one or a combination of the following drugs: high dose Methotrexate (MTX), Doxorubicin (DOX) and Cisplatin (CDDP). The most popular treatment regimen for adolescents is the MAP regimen, consisting of all those three drugs [209, 210], and a widely used treatment for older adults is a two-drug regimen of Doxorubicin and Cisplatin [210]. This study investigates the response to these regimens through a data driven

mathematical model, suggests optimal chemotherapy dosages, as well as compares the behaviors of immune and cancer cells under various conditions such as resistance to chemotherapy and different treatment start times.

5.1 Materials and Methods

5.1.1 Mathematical Model

We build upon the model in the previous chapter by adding the interactions of the variables in that model with the following chemotherapy drugs: Methotrexate, Doxorubicin and Cisplatin. The interaction network of these drugs with cells and cytokines of osteosarcoma tumor microenvironment is shown in Figure 5.1. We use exponential kill model, as introduced in [211], to describe how chemotherapy affects the cancer microenvironment and model the change in population of the new model's variables throughout time in unit of day. The details of the effects of chemotherapy drugs on immune cells and cancer cells are explained below (changes to equations (A.1)-(A.14) are in bold).

Cancer cells

All chemotherapy drugs in our model aim to kill tumor cells, though they have different mechanisms of action. Methotrexate hinders DNA synthesis in fast dividing cancer cells by inhibiting folate dependent pathways [212]. Doxorubicin can kill cancer cells by binding to DNA-associated enzymes, intercalating the base pair of DNA's double helix, and targeting many molecular targets such as topoisomerase enzymes I and II, which results in DNA damage [213]. Cisplatin binds platinum to DNA by forming inter-stranded and intra-stranded crosslinks, thus induces DNA

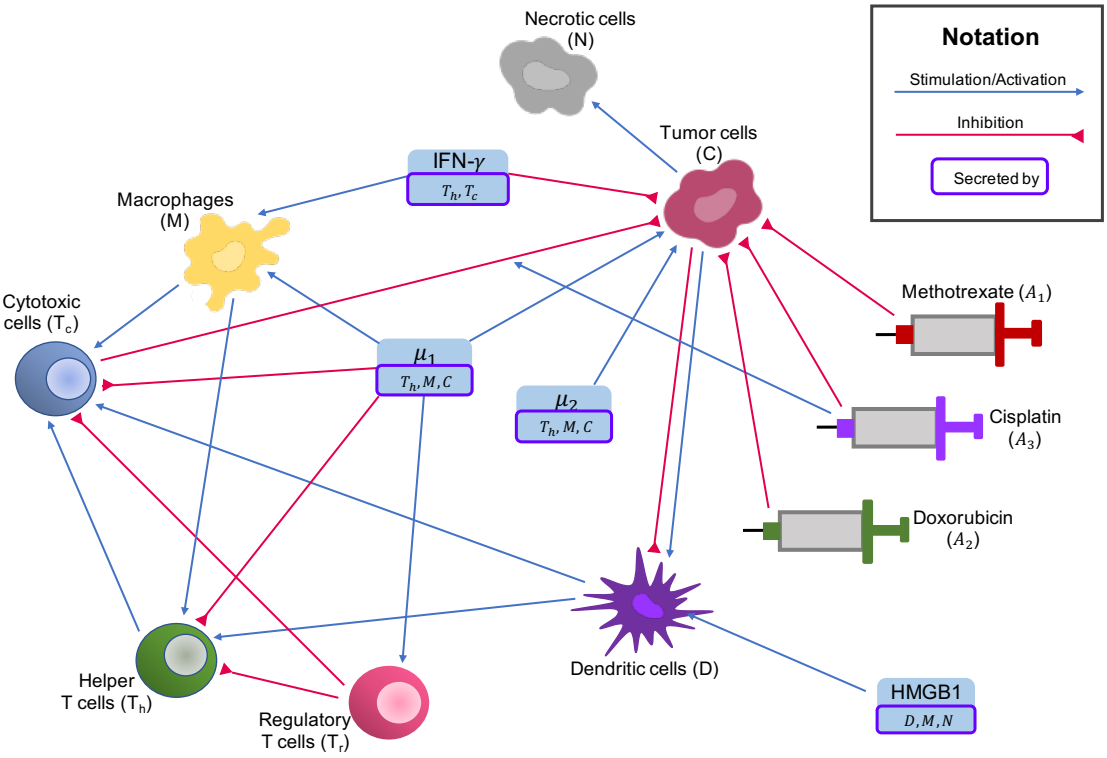

Figure 5.1. **Interaction network with chemotherapy drugs.** Activations, proliferations or stimulations are indicated by blue arrows, and inhibitions are indicated by red arrows. Chemotherapy drugs also inhibit all immune cells (red arrows from drugs to immune cells not shown).

Table 5.1: Model Variables. Names and descriptions of the variables used in the model.

Variable	**Name**	**Description**
T_N	Naive T-cells	
T_h	Helper T-cells	
T_C	Cytotoxic cells	includes CD8+ T-cells and NK cells
T_r	Regulatory T-cells	
D_n	Naive dendritic cells	
D	Activated dendritic cells	antigen presenting cells
M_N	Naive macrophages	includes naive macrophages and monocytes
M	Macrophages	includes M1 macrophages and M2 macrophages
C	Cancer cells	
N	Nectrotic cells	
H	HMGB1	
μ_1	Cytokines group μ_1	includes effects of TGF-β, IL-4, IL-10 and IL-13
μ_2	Cytokines group μ_2	includes effects of IL-6 and IL-17
I_γ	IFN-γ	
A_1	Methotrexate	Methotrexate concentration at tumor site
A_2	Doxorubicin	Doxorubicin concentration at tumor site
A_3	Cisplatin	Cisplatin concentration at tumor site

damage which leads to cell death in rapidly proliferating cells [214, 215].

Similar to [211, 216], we use saturation kill term $(1 - e^{\beta A})$ to model the direct cytotoxic effects of chemotherapy drugs on cancer cells, where β is the drug efficacy parameter, and A is the drug concentration at the tumor site. This is based on the observation that at low concentration, the cancer killing effect of these drugs are

almost linear, but at very high concentration the cancer killing effect plateaus. Unlike Doxorubicin and Cisplatin, Methotrexate can only eliminate cancer cells during certain phases of the cell cycle, so we add the term $(f - \frac{\tau}{a} + \frac{1}{24a})$ to Methotrexate's cytotoxic effect to account for this phenomenon, as modeled in [211]. Here, f denotes the fraction of cells in the vulnerable phase of the cell cycle for Methotrexate, a denotes cell cycle time in days, and τ is defined to be $minimum(T, fa)$ with T being drug exposure time in days.

Besides its direct role in targeting tumor cells, Cisplatin has also been reported to increase cytotoxic cells' cancer killing capability up regulating MHC-1 expression of cancer cells [81, 214, 217, 218]. We also use saturation term to describe this effect, as it is very likely that high concentration of Cisplatin can also plauteau at up-regulating MHC-1 in cancer cells. We make the assumption that the concentration of Cisplatin at which this effect slows down is about the same concentration at which the cancer killing effect of Cisplatin slows down, so we use the same drug efficacy parameter β_3 in both terms, resulting in the following equation for cancer cells:

$$\begin{aligned}
\frac{d\,[C]}{dt} = \; & \left(\lambda_C + \lambda_{C\mu_1}\,[\mu_1] + \lambda_{C\mu_2}\,[\mu_2]\right)[C]\left(1 - \frac{[C]}{C_0}\right) \\
& - \left(\delta_{CI_\gamma}\,[I_\gamma] + \delta_C + \delta_{CT_c}\left(1 + \boldsymbol{\delta_{CT_cA_3}(1 - e^{-\beta_3 A_3})}\right)[T_c]\right)[C] \\
& - \left(\boldsymbol{K_C}\left(\boldsymbol{f - \frac{\tau}{a} + \frac{1}{24a}}\right)\boldsymbol{(1 - e^{-\beta_1 A_1})} + \boldsymbol{K_C(1 - e^{-\beta_2 A_2})}\right. \\
& \left. + \boldsymbol{K_C(1 - e^{-\beta_3 A_3})}\right)\boldsymbol{[C]}
\end{aligned} \tag{5.1}$$

where $\delta_{CT_cA_3}$ represents the effect of Cisplatin to promote cytotoxic cell's cancer killing ability; K_C is rate of chemotherapy-induced tumor death; and β_1, β_2, β_3 are

medicine efficacy coefficients of Methotrexate, Doxorubicin, Cisplatin, respectively. A description of every chemotherapy-related parameter in our model is given in Table A.1.

Necrotic cells

As a proportion of cancer cells killed by chemotherapy drugs become necrotic cells, we describe the change in population of necrotic cells with the presence of chemotherapy as follows:

$$\begin{aligned}\frac{d\,[N]}{dt} = \; & \alpha_{NC}\left(\delta_{CI_\gamma}\,[I_\gamma] + \delta_C + \delta_{CT_c}\left(1 + \boldsymbol{\delta_{CT_cA_3}(1 - e^{-\beta_3 A_3})}\right)[T_c]\right)[C] \\ & + \boldsymbol{\alpha_{NCA}}\left(\boldsymbol{K_C}\left(\boldsymbol{f - \frac{\tau}{a} + \frac{1}{24a}}\right)\boldsymbol{(1 - e^{-\beta_1 A_1})} + \boldsymbol{K_C(1 - e^{-\beta_2 A_2})}\right. \\ & \left. + \boldsymbol{K_C(1 - e^{-\beta_3 A_3})}\right)[\boldsymbol{C}] - \delta_N\,[N] \end{aligned} \tag{5.2}$$

where α_{NCA} is the fraction of dying cancer cells induced by chemotherapy that turn into necrotic cells.

Immune cells

Since chemotherapy does not only eliminate tumor cells but also kills immune cells, we include the effects of chemotherapy in the equations of immune cells as well. Similar to [216], we assume that the same amount of chemotherapy drugs is required to affect cancer cells and immune cells, even when the rate at which chemotherapy kills cancer cells is different than when it kills immune cells. Hence, we use the same drug efficacy coefficients for cancer and immune cells, but different rates of drug-induced cell death between them, leading to the following modified immune cells' equations:

$$
\begin{aligned}
\frac{d\left[M_N\right]}{dt} = & A_{M_N} - \left(\lambda_{MI_\gamma}\left[I_\gamma\right] + \lambda_{M\mu_1}\left[\mu_1\right]\right)\left[M_N\right] - \delta_{M_N}\left[M_N\right] \\
& - \left(\boldsymbol{K_{M_N}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_{M_N}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right. \\
& \left. + \boldsymbol{K_{M_N}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{M_N}\right] \qquad (5.3) \\
\frac{d\left[M\right]}{dt} = & \left(\lambda_{MI_\gamma}\left[I_\gamma\right] + \lambda_{M\mu_1}\left[\mu_1\right]\right)\left[M_N\right] - \delta_M\left[M\right] - \left(\boldsymbol{K_M}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)\right. \\
& \left.\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_M}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right) + \boldsymbol{K_M}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{M}\right] \qquad (5.4) \\
\frac{d\left[T_N\right]}{dt} = & A_{T_N} - \left(\lambda_{T_hM}\left[M\right] + \lambda_{T_hD}\left[D\right]\right)\left[T_N\right] - \lambda_{T_r\mu_1}\left[\mu_1\right]\left[T_N\right] \\
& - \left(\lambda_{T_cT_h}\left[T_h\right] + \lambda_{T_cM}\left[M\right] + \lambda_{T_cD}\left[D\right]\right)\left[T_N\right] - \delta_{T_N}\left[T_N\right] \\
& - \left(\boldsymbol{K_{T_N}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_{T_N}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right. \\
& \left. + \boldsymbol{K_{T_N}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{T_N}\right] \qquad (5.5) \\
\frac{d\left[T_h\right]}{dt} = & \left(\lambda_{T_hM}\left[M\right] + \lambda_{T_hD}\left[D\right]\right)\left[T_N\right] - \left(\delta_{T_hT_r}\left[T_r\right] + \delta_{T_h\mu_1}\left[\mu_1\right] + \delta_{T_h}\right)\left[T_h\right] \\
& - \left(\boldsymbol{K_{T_h}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_{T_h}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right. \\
& \left. + \boldsymbol{K_{T_h}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{T_h}\right] \qquad (5.6) \\
\frac{d\left[T_r\right]}{dt} = & \left(\lambda_{T_r\mu_1}\left[\mu_1\right]\right)\left[T_N\right] - \delta_{T_r}\left[T_r\right] - \left(\boldsymbol{K_{T_r}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right)\right. \\
& \left. + \boldsymbol{K_{T_r}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right) + \boldsymbol{K_{T_r}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{T_r}\right] \qquad (5.7) \\
\frac{d\left[T_c\right]}{dt} = & \left(\lambda_{T_cT_h}\left[T_h\right] + \lambda_{T_cM}\left[M\right] + \lambda_{T_cD}\left[D\right]\right)\left[T_N\right] - \left(\delta_{T_cT_r}\left[T_r\right] + \delta_{T_c\mu_1}\left[\mu_1\right] + \delta_{T_c}\right)\left[T_c\right] \\
& - \left(\boldsymbol{K_{T_c}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_{T_c}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right.
\end{aligned}
$$

$$+ \boldsymbol{K_{T_c}}(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}})\Bigg)[\boldsymbol{T_c}] \tag{5.8}$$

$$\begin{aligned}\frac{d[D_N]}{dt} = \; & A_{D_N} - (\lambda_{DC}[C] + \lambda_{DH}[H])[D_N] - \delta_{D_N}[D_N] \\ & - \Bigg(\boldsymbol{K_{D_N}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}) + \boldsymbol{K_{D_N}}(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}) \\ & + \boldsymbol{K_{D_N}}(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}})\Bigg)[\boldsymbol{D_N}]\end{aligned} \tag{5.9}$$

$$\begin{aligned}\frac{d[D]}{dt} = \; & (\lambda_{DC}[C] + \lambda_{DH}[H])[D_N] - (\delta_{DC}[C] + \delta_D)[D] \\ & - \Bigg(\boldsymbol{K_D}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}) + \boldsymbol{K_D}(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}) \\ & + \boldsymbol{K_D}(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}})\Bigg)[\boldsymbol{D}]\end{aligned} \tag{5.10}$$

where K_{M_N}, K_M, K_{T_N}, K_{T_h}, K_{T_r}, K_{T_c}, K_{D_N} and K_D are the rate of chemotherapy-induced cell death of naive macrophages, macrophages, naive T cells, helper T cells, regulatory T cells, cytotoxic cells, naive dendritic cells and dendritic cells, respectively.

Chemotherapy drugs

Chemotherapy drugs are given through IV infusion in osteosarcoma treatments, so their bioavailability is 100%. Thus, we use the following equations to model the change in concentration of each chemotherapy drug at the tumor site over time:

$$\frac{d[A_1]}{dt} = v_{A_1}(t) - \delta_{A_1}[A_1] \tag{5.11}$$

$$\frac{d[A_2]}{dt} = v_{A_2}(t) - \delta_{A_2}[A_2] \tag{5.12}$$

$$\frac{d[A_3]}{dt} = v_{A_3}(t) - \delta_{A_3}[A_3] \tag{5.13}$$

Here, $v_{A_1}(t)$, $v_{A_2}(t)$ and $v_{A_3}(t)$ are the amount of Methotrexate, Doxorubicin and Cisplatin injected per day per liter of body volume, with the unit of mg/l per day;

while δ_{A_1}, δ_{A_2} and δ_{A_3} are respectively the decay rate of Methotrexate, Doxorubicin and Cisplatin.

5.1.2 Data of the Model

The populations of cancer cells, necrotic cells, immune cells and cytokines are obtained from the previous chapter. Then, given a treatment regimen of interest, we apply its standard dosage to our model. Most doses of chemotherapy drugs for osteosarcoma are measured in mg/m^2, but we model the drug concentration at tumor site in mg per liter of body volume. We therefore need to convert drug doses from mg/m^2 to mg/l. We use an average body surface area of a human male of 1.9 m^2 [219] and an average male body volume of 59.7 liters [220] for conversion. That is, for example, 75 mg/m^2 would be equivalent to:

$$75\ mg/m^2 = \frac{75\ mg}{m^2} \times \frac{1.9\ \text{m}^2}{59.7\ l} = 2.3869\ mg/l \tag{5.14}$$

5.1.3 Parameter values

The drug efficacy coefficients, as well as cell cycle time and fraction of cells in the vulnerable phase of the cell cycle are taken from [211]. Using the molecular mass of chemotherapy drugs [216, 221, 223, 222], we convert the drug efficacy coefficients given in [211] to units of (mg/l)$^{-1}$:

$$\beta_1 = \left(\frac{1.126\ l}{\mu mol}\right)\left(\frac{10^6\ \mu mol}{1\ mol}\right)\left(\frac{1\ mol}{454.4\ g\ \text{MTX}}\right)\left(\frac{1\ g}{1000\ mg}\right) = 2.4780\ l/mg$$

$$\beta_2 = \left(\frac{1.063\ l}{\mu mol}\right)\left(\frac{10^6\ \mu mol}{1\ mol}\right)\left(\frac{1\ mol}{580\ g\ \text{DOX}}\right)\left(\frac{1\ g}{1000\ mg}\right) = 1.8328\ l/mg$$

$$\beta_3 = \left(\frac{0.044\ l}{\mu mol}\right)\left(\frac{10^6\ \mu mol}{1\ mol}\right)\left(\frac{1\ mol}{300\ g\ \text{CDDP}}\right)\left(\frac{1\ g}{1000\ mg}\right) = 0.1467\ l/mg$$

Drug efficacy coefficients for Doxorubicin-resistant and Cisplatin-resistant cells are also included in [211], and can be converted in a similar way. The values for

all chemotherapy-related parameters in our model and their sources are given in Table A.1.

The fractional cancer cell killed by chemotherapy, K_C, is taken from [216, 224] to be 0.9, based on the notion that chemotherapy strength is one log kill [225]. Since chemotherapy is more effective at eliminating fast proliferating cells, it is safe to assume that the rates of chemotherapy-induced death of immune cells are smaller than that of cancer cells. Hence, K_{M_N}, K_M, K_{T_N}, K_{T_h}, K_{T_r}, K_{T_c}, K_{D_N} and K_D are assumed to be smaller than but on the same order of magnitude as K_C, and we use a value of 0.6 for them, similar to [224].

Decay rates of chemotherapy drugs are derived from their elimination half lives in the following way:

$$\delta_A = \frac{ln2}{\text{half life of } A \text{ in days}} \tag{5.15}$$

where δ_A is the decay rate of A. On average, the elimination half lives of Doxorubicin, Cisplatin and high dose Methotrexate are 2 hours, 25 minutes and 11.5 hours respectively [226, 227, 228], resulting in the corresponding decay rates of 8.3178, 39.9253 and 1.4466.

As no values for α_{NCA} and $\delta_{CT_cA_3}$ can be found in literature, we assume biologically reasonable values for these parameters. α_{NCA} is the fraction of dying cancer cells induced by chemotherapy that become necrotic cells, so it is bounded between 0 and 1. We make the assumption that a large proportion of dying cancer cells from treatment turn into necrotic cells, and set $\alpha_{NCA} = 0.8$. For $\delta_{CT_cA_3}$, we assume that Cisplatin at maximum effect can double the cancer killing ability of cytotoxic cells, or equivalently $\delta_{CT_cA_3} = 1$. In order to investigate whether our assumptions on these parameters have a large impact on the cancer population, in the later section we perform sensitivity analysis, as well as study the change in cancer population

after treatment while varying these two parameters.

Table 5.2: Chemotherapy Parameters. Name, unit, description, value and source of chemotherapy-related parameters used in the model.

Parameter	**Unit**	**Description**	**Value**	**Source**
f	none	Initial fraction of cells in vulnerable phase of the cell cycle	0.5	[211]
a	day	Cell cycle time	0.6667	[211]
T	day	Duration of drug exposure		
τ	day		$\min(T, fa)$	[211]
β_1	mg/l^{-1}	Methotrexate efficacy coefficient	2.4780	[211]
β_2	mg/l^{-1}	Doxorubicin efficacy coefficient	1.8328	[211]
β_3	mg/l^{-1}	Cisplatin efficacy coefficient	0.1467	[211]
K_C	day^{-1}	Rate of chemo-induced tumor death	0.9	[216, 224]
K_{M_N}	day^{-1}	Rate of chemo-induced death of naive macrophages	0.6	[224]
K_M	day^{-1}	Rate of chemo-induced death of macrophages	0.6	[224]
K_{T_N}	day^{-1}	Rate of chemo-induced death of naive T-cells	0.6	[224]
K_{T_h}	day^{-1}	Rate of chemo-induced death of helper T-cells	0.6	[224]
K_{T_r}	day^{-1}	Rate of chemo-induced death of regulatory T-cells	0.6	[224]
K_{T_c}	day^{-1}	Rate of chemo-induced death of cytotoxic cells	0.6	[224]

K_{D_N}	day^{-1}	Rate of chemo-induced death of naive dendritic cells	0.6	[224]
K_D	day^{-1}	Rate of chemo-induced death of dendritic cells	0.6	[224]
$\delta_{CT_cA_3}$	none	Effect of Cisplatin to promote cancer killing ability of cytotoxic cells	1	Assumed
α_{NCA}	none	Fraction of chemo-induced dying tumor cells that become necrotic cells	0.8	Assumed
δ_{A_1}	day^{-1}	Decay rate of Methotrexate	1.4466	[228]
δ_{A_2}	day^{-1}	Decay rate of Doxorubicin	8.3178	[226]
δ_{A_3}	day^{-1}	Decay rate of Cisplatin	39.9253	[227]

5.1.4 Non-Dimensionalization

To achieve additional numerical stability, non-dimensionalization of the whole system is carried out. For each variable X of the original system in chapter 4, its dimensionless form can be written similar to chapter 4:

$$\overline{X} = \frac{X}{X^{\infty}}, \tag{5.16}$$

where X^{∞} is the steady state value of X given in Table 4.2. For the newly added variables, which are the chemotherapy drugs, we introduce the following non-dimensional variables:

$$\overline{A} = \frac{A\ \delta_A}{v_A^*}, \tag{5.17}$$

where A is the dimensional variable, δ_A is the decay rate of A, and v_A^* is the amount of drug A injected on its first injection day of the treatment. Further details on non-dimensionalization are given in Appendix B.2.

To solve the non-dimensional system of ordinary differential equations, we use *solve_ivp* function from Scipy package in python [188], with initial conditions from a chosen data point of interest in each cluster.

5.1.5 Sensitivity Analysis

We perform local gradient-based sensitivity analysis on all chemotherapy-related parameters to study their impacts on the outputs of the system. For the non-dimensional system $\frac{d\overline{X}}{dt} = F(\overline{X}, \theta, t)$ with model parameters $\theta = \theta_1, ..., \theta_N$, the local (first order) sensitivity of parameter θ_i with respect to the variable X is defined as [189]:

$$s_i = \frac{\partial \overline{X}}{\partial \theta_i} \tag{5.18}$$

As we are mainly interested in how drug-related parameters affect the number of cancer cells, we calculate sensitivity of treatment parameters with respect to cancer and total cell population. Since the effects of the treatment does not reach the steady state, we consider time-dependent sensitivity. That is, we measure sensitivity of parameters in every time step throughout the treatment and some time after. The change in sensitivity of θ_i over time can be derived as follows:

$$\frac{\partial s_i}{\partial t} = \frac{\partial}{\partial t}\left(\frac{\partial \overline{X}}{\partial \theta_i}\right) = \frac{\partial}{\partial \theta_i}\left(\frac{\partial \overline{X}}{\partial t}\right) = \frac{\partial F(\overline{X}, \theta_i, t)}{\partial \theta_i} \tag{5.19}$$

By applying the chain rule, we have:

$$\frac{\partial s_i}{\partial t} = \frac{\partial F}{\partial \theta_i} + \frac{\partial F}{\partial \overline{X}} s_i \tag{5.20}$$

In addition, we also look at the relative sensitivity, which is commonly used in metabollic control analysis of biological networks [189]:

$$\overline{s}_i(t) = s_i(t)\frac{\theta_i}{\overline{X}(t)} \tag{5.21}$$

Then, we compute the average sensitivity of each type over a period of time T:

$$S_i = \frac{1}{T}\int_0^T s_i(t)dt, \qquad \overline{S}_i = \frac{1}{T}\int_0^T \overline{s}_i(t)dt \tag{5.22}$$

The sensitivity varies for different values of the parameters, so we consider a small neighborhood $\Omega(\theta)$ of the given parameter set and calculate:

$$\mathbf{S}_i = \int_\Omega S_i(\theta)d\theta, \qquad \overline{\mathbf{S}}_i = \int_\Omega \overline{S}_i(\theta)d\theta \tag{5.23}$$

where the integrals are computed using a numerical technique called sparse grid points [190, 191].

5.1.6 Optimization of drug dosage

We introduce a framework to find the appropriate dose of a given treatment regimen for each patient. To find the optimal dosage to achieve a target cancer population after treatment, we minimize the following loss function:

$$L(v,t) = \left(\hat{C}(v,t) - C_{target}\right)^2 + \kappa\sum_{i=1}^{M}|v_i| \tag{5.24}$$

$$\text{subject to } \ 0 \le v_i \le U_i, i = 1, ..., M$$

Here, v is a vector of length M, denoting the doses of the M drugs in the given treatment, t is the time of interest at which cancer population is evaluated for optimization, typically right after treatment or at the time of surgery. $\hat{C}(v,t)$ is the cancer population with drug doses v at time t of interest, and is computed via our ODE solver. C_{target} is the target cancer population at time t, and is chosen by the user. As very high doses of chemotherapy are known to induce high toxicity to the patient, we put an upper bound constraint U_i on the dosage of each drug.

We utilize the least square error to describe the difference between $\hat{C}(v,t)$ and the target cancer population C_{target}. The regularization term $\kappa\sum_{i=1}^{M}|v_i|$ is added

so that we can reach the approximate target cancer population with the smallest possible drug doses. The higher the regularizaztion parameter κ is, the more the optimizer focuses on achieving small doses and less on achieving small error between $\hat{C}(v,t)$ and C_{target}.

The *optimize.minimize* function from Scipy package in python is used to solve this optimization problem, with the outputs being the optimal doses.

5.2 Results

5.2.1 Dynamics of cancer microenvironment with MAP treatment

Typical treatments for osteosarcoma include neoadjuvant chemotherapy, usually for 10 weeks, then surgery, and adjuvant chemotherapy after the surgery for up to a year [210]. The most common chemotherapy regimen for osteosarcoma in children and young adults is the MAP regimen, which is a combination of Doxorubicin, Cisplatin and high dose Methotrexate [209]. This regimen consists of six 35-day cycles, where two cycles are applied before surgery and the remaining four are applied after surgery. In each cycle, 37.5 mg/m^2 of Doxorubicin and 60 mg/m^2 of Cisplatin are administered through IV per day on day 1 and 2, and 12000 mg/m^2 of Metrotrexate is administered through IV over 4 hours per day on day 22 and 29 [229, 230]. Different infusion schedules have been used for Doxorubicin and Cisplatin: Doxorubicin can be injected as a bolus or a 4-hour infusion each day, or a continuous infusion over 48 hours, while Cisplatin can be injected over 2 or 4 hours each day, or continuously over 72 hours [230]. Here, we study the dynamics of cells and cytokines populations in large osteosarcoma tumors during neoadjuvant MAP treatment, which includes two 35-day cycles. In particular, we use the steady state

values of cells and cytokines in chapter 4 as initial conditions, the typical dosage of the MAP regimen as the drug inputs, and 4-hour infusions on previously specified days as the injection schedule for each drug. We set the start of chemotherapy treatment to be 7 days after biopsy, as it usually takes a few days to receive the results of the biopsy.

Figure 5.2 shows that for all clusters, cancer populations are reduced significantly after two cycles of MAP treatment. It is important to note that the cancer populations do not reach zero after chemotherapy so cancer cells will start growing again after chemotherapy. However, the goal of neoadjuvant therapy is not to eradicate cancer cells completely, but only to reduce the boundaries of the tumor and to remove any small metastases that have not been detected [231].

Cluster 2 has the highest cancer population at the start of treatment, so naturally cluster 2 also has the highest cancer population left after neoadjuvant therapy. Interestingly, cluster 1 has approximately the same number of cancer cells as cluster 3 at the start of treatment, but ends up with higher cancer population than cluster 3 after treatment. This is because in each chemotherapy cycle, there are a few weeks where no chemotherapy drugs are administered in order to allow the patient to recover from the drugs' toxicity, and thus during these few weeks, cancer cells can start growing again. Cluster 1's cancer population, which is reported in chapter 4 to have the highest growth rate in the three clusters, grow more during the weeks with no drugs given, resulting in higher number of cancer cells after treatment compared to cluster 3. This observation suggests that we should take the patient's cancer growth rate into account when choosing their chemotherapy dosage.

During the MAP treatment, necrotic cells, dendritic cells and HMGB1 oscillate between increasing and decreasing in abundance. Since chemotherapy drugs aim

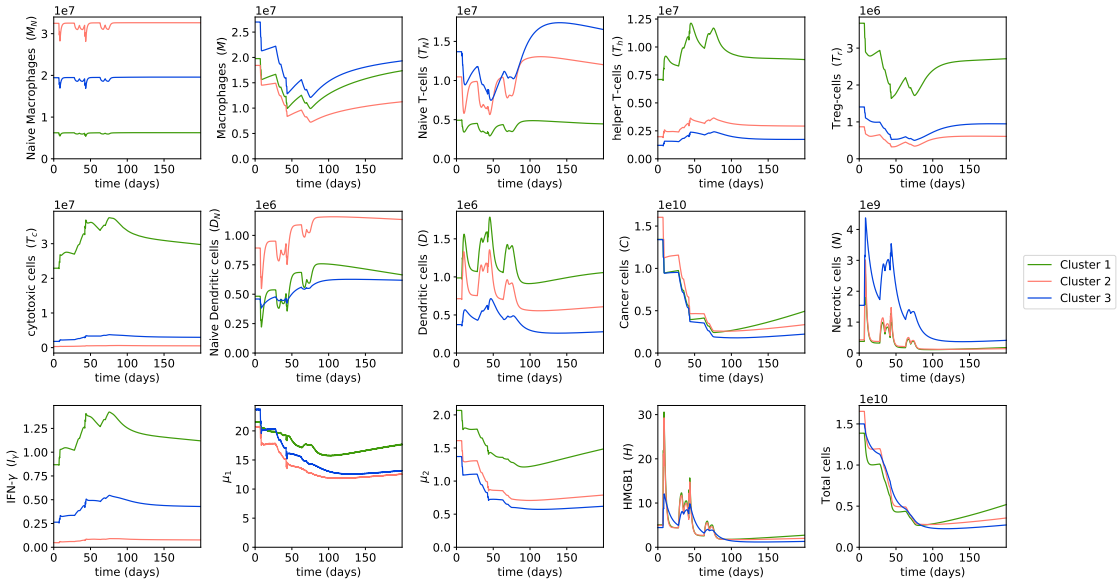

Figure 5.2. **Dynamics with MAP treatment.** Behaviors of cells and cytokines in osteosarcoma tumors during the MAP treatment and a few months after treatment. Initial conditions are large tumors in each cluster, i.e. the without-treatment steady state values of each cluster. Drug doses are the typical doses of the MAP regimen. The different color lines indicate the dynamics of different clusters.

to kill cancer cells, and a fraction of dying cancer cells become necrotic cells, the population of necrotic cells increases on the days the drugs are injected and there are many drug-induced dying cancer cells. However, during the weeks where no chemotherapy drugs are administered, only drug-free dying cancer cells can become necrotic cells, and with cancer population being already reduced by the previously given drugs, the number of drug-free dying cancer cells is small, leading to a decrease in necrotic cell population.

HMGB1 is mainly produced by necrotic cells, so HMGB1 abundance increases when necrotic population increases and decreases when necrotic population decreases. Meanwhile, dendritic cells are activated largely by HMGB1, so the dynamics of dendritic cells share the same trend with the dynamics of HMGB1. That means both HMGB1 and dendritic cells increase on the days chemotherapy drugs are administered and decrease on the weeks with no drugs given. An increase in dendritic cell population right after Doxorubicin [232, 233, 234, 235] and Cisplatin [81, 82, 214] introduction, as well as a rise in HMGB1 production following Doxorubicin [234, 235, 236], has been shown in several studies, which aligns with our results.

We observe that in general, helper T cells, cytotoxic cells and IFN-γ increase in population during chemotherapy. There are many studies that report an increase in helper T cells and/or cytotoxic T cells' abundance due to Doxorubicin [235, 236, 237, 238, 239, 240], Cisplatin [81, 214, 217, 241] and Methotrexate [242], and thus support our findings. Especially, Doxorubicin has been known to induce immunogenic cell death, which leads to the maturation of dendritic cells and accordingly the activation of helper and cytotoxic T cells [232, 238, 243]. IFN-γ is produced by helper T cells and cytotoxic cells, thus IFN-γ abundance also increases as these two cells increase in population during MAP treatment. The increase in

IFN-γ level after administration of Doxorubicin and Cisplatin has also been observed in multiple studies [214, 234, 235, 244].

On the other hand, population of macrophages, regulatory T cells, cytokines groups μ_1 and μ_2 mainly decrease during MAP treatment. These immune cells are not affected by the necrotic cell death process caused by chemotherapy, so they decrease in population during chemotherapy as they are also killed by the drugs. μ_1 and μ_2 are produced by helper T cells, macrophages and cancer cells. Even though helper T cell population increases during treatment, macrophages and cancer cell populations decrease at a greater magnitude, which leads to an overall decrease in μ_1 and μ_2 throughout MAP treatment. Several other studies have also reported a reduction in regulatory T cell number due to Cisplatin [81, 214, 217] and a decrease in the level of IL-6, which is the main component of μ_2, due to Methotrexate and Doxorubicin [242, 244, 245, 246].

5.2.2 Sensitivity analysis

To study the impact of the newly introduced parameters on the outputs of the model, we perform local sensitivity analysis on the chemotherapy-related parameters using the non-dimensional system. The initial conditions for sensitivity analysis are the large tumors in each cluster, which we use without-treatment steady state values to represent. It is worth noting that the cell cycle time, a, is not included in this sensitivity analysis, because it is a simple measurement rather than a parameter that needs to be estimated or fitted to the experimental data. The most sensitive time-averaged parameters in terms of sensitivity and relative sensitivity are presented in Figure 5.3.

In all clusters, the initial fraction of cells in the vulnerable phase of the cell cycle f has the largest impact on cancer population among treatment-related pa-

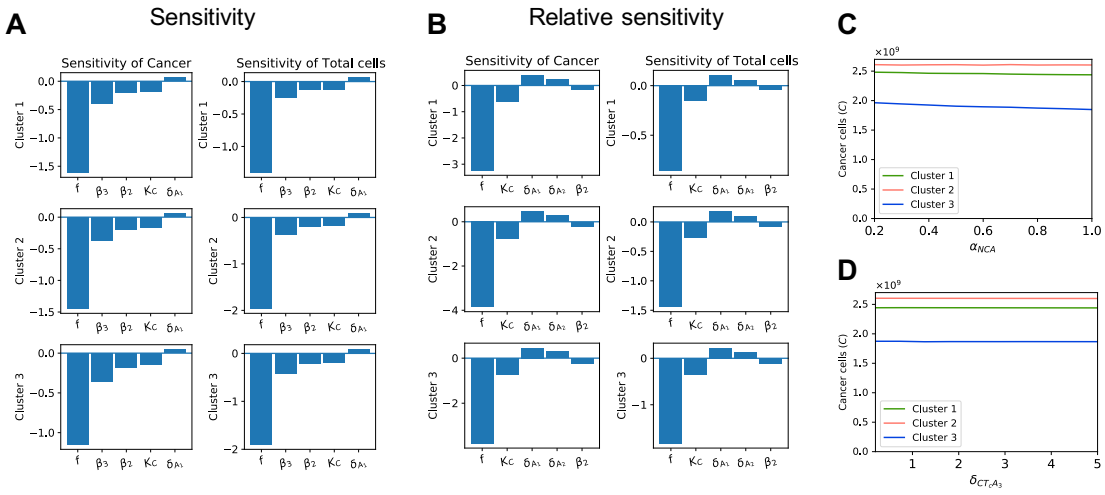

Figure 5.3. **Sensitivity of chemotherapy-related parameters.** Sub-figure A shows the local sensitivity of 5 most sensitive treatment-related parameters with respect to cancer population and total cell population. Sub-figure B shows the local relative sensitivity of 5 most sensitive treatment-related parameters with respect to cancer population and total cell population. Sub-figures C and D display the cancer population after treatment with different values of α_{NCA} and $\delta_{CT_cA_3}$, respectively.

rameters according to both the sensitivity and relative sensitivity analyses. The rate of chemotherapy-induced cancer cell death, K_C, and the drug efficacy coefficients of Doxorubicin and Cisplatin, β_2 and β_3, are also sensitive to cancer and total cell population during treatment. Meanwhile, the drug efficacy coefficient of Methotrexate, β_1, does not seem as sensitive, but the decay rate of Methotrexate is.

We notice that the parameters whose values are assumed, α_{NCA} and $\delta_{CT_cA_3}$, do not have a large effect on the cancer population or total cell population based on the results of the sensitivity analysis. To further confirm this, we also plot the cancer population after treatment with different values of these parameters. We choose α_{NCA} ranging from 0.2 to 1, because it is a fraction and thus is bounded between 0 and 1, and $\delta_{CT_cA_3}$ ranging from 0.2 to 5 times its original value. Figure 5.3C and D show that varying either of these parameters results in negligible changes to cancer population after treatment.

5.2.3 Dynamics of cancer microenvironment in chemo-resistant tumors with MAP treatment

The effectiveness of chemotherapy is highly dependent on the existence of resistant cancer cells. We are interested in studying the change in population of cells and cytokines when osteosarcoma cells are resistant to one or multiple drugs within the MAP regimen. As mentioned in section 2, we obtain drug efficacy coefficients from [211], where these values were estimated to fit the survival data of cancer cells under different chemotherapy drugs. The same study [211] also includes the estimated drug efficacy coefficients of Doxorubicin and Cisplatin in Doxorubicin-resistant and Cisplatin-resistant cancer cells, respectively. Using these parameter values, we plot the dynamics in osteosarcoma microenvironment during MAP treatment when can-

cer cells are resistant to either Doxorubicin or Cisplatin, or to both drugs. Since Methotrexate-resistant cells were not used in [211], and hence no parameter values are available for them, we do not model the dynamics with Methotrexate-resistant cells.

Figure 5.4A shows that MAP treatment is not as effective in shrinking the tumor, when cancer cells are resistant to Doxorubicin, with the cancer population after treatment about 60% to 70% higher in Doxorubicin-resistant cells than in non-Doxorubicin-resistant cells (Table 5.3). The smaller reduction in cancer population of Doxorubicin-resistant cells during Doxorubicin administration means fewer necrotic cells are produced in the process, and accordingly lower level of dendritic cells (Figure 5.4A), as necrotic cells indirectly promote dendritic cell maturation through the release of HMGB1. We notice no clear difference in the dynamics of T cells and macrophages compared to the microenvironment of non-Doxorubicin-resistant cells. It is worth noting that we model only cancer cells to be resistant to chemotherapy drugs, so immune cells are by design not resistant to these drugs.

Table 5.3: Cancer population after MAP treatment with chemotherapy-resistant cells.

Cluster	**Initial cancer population**	**Cancer population after treatment**			
		Chemotherapy sensitive	Resistant to DOX	Resistant to CDDP	Resistant to DOX + CDDP
1	1.34×10^{10}	2.44×10^{9}	3.82×10^{9}	2.49×10^{9}	3.89×10^{9}
2	1.6×10^{10}	2.6×10^{9}	4.32×10^{9}	2.66×10^{9}	4.41×10^{9}
3	1.34×10^{10}	1.87×10^{9}	3.23×10^{9}	1.92×10^{9}	3.29×10^{9}

On the other hand, with Cisplatin-resistant cells, we observe little difference in

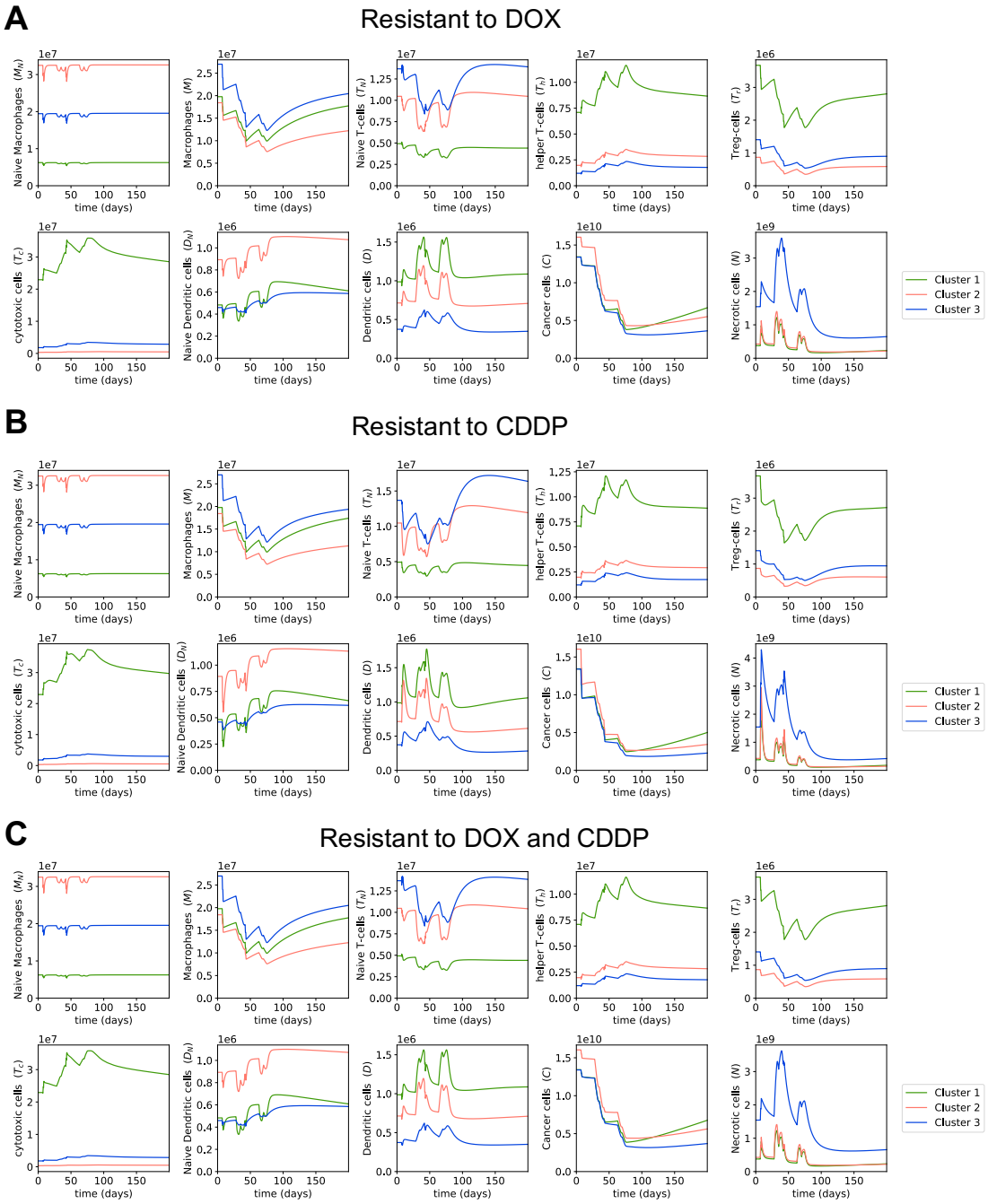

Figure 5.4. **Dynamics in chemotherapy-resistant cells with MAP treatment.** Sub-figures A-C show the dynamics of immune, cancer and necrotic cells in osteosarcoma during the MAP treatment and a few months after treatment when cancer cells are resistant to Doxorubicin, Cisplatin, and both Doxorubicin and Cisplatin, respectively.

the reduction of cancer population compared to non-Cisplatin-resistant cells (Figure 5.4B, Table 5.3). This is due to the fact that Cisplatin's drug efficacy parameter, β_3, is small compared to Methotrexate and Doxorubicin's drug efficacy parameters, resulting in Cisplatin having a rather minor effect on cancer reduction in the MAP treatment. Hence, the resistance to Doxorubicin matters more than the resistance to Cisplatin. Since the Cisplatin-resistance does not have a large impact on the effectiveness of MAP treatment, cancer cells that are resistant to both Doxorubicin and Cisplatin have similar dynamics to Doxorubicin-resistant cancer cells (Figure 5.4A, C).

5.2.4 Varying treatment start time

We study the effect of delays in the starting time of treatments on the tumors' responses to the treatments. Since the tumor growth rate depends on tumor size, we investigate the effects of delaying the chemotherapy treatment in small, medium and large tumors, separately. Small tumors are chosen as follows: we first choose the tumor with the smallest cancer population in cluster 1, then find the tumors in cluster 2 and 3 that have cancer population closest to the chosen tumor in cluster 1. Medium tumors are taken to be the mean values of all patients in each cluster. For large tumors, we take the without-treatment steady state values. We plot the dynamics of cancer population in each cluster when chemotherapy is started 1 week, which we assume is the earliest start time as it takes a few days to obtain biopsy results, 1 month, 3 months, and 6 months after the initial diagnosis.

Figure 5.5 and Table 5.4 show that in small and medium tumors, the cancer population after treatment is higher the longer we wait to start the chemotherapy. Thus, the earlier start of the chemotherapy leads to the better outcomes in these tumors. On the contrary, cancer population stays the same after treatment in

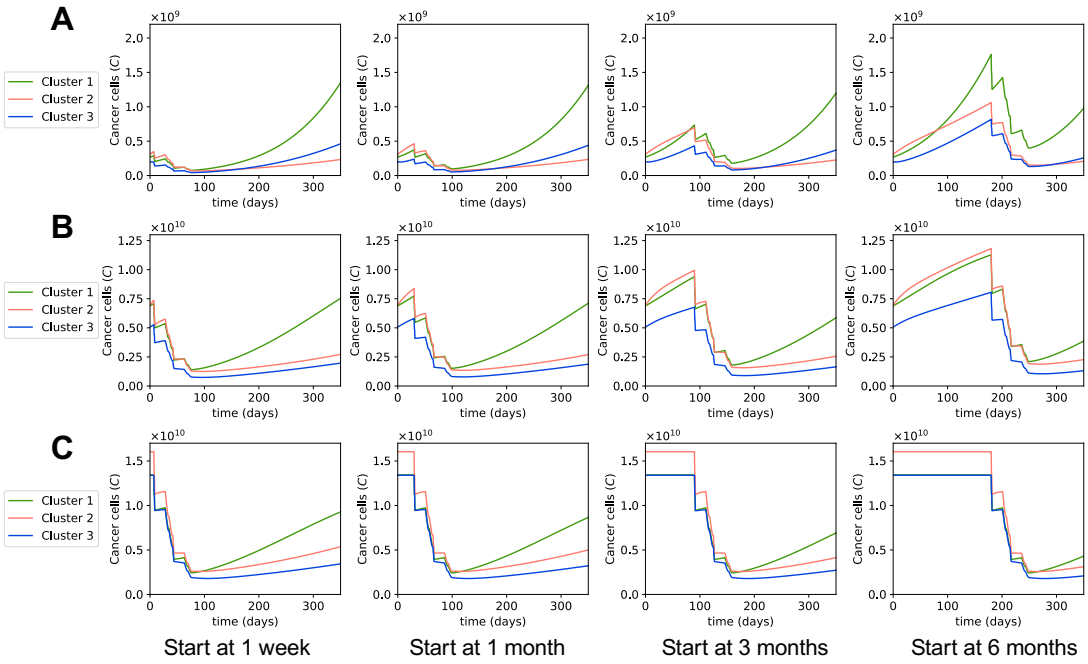

Figure 5.5. **Dynamics with different start times of MAP treatment.** Sub-figures A-C show the dynamics of cancer population for different MAP treatment's start times in small, medium, and large tumors, respectively. In each sub-figure, from left to right: the treatment starts at 1 week, 1 month, 3 months, and 6 months after initial diagnosis.

large tumors regardless of treatment start time. This is because these large tumors are at their steady states and do not grow more while the patient waits for the treatment. Then theoretically the treatment start time does not matter as much for tumors at the steady state or close to reaching the steady state. However, in reality, when tumors are large, the functionality of the cancerous body part is likely compromised, and the quality of the patient's life is affected, which makes us want to start the chemotherapy promptly for large tumors.

Table 5.4: Cancer population after MAP treatment with different treatment start times.

Tumor size	**Cluster**	**Initial cancer population**	**Cancer population after treatment**			
			Start at 1 week	Start at 1 month	Start at 3 months	Start at 6 months
Small	1	2.7×10^8	7.64×10^7	9.81×10^7	1.82×10^8	4.04×10^8
	2	3.07×10^8	6.73×10^7	7.76×10^7	1.05×10^8	1.52×10^8
	3	1.93×10^8	4.29×10^7	5.26×10^7	8.13×10^7	1.31×10^8
Medium	1	6.9×10^9	1.41×10^9	1.53×10^9	1.81×10^9	2.12×10^9
	2	6.96×10^9	1.27×10^9	1.37×10^9	1.6×10^9	1.9×10^9
	3	5.05×10^9	7.55×10^8	8.02×10^8	9.19×10^8	1.09×10^9
Large	1	1.34×10^{10}	2.44×10^9	2.44×10^9	2.44×10^9	2.44×10^9
	2	1.6×10^{10}	2.6×10^9	2.6×10^9	2.6×10^9	2.6×10^9
	3	1.34×10^{10}	1.87×10^9	1.87×10^9	1.87×10^9	1.87×10^9

It was previously observed in chapter 4 that tumors in cluster 1 grow fast even when the tumor is small, while tumors in clusters 2 and 3 start growing fast when the tumor is a bit bigger. Hence, when we delay the treatment for a long time, the small tumor in cluster 1 grows quickly and ends up with much higher cancer

population after the treatment than in other clusters, as seen in the treatments starting at 3 months and 6 months (Figure 5.5A, Table 5.4). For small tumors in clusters 2 and 3, even though they do grow while waiting for treatment, their growths are not as fast and their cancer populations after the treatment are still relatively small with the treatment delay. Therefore, it is important to start the treatment early for small tumors of cluster 1, and it would be ideal but not as urgent to start the treatment early for small tumors in clusters 2 and 3.

Figure 5.5B indicates that medium tumors in all three clusters grow comparably fast, and since their cancer populations after treatment are not very small, we should start chemotherapy for them as early as possible. We also notice that for small and medium tumors in all clusters, the difference in cancer population after the treatment is not significant between starting the treatment after 1 week or 1 month from the diagnosis. However, treatments starting at 3 months or 6 months result in much bigger cancer population after the treatment. Based on our model, it is thus not recommended to start the chemotherapy several months after the diagnosis, but rather to start within a month of the initial diagnosis.

5.2.5 Dynamics of cancer microenvironment with different treatment regimens

We investigate the effects of two other chemotherapy regimens on the osteosarcoma microenvironment. A combination of Doxorubicin and Cisplatin (AP) is a very common treatment of osteosarcoma tumors in older adults, as they are less likely to be able to tolerate high dose Methotrexate. This regimen consists of three preoperative 21-day cycles, where 25 mg/m^2 of Doxorubicin is given as a bolus once per day from day 1 to 3, and 100 mg/m^2 of Cisplatin is given as a continuous infusion over 24 hours on day 1 in each cycle [247, 248]. High dose Methotrexate

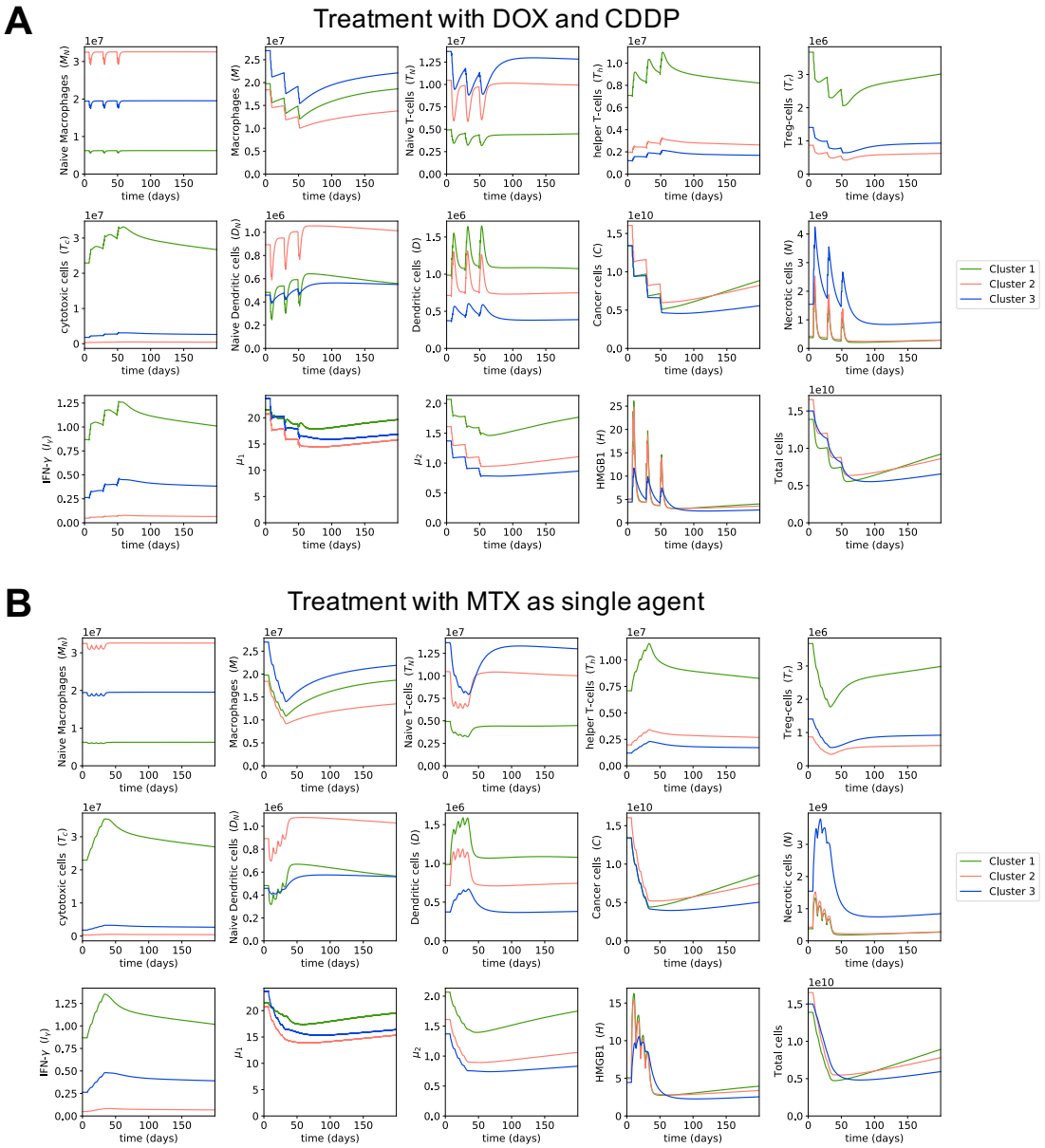

Figure 5.6. **Dynamics with different treatment regimens.** Sub-figure A shows the dynamics of cells and cytokines in osteosarcoma microenvironment in response to the combination of Doxorubicin and Cisplatin. Sub-figure B shows the dynamics of cells and cytokines in osteosarcoma microenvironment in response to a high dose of Methotrexate as a single agent.

(MTX) has also been used as a single agent to treat osteosarcoma, with 4 courses of 8 to 12 mg/m^2 given weekly before surgery [249]. In this study, we use the average dose which is 10 mg/m^2 of Methotrexate injected over 4 hours on day 1 every week for this regimen.

Figure 5.6 shows that MTX and AP regimens both have higher cancer population after treatment than MAP. This agrees with the finding from [250] that AP regimen is less effective but safer than MAP regimen. Meanwhile, MTX as a single agent is reported to be insufficient as a neoadjuvant therapy for osteosarcoma in [249], which uses the same MTX dosages and schedules as this study. Overall, according to our model, MAP is the superior treatment to MTX and AP in terms of cancer-killing ability. In fact, a recent study reports that MAP is still the favorable option for osteosarcoma among various combinations of chemotherapy drugs [251].

The AP regimen has relatively similar dynamics of cells and cytokines as the MAP regimen. That is, the populations of HMGB1, necrotic and dendritic cells increase when drugs are given and decrease when no drugs are given; while populations of helper T cells, cytotoxic cells and IFN-γ decrease at a smaller magnitude than their increase, so in general they increase during treatment; and regulatory T cells, macrophages, cytokines groups μ_1 and μ_2 generally decrease in abundance during treatment. The MTX treatment is given at closer intervals than AP and MAP treatments, so there is always some drug at the tumor site during MTX treatment. Therefore, the change in population of cells and cytokines over time for MTX regimen is smoother and does not fluctuate as much as the other two treatments, even though the dynamics of MTX regimen follow the same trend as them.

5.2.6 Optimal dosage for MAP treatment

Since neoadjuvant chemotherapy tries to reduce the boundaries of the tumor before surgery, we can choose the desired size of tumor for surgery and run our optimization framework to find the optimal dosage of chemotherapy drugs for the tumor to reach this size at a specific time. Osteosarcoma sizes vary greatly between patients at first diagnosis, and large tumors cannot reduce to the same size as small tumors after neoadjuvant treatment without exceeding the safe dosage of chemotherapy drugs. Thus, we choose different desired cancer population to optimize for depending on the size of tumor at first diagnosis. With MAP being the preferable treatment for osteosarcoma as mentioned in the previous section, here we present the optimal dosage of the MAP regimen for a large and a small tumor in each cluster, where the desired cancer population is 2.916×10^9 for large tumors and 1.36×10^8 for small tumors, which is equivalent to about 5cm per dimension (length, width, depth) for large tumors and 1.8cm per dimension for small tumors. We use 20000 mg/m^2 of Methotrexate, 45 mg/m^2 of Doxorubicin, and 75 mg/m^2 of Cisplatin per infusion day as maximum potential dosage, or equivalently 40000 mg/m^2 of Methotrexate, 90 mg/m^2 of Doxorubicin, and 150 mg/m^2 of Cisplatin per 35-day cycle.

Table 5.5: Optimal MAP dosages for large tumors.

Cluster	**Initial cancer population**	**Cancer population after treatment**	**MTX** (mg/m^2)	**DOX** (mg/m^2)	**CDDP** (mg/m^2)
1	1.34×10^{10}	2.916×10^9	8993	28	45
2	1.6×10^{10}	2.916×10^9	10134	32	51
3	1.34×10^{10}	2.916×10^9	6176	19	31

Table 5.6: Optimal MAP dosages for small tumors.

Cluster	**Initial cancer population**	**Cancer population after treatment**	**MTX** (mg/m^2)	**DOX** (mg/m^2)	**CDDP** (mg/m^2)
1	2.7×10^8	1.36×10^8	4926	15	25
2	3.07×10^8	1.36×10^8	4196	13	21
3	1.93×10^8	1.36×10^8	1305	3	6

The optimal dosages for large tumors are given in Table 5.5. Large tumors are taken to be the steady state values of each cluster. As cluster 2 has the highest cancer population at the steady state, it has the highest optimal dosage for each drug of the MAP treatment among all clusters. Interestingly, cluster 1 and 3 have the same cancer population before treatment, but cluster 1 needs higher dosage to achieve the same cancer population after treatment as cluster 3. This is due to the fact that cluster 1's cancer cells grow faster during treatment, so the same dosage of drugs would result in higher cancer population in cluster 1 than in cluster 3 after treatment, as seen in section 3.1. Thus, it is important to take tumor growth rate of the patient into account while finding the optimal dosage of chemotherapy.

Figure 5.7A shows that cancer cells in cluster 1 also grow fast after treatment, so it would be ideal to perform surgery quickly after neoadjuvant therapy. If it is impossible to start surgery promptly, we can choose a later time point to optimize for, so that at the time of surgery we still have the desired tumor size for resection. For example, if we cannot perform surgery until a month after chemotherapy, instead of using cancer population at day 80 for optimization, which is 3 days after the second cycle of chemotherapy, we can use cancer population at day 107 for optimization to find optimal dosage, which is 30 days after the second cycle of

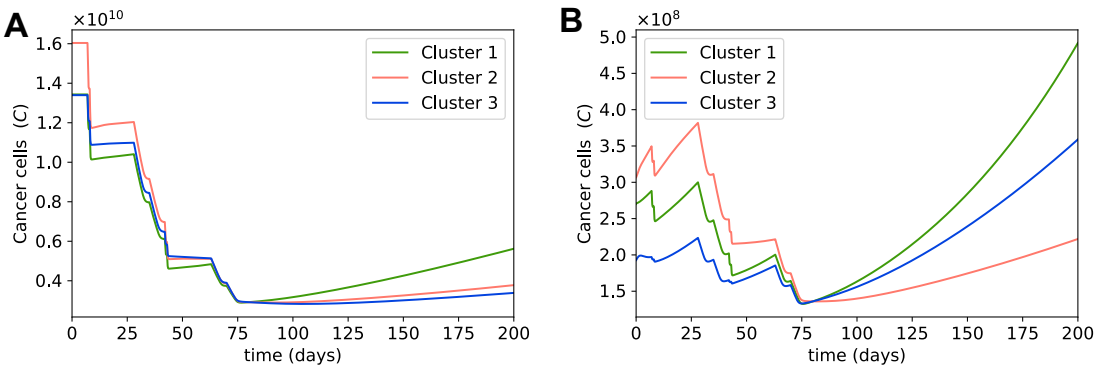

Figure 5.7. **Dynamics with optimal dosages for MAP treatment.** Sub-figure A shows the dynamics of cancer population in a large tumor in each cluster, where MAP dosages were optimized to obtain 2.916×10^9 cancer cells after treatment. Sub-figure B shows the dynamics of cancer population in a small tumor in each cluster, where MAP dosages were optimized to obtain 1.36×10^8 cancer cells after treatment.

chemotherapy. Then with the estimated optimal dosage, we will have the desired cancer population at day 107, which is the time of surgery.

The optimal dosages for a small tumor in each cluster are given in Table 5.6. Small tumors are chosen in the same method as described in section 5.2.4. The cancer populations in these small tumors are not much bigger than the desired cancer population after treatment, so in all clusters the optimal dosages for small tumors are much smaller than the optimal dosages for large tumors. Especially cluster 3, with cancer population before treatment very close to desired cancer population, has very small optimal dosages.

In many cases, even though the tumor is small enough for resection, neoadjuvant chemotherapy is still given to remove any potential metastases that are too small to be yet detected. Another reason for neoadjuvant chemotherapy in small tumors is to allow evaluation of the tumor response [209]. Figure 5.7B suggests that although chemotherapy does not reduce cancer populations significantly as

the cancer populations are already small to begin with, it helps prevent the cancer populations from growing bigger before surgery. Therefore, chemotherapy can also be used to control the growth of tumor while the patient waits for surgery.

Overall, with our optimization framework, we can find the optimal chemotherapy dosage to obtain the desired cancer population on the day of surgery. Our results show that it is important to consider each individual patient's cancer growth rate while computing optimal dosage, as patients with faster growth rate would need a higher dose.

5.3 Discussion

Our results indicate that besides reducing the number of cancer cells, chemotherapy induces specific behaviors in certain immune cells and cytokines by causing necrosis of cancer cells. In particular, the population of HMGB1 and dendritic cells increase when chemotherapy drugs are administered and decrease when these drugs are not given. In addition, helper T cells, cytotoxic cells and IFN-γ generally increase in population during treatment, which aligns with the findings from [81, 82, 214, 217, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 244]. Meanwhile, cells and cytokines that are not affected by this necrotic cell death, decrease in abundance due to being killed by chemotherapy drugs.

We note that it would be good to start chemotherapy early, unless the tumor is close to its steady state, as tumors of small and medium size will grow more while the patient waits for treatment. It is especially important to start chemotherapy promptly for tumors that grow fast such as those in cluster 1. Interestingly, we also notice that with the same initial cancer population and the same dosage, the cancer population after treatment is higher in cluster 1 than in cluster 3, where cluster

3 has slower cancer growth rate than cluster 1. All of these observations suggest that it is necessary to take the unique growth rate of the tumor into consideration when choosing the dosage and treatment start time for the patient, inferring the importance of personalized medicine.

In this study, we introduced a simple optimization framework to find the appropriate drug dosages to achieve a desired cancer population on a chosen day, such as the day of surgery. The results of our optimization also agrees with the above observation that the individual's cancer growth rate is essential for calculating optimal chemotherapy dosages. Since high doses of chemotherapy are known to have high toxicity and to induce many serious health problems [252, 253], it could be useful to use a mathematical model such as ours to estimate the appropriate dose rather than to give the standard dose to all tumor sizes, especially when small tumors are likely to need much smaller doses than the standard ones. Moreover, our model divides patients into groups based on their immune compositions, and thus can estimate their cancer growth more accurately than having one model for all patients, resulting in a more customized dosage recommendation for each patient.

Finding the right parameter values is a big challenge in mathematical modelling of cancers. While it would be ideal to acquire parameters by performing in vivo and in vitro experiments, these experiments are often expensive and time-consuming. Here, we use most chemotherapy-related parameters from a study that fitted these values to experimental data [211]. Therefore, our treatment parameters should be close to parameter values obtained by performing experiments, and more accurate than values chosen based on biological rationality or derived from assumptions. For the two parameters that we have to assume appropriate values for, we study their impact on the results through sensitivity analysis as well as by varying them, and find that different values of these parameters result in fairly similar outputs of our

model.

There are still some factors that our model does not account for. For instance, there are multiple levels of sensitivity of cancer cells to chemotherapy, which means two different patients can both be resistant to a chemotherapy drug, but one patient might be more sensitive than the other. Thus, the drug efficacy coefficient of Doxorubicin-/Cisplatin-resistant cells used in our model do not represent all Doxorubicin-/Cisplatin-resistant drug efficacies, as these parameters vary based on the level of resistance of the cells. However, our model is still useful for dose recommendations or for physicians to take into consideration while choosing between treatment options. Based on our model, a physician can monitor the tumor reduction throughout the treatment and adjust parameters such as drug efficacy coefficients according to how the tumor responds to treatment.

Another factor that our model did not consider is the different rates of drug-induced immune cell death. Since chemotherapy target cells with a faster metabolic rate more successfully [216], it is reasonable to expect that the death rates by chemotherapy differ between types of immune cells. Therefore, one idea to better our model is to update the rates of chemotherapy-induced death of immune cells in proportion to their metabolic rates. Other ways to improve upon this work include adding other chemotherapy drugs like Ifosfamide, which is also a commonly used drug for osteosarcoma [209, 210]; extending to a different treatment option besides chemotherapy such as radiotherapy and immunotherapy; or using partial differential equations to take into account the spatial distribution of the tumors as well.

CHAPTER 6

CONCLUSION

Due to the high cost of flow and mass cytometry, there has been a recent surge in the development of computational methods for estimating the relative distributions of cell types from the gene expression profile of a bulk of cells. These computational methods are referred to as digital cytometry methods. In chapter 2, we compare five common digital cytometry methods, including three linear models and two rank-based methods, on simulation data, whole blood RNA-Seq data, and PBMC microarray data.

Rank-based methods ssGSEA DM and SingScore DM give conflicting results between sample-level and cell-level correlation with ground truth fractions, and overall perform worse than linear methods. DeconRNASeq's performance depends heavily on how comparable the signature matrix and mixture data are. CIBERSORT and CIBERSORTx B-mode perform the best among all mentioned methods based on sample-level and cell-level Pearson and Spearman correlation with ground truth cell fractions for all three data sets, regardless of the signature matrix used. CIBERSORTx B-mode, which uses batch correction on the mixture data, tends to slightly outperform CIBERSORT, especially when signature matrix and mixture data come from different platforms. However, CIBERSORTx S-mode, which uses batch correction on the signature matrix, does not perform as well. This suggests

further investigation into the way batch correction is used for adjusting the signature matrix in order to eliminate the technical variations between signature matrix and mixture data.

As the tumor immune microenvironment has been shown to be important in predicting the tumor progression and the outcome of treatments, in chapter 3, we study different immune patterns in osteosarcoma and their clinical characteristics. We use the best performing digital cytometry method reported in chapter 2, CIBERSORTx B-mode, to obtain the relative abundance of 22 immune cells in osteosarcoma primary tumors. Then we cluster patients based on their estimated immune abundance and analyze the characteristics of these clusters, along with the relationship between immune infiltration and outcome of patients.

Three clusters of osteosarcoma patients with distinct immune compositions are found. We observe that abundance of cytotoxic T cells, NK cells and M1 Macrophages have a positive association with prognosis, while abundance of $\gamma\delta$ T cells, Mast cells, M0 Macrophages and Dendritic cells have a negative association with prognosis. Accordingly, cluster 2, with the highest proportion of M0 macrophages and the lowest proportion of cytotoxic T cells and M1 macrophages, has the worst outcome among clusters. Cluster 3, with higher levels of cytotoxic T cells, M1 macrophages and lower level of M0 macrophages than cluster 2, appears to have the best outcome. Cluster 1, which has both high levels of immune cells with good prognosis and of immune cells with poor prognosis, has better outcome than cluster 2 but worse outcome than cluster 3.

Using the same cluster assignment as chapter 3, in chapter 4, we build a data-driven mathematical model of osteosarcoma progression while taking into account the interactions between immune cells and cancer cells. Based on our model, out of the three clusters, cluster 3 appears to have the most favorable tumor growth, and

cluster 2 has the least favorable growth. This observation agrees with the findings from clinical information of these clusters in chapter 3, as mentioned above. We also find that during osteosarcoma progression, the number of dendritic cells, helper T cells, cytotoxic cells, and the amount of IFN-γ first increase and then decrease, while the regulatory T cell population decreases and then increases. This switch in the dynamics of immune cells and cytokines happens around the same time that cancer cells have the fastest growth.

The global sensitivity analysis in chapter 4 indicates that the cancer death rates by cytotoxic cells and IFN-γ, the cancer proliferation rates by cytokines groups μ_1 and μ_2, as well as the cancer self-proliferation and apoptosis rates are the most impactful parameters on cancer growth. Additionally, among all immune parameters, the activation and decay rates of macrophages and regulatory T cells have the most impact on cancer growth. The results of chapter 4 also suggest that it is necessary to investigate the complex interactions between immune cells and cancer cells instead of purely looking at the abundance of certain immune cells as a marker for the disease's progression.

In chapter 5, we extend the model in chapter 4 to include the interactions between the key components of osteosarcoma microenvironment and common chemotherapy drugs: Methotrexate, Doxorubicin and Cisplatin. We find that during the chemotherapy treatment, dendritic cells and HMGB1 increase in population when drugs are given and decrease in population while the patient waits for the next dose of drugs, while helper T cells, cytotoxic cells and IFN-γ have an overall increase in abundance. Other cells and cytokines of the microenvironment, which are not under the impact of necrotic cell death, have reduced populations after the treatment. Overall, the MAP regimen, which consists of all three drugs above, is effective at minimizing the number of cancer cells, and works better than Methotrexate alone

or a combination of Doxorubicin and Cisplatin.

We observe that it is important to start chemotherapy treatment quickly after diagnosis, unless the tumor is close to its steady state, as the tumor can grow while the patient waits for treatments. Interestingly, with the same dosage and treatment start time, tumors in cluster 1 have a higher cancer population after treatment than those in cluster 3, because cluster 1's tumors have a faster growth rate. Accordingly, tumors in cluster 1 need higher dosages to achieve the same results as tumors in cluster 3. These findings from chapter 5 highlight the importance of considering the individual growth of the tumor when deciding on the dosage and treatment start time for a patient, as fast growing tumors require higher dosages and an earlier start to treatment than slow growing tumors.

A P P E N D I X A

SUPPORTING INFORMATION FOR CHAPTER 4

A.1 System Analysis

A.1.1 Full dimensional system of ODEs

Combining Equations (4.1)–(4.16) we obtain the following system of ODEs:

$$\frac{d\left[M_N\right]}{dt} = A_{M_N} - \left(\lambda_{MI_\gamma}\left[I_\gamma\right] + \lambda_{M\mu_1}\left[\mu_1\right]\right)\left[M_N\right] - \delta_{M_N}\left[M_N\right], \tag{A.1}$$

$$\frac{d\left[M\right]}{dt} = \left(\lambda_{MI_\gamma}\left[I_\gamma\right] + \lambda_{M\mu_1}\left[\mu_1\right]\right)\left[M_N\right] - \delta_{M}\left[M\right], \tag{A.2}$$

$$\begin{aligned}\frac{d\left[T_N\right]}{dt} = {} & A_{T_N} - \left(\lambda_{T_hM}\left[M\right] + \lambda_{T_hD}\left[D\right]\right)\left[T_N\right] \\ & - \lambda_{T_r\mu_1}\left[\mu_1\right]\left[T_N\right] \\ & - \left(\lambda_{T_cT_h}\left[T_h\right] + \lambda_{T_cM}\left[M\right] + \lambda_{T_cD}\left[D\right]\right)\left[T_N\right] - \delta_{T_N}\left[T_N\right],\end{aligned} \tag{A.3}$$

$$\frac{d\left[T_h\right]}{dt} = \left(\lambda_{T_hM}\left[M\right] + \lambda_{T_hD}\left[D\right]\right)\left[T_N\right] - \left(\delta_{T_hT_r}\left[T_r\right] + \delta_{T_h\mu_1}\left[\mu_1\right] + \delta_{T_h}\right)\left[T_h\right], \tag{A.4}$$

$$\frac{d\left[T_r\right]}{dt} = \left(\lambda_{T_r\mu_1}\left[\mu_1\right]\right)\left[T_N\right] - \delta_{T_r}\left[T_r\right], \tag{A.5}$$

$$\begin{aligned}\frac{d\left[T_c\right]}{dt} = {} & \left(\lambda_{T_cT_h}\left[T_h\right] + \lambda_{T_cM}\left[M\right] + \lambda_{T_cD}\left[D\right]\right)\left[T_N\right] \\ & - \left(\delta_{T_cT_r}\left[T_r\right] + \delta_{T_c\mu_1}\left[\mu_1\right] + \delta_{T_c}\right)\left[T_c\right],\end{aligned} \tag{A.6}$$

$$\frac{d\left[D_N\right]}{dt} = A_{D_N} - \left(\lambda_{DC}\left[C\right] + \lambda_{DH}\left[H\right]\right)\left[D_N\right] - \delta_{D_N}\left[D_N\right], \tag{A.7}$$

$$\frac{d\left[D\right]}{dt} = \left(\lambda_{DC}\left[C\right] + \lambda_{DH}\left[H\right]\right)\left[D_N\right] - \left(\delta_{DC}\left[C\right] + \delta_{D}\right)\left[D\right], \tag{A.8}$$

$$\frac{d[C]}{dt} = (\lambda_C + \lambda_{C\mu_1}[\mu_1] + \lambda_{C\mu_2}[\mu_2])[C]\left(1 - \frac{[C]}{C_0}\right) - \left(\delta_{CT_c}[T_c] + \delta_{CI_\gamma}[I_\gamma] + \delta_C\right)[C], \tag{A.9}$$

$$\frac{d[N]}{dt} = \alpha_{NC}\left(\delta_{CT_c}[T_c] + \delta_{CI_\gamma}[I_\gamma] + \delta_C\right)[C] - \delta_N[N], \tag{A.10}$$

$$\frac{d[I_\gamma]}{dt} = \lambda_{I_\gamma T_h}[T_h] + \lambda_{I_\gamma T_c}[T_c] - \delta_{I_\gamma}[I_\gamma], \tag{A.11}$$

$$\frac{d[\mu_1]}{dt} = \lambda_{\mu_1 T_h}[T_h] + \lambda_{\mu_1 M}[M] + \lambda_{\mu_1 C}[C] - \delta_{\mu_1}[\mu_1], \tag{A.12}$$

$$\frac{d[\mu_2]}{dt} = \lambda_{\mu_2 T_h}[T_h] + \lambda_{\mu_2 M}[M] + \lambda_{\mu_2 C}[C] - \delta_{\mu_2}[\mu_2], \tag{A.13}$$

$$\frac{d[H]}{dt} = \lambda_{HM}[M] + \lambda_{HD}[D] + \lambda_{HN}[N] - \delta_H[H]. \tag{A.14}$$

A.1.2 Proof of positivity

To prove that the system with positive coefficients and positive initial conditions has a positive solution, let us consider the set of integrating factors, one for each variable:

$$\eta_{M_N}(t) = \exp\int_0^t \left(\lambda_{MI_\gamma}[I_\gamma] + \lambda_{M\mu_1}[\mu_1] + \delta_{M_N}\right) ds$$

$$\eta_{T_N}(t) = \exp\int_0^t (\lambda_{T_hM}[M] + \lambda_{T_hD}[D] + \lambda_{T_r\mu_1}[\mu_1] + \lambda_{T_cT_h}[T_h] + \lambda_{T_cM}[M] + \lambda_{T_cD}[D] + \delta_{T_N})\, ds$$

$$\eta_{T_h}(t) = \exp\int_0^t (\delta_{T_hT_r}[T_r] + \delta_{T_h\mu_1}[\mu_1] + \delta_{T_h})\, ds$$

$$\eta_{T_C}(t) = \exp\int_0^t (\delta_{T_cT_r}[T_r] + \delta_{T_c\mu_1}[\mu_1] + \delta_{T_c})\, ds$$

$$\eta_{D_N}(t) = \exp\int_0^t (\lambda_{DC}[C] + \lambda_{DH}[H] + \delta_{D_N})\, ds$$

$$\eta_D(t) = \exp \int_0^t \left(\delta_{DC}[C] + \delta_D\right) ds$$

$$\eta_C(t) = \exp \int_0^t \left(\delta_{CT_c}[T_c] + \delta_{CI_\gamma}[I_\gamma] + \delta_C - \left(\lambda_C + \lambda_{C\mu_1}[\mu_1] + \lambda_{C\mu_2}[\mu_2]\right)\left(1 - \frac{[C]}{C_0}\right)\right) ds$$

$$\eta_M(t) = \exp\left(\delta_M t\right), \quad \eta_{T_r}(t) = \exp\left(\delta_{T_r} t\right), \quad \eta_N(t) = \exp\left(\delta_N t\right), \quad \eta_H(t) = \exp\left(\delta_H t\right),$$

$$\eta_{I_\gamma}(t) = \exp\left(\delta_{I_\gamma} t\right), \quad \eta_{\mu_1}(t) = \exp\left(\delta_{\mu_1} t\right), \quad \eta_{\mu_2}(t) = \exp\left(\delta_{\mu_2} t\right).$$

These integrating factors will not allow us to derive explicit solution as some of them are defined through the unknown variables. However, it is important to note that the factors are strictly positive and allow us to rewrite the system as:

$$\begin{aligned}
\frac{d\left([M_N]\,\eta_{M_N}\right)}{dt} &= A_{M_N}\eta_{M_N}, \\
\frac{d\left([M]\,\eta_M\right)}{dt} &= \left(\lambda_{MI_\gamma}[I_\gamma] + \lambda_{M\mu_1}[\mu_1]\right)[M_N]\,\eta_M, \\
\frac{d\left([T_N]\,\eta_{T_N}\right)}{dt} &= A_{T_N}\eta_{T_N}, \\
\frac{d\left([T_h]\,\eta_{T_h}\right)}{dt} &= \left(\lambda_{T_hM}[M] + \lambda_{T_hD}[D]\right)[T_N]\,\eta_{T_h}, \\
\frac{d\left([T_r]\,\eta_{T_r}\right)}{dt} &= \left(\lambda_{T_r\mu_1}[\mu_1]\right)[T_N]\,\eta_{T_r}, \\
\frac{d\left([T_c]\,\eta_{T_c}\right)}{dt} &= \left(\lambda_{T_cT_h}[T_h] + \lambda_{T_cM}[M] + \lambda_{T_cD}[D]\right)[T_N]\,\eta_{T_c}, \\
\frac{d\left([D_N]\,\eta_{D_N}\right)}{dt} &= A_{D_N}\eta_{D_N}, \\
\frac{d\left([D]\,\eta_D\right)}{dt} &= \left(\lambda_{DC}[C] + \lambda_{DH}[H]\right)[D_N]\,\eta_D, \\
\frac{d\left([C]\,\eta_C\right)}{dt} &= 0, \\
\frac{d\left([N]\,\eta_N\right)}{dt} &= \alpha_{NC}\left(\delta_{CT_c}[T_c] + \delta_{CI_\gamma}[I_\gamma] + \delta_C\right)[C]\,\eta_N, \\
\frac{d\left([I_\gamma]\,\eta_{I_\gamma}\right)}{dt} &= \left(\lambda_{I_\gamma T_h}[T_h] + \lambda_{I_\gamma T_c}[T_c]\right)\eta_{I_\gamma}, \\
\frac{d\left([\mu_1]\,\eta_{\mu_1}\right)}{dt} &= \left(\lambda_{\mu_1 T_h}[T_h] + \lambda_{\mu_1 M}[M] + \lambda_{\mu_1 C}[C]\right)\eta_{\mu_1}, \\
\frac{d\left([\mu_2]\,\eta_{\mu_2}\right)}{dt} &= \left(\lambda_{\mu_2 T_h}[T_h] + \lambda_{\mu_2 M}[M] + \lambda_{\mu_2 C}[C]\right)\eta_{\mu_2},
\end{aligned}$$

$$\frac{d\left([H]\,\eta_H\right)}{dt} = \left(\lambda_{HM}\,[M] + \lambda_{HD}\,[D] + \lambda_{HN}\,[N]\right)\eta_H.$$

We see that the right-hand side of each equation in this system is non-negative, which means that the variable-factor product $([X]\,\eta_X)$ is non-decreasing for each variable $[X]$, and thus, if positive, initially remains positive at all times. Since the integrating factor is positive by design, the positivity of the variables follows.

A.1.3 Proof of boundedness

Macrophages

Adding Equations (A.1) and (A.2), we obtain

$$\frac{d\left([M_N] + [M]\right)}{dt} = A_{M_N} - \delta_{M_N}\,[M_N] - \delta_M\,[M] \leq A_{M_N} - \min\left(\delta_{M_N},\,\delta_M\right)\left([M_N] + [M]\right).$$

Thus, integrating, we obtain

$$\begin{aligned}[M_N] + [M] \leq & \frac{A_{M_N}}{\min\left(\delta_{M_N},\,\delta_M\right)}\left(1 - e^{-\min\left(\delta_{M_N},\delta_M\right)t}\right) \\ & + e^{-\min\left(\delta_{M_N},\delta_M\right)t}\left([M_N]\,(0) + [M]\,(0)\right).\end{aligned}$$

Since the right-hand side is bounded and each variable is positive, this proves that each variable is bounded.

T-cells

Adding Equations (A.3)–(A.6) and using the positivity of all variables, we obtain

$$\begin{aligned}\frac{d\left([T_N] + [T_h] + [T_r] + [T_c]\right)}{dt} = & A_{T_N} - \delta_{T_N}\,[T_N] - \delta_{T_r}\,[T_r] \\ & - \left(\delta_{T_hT_r}\,[T_r] + \delta_{T_h\mu_1}\,[\mu_1] + \delta_{T_h}\right)[T_h] \\ & - \left(\delta_{T_cT_r}\,[T_r] + \delta_{T_c\mu_1}\,[\mu_1] + \delta_{T_c}\right)[T_c]\end{aligned}$$

$$\leq A_{T_N} - \min\left(\delta_{T_N},\, \delta_{T_h},\, \delta_{T_c},\, \delta_{T_r}\right)\left([T_N] + [T_h] + [T_r] + [T_c]\right).$$

Then, integrating, we obtain

$$\begin{aligned}[T_N] + [T_h] + [T_r] + [T_c] \leq & \frac{A_{T_N}}{\min\left(\delta_{T_N},\, \delta_{T_h},\, \delta_{T_c},\, \delta_{T_r}\right)}\left(1 - e^{-\min\left(\delta_{T_N}, \delta_{T_h}, \delta_{T_c}, \delta_{T_r}\right)t}\right) \\ & + e^{-\min\left(\delta_{T_N}, \delta_{T_h}, \delta_{T_c}, \delta_{T_r}\right)t}\left([T_N](0) + [T_h](0) + [T_r](0) + [T_c](0)\right).\end{aligned}$$

Since the right-hand side is bounded and each variable is positive, this proves that each variable is bounded.

Dendritic cells

Adding Equations (A.7) and (A.8) and using the positivity of $[C]$, we obtain

$$\begin{aligned}\frac{d\left([D_N] + [D]\right)}{dt} = & A_{D_N} - \delta_{D_N}[D_N] - \left(\delta_{DC}[C] + \delta_D\right)[D] \\ \leq & A_{D_N} - \min\left(\delta_{D_N},\, \delta_D\right)\left([D_N] + [D]\right).\end{aligned}$$

Similar to the previous cases, integrated bound

$$[D_N]+[D] \leq \frac{A_{D_N}}{\min\left(\delta_{D_N},\, \delta_D\right)}\left(1 - e^{-\min\left(\delta_{D_N}, \delta_D\right)t}\right)+e^{\min\left(\delta_{D_N}, \delta_D\right)t}\left([D_N](0) + [D](0)\right)$$

proves the upper bound on $[D_N]$ and $[D]$.

Cancer cells

Let us rewrite Equation (A.9) as follows

$$\begin{aligned}\frac{d\left([C] - C_0\right)}{dt} + & \frac{\left(\lambda_C + \lambda_{C\mu_1}[\mu_1] + \lambda_{C\mu_2}[\mu_2]\right)[C]}{C_0}\left([C] - C_0\right) \\ & = -\left(\delta_{CT_c}[T_c] + \delta_{CI_\gamma}[I_\gamma] + \delta_C\right)[C] \leq 0.\end{aligned}$$

Integrating the inequality with implicit dependence on $[C]$, $[\mu_1]$, and $[\mu_2]$, we obtain

$$[C] \leq C_0 - \left(C_0 - [C](0)\right)\exp\left(-\int_0^t \frac{\left(\lambda_C + \lambda_{C\mu_1}[\mu_1](s) + \lambda_{C\mu_2}[\mu_2](s)\right)[C](s)}{C_0} ds\right).$$

Since $[C]$, $[\mu_1]$, and $[\mu_2]$ are proven to be positive, the right-hand side is bounded and, thus, $[C]$ is bounded.

Interferon-γ

We require the bound on interferon before proving the bound on necrotic cells. Since $[T_h]$ and $[T_c]$ are proven to be bounded, we could claim that

$$\lambda_{I_\gamma T_h}[T_h] + \lambda_{I_\gamma T_c}[T_c] \leq \lambda_{I_\gamma}^{\max}.$$

This, together with Equation (A.11), yields the following inequality:

$$\frac{d[I_\gamma]}{dt} + \delta_{I_\gamma}[I_\gamma] \leq \lambda_{I_\gamma}^{\max},$$

which, when integrated, gives the upper bound on $[I_\gamma]$ as follows:

$$[I_\gamma] \leq \frac{\lambda_{I_\gamma}^{\max}}{\delta_{I_\gamma}}\left(1 - e^{-\delta_{I_\gamma} t}\right) + e^{-\delta_{I_\gamma} t}[I_\gamma](0).$$

Remaining variables

For each of the remaining variables, the bounds proven above result in the upper bounds for the positive parts of the right-hand side in each equation as follows

$$\begin{aligned}
\alpha_{NC}\left(\delta_{CT_c}[T_c] + \delta_{CI_\gamma}[I_\gamma] + \delta_C\right)[C] &\leq \lambda_N^{\max},\\
\lambda_{\mu_1 T_h}[T_h] + \lambda_{\mu_1 M}[M] + \lambda_{\mu_1 C}[C] &\leq \lambda_{\mu_1}^{\max},\\
\lambda_{\mu_2 T_h}[T_h] + \lambda_{\mu_2 M}[M] + \lambda_{\mu_2 C}[C] &\leq \lambda_{\mu_2}^{\max},\\
\lambda_{HM}[M] + \lambda_{HD}[D] + \lambda_{HN}[N] &\leq \lambda_H^{\max}.
\end{aligned}$$

Then, Equations (A.10) and (A.12)–(A.14) result in the following differential inequalities

$$\frac{d[N]}{dt} + \delta_N[N] \leq \lambda_N^{\max},$$

$$\frac{d\left[\mu_1\right]}{dt} + \delta_{\mu_1}\left[\mu_1\right] \leq \lambda_{\mu_1}^{\max},$$
$$\frac{d\left[\mu_2\right]}{dt} + \delta_{\mu_2}\left[\mu_2\right] \leq \lambda_{\mu_2}^{\max},$$
$$\frac{d\left[H\right]}{dt} + \delta_{H}\left[H\right] \leq \lambda_{H}^{\max}.$$

Integrating, we obtain

$$[N] \leq \frac{\lambda_N^{\max}}{\delta_N}\left(1 - e^{-\delta_N t}\right) + e^{-\delta_N t}\left[N\right](0),$$
$$[\mu_1] \leq \frac{\lambda_{\mu_1}^{\max}}{\delta_{\mu_1}}\left(1 - e^{-\delta_{\mu_1} t}\right) + e^{-\delta_{\mu_1} t}\left[\mu_1\right](0),$$
$$[\mu_2] \leq \frac{\lambda_{\mu_2}^{\max}}{\delta_{\mu_2}}\left(1 - e^{-\delta_{\mu_2} t}\right) + e^{-\delta_{\mu_2} t}\left[\mu_2\right](0),$$
$$[H] \leq \frac{\lambda_H^{\max}}{\delta_H}\left(1 - e^{-\delta_H t}\right) + e^{-\delta_H t}\left[H\right](0),$$

thus, proving the upper bounds.

A.2 Derivation of the parameter set

A.2.1 Assumptions on parameters

We adopt natural degradation/decay rates of immune cells and cytokines based on information about their half life from the literature (see Table A.1). For example, the degradation/decay rate of X is calculated as:

$$\delta_X = \frac{ln2}{\text{half life of X in days}}$$

The decay rate of μ_1 is estimated to be a weighted average of the decay rates of cytokines within μ_1, where the weights are proportional to the abundance of these cytokines. A similar procedure is carried out for μ_2. The obtained natural decay rates are as follows:

$$\delta_{M_N} = 0.693, \qquad \delta_M = 0.015, \qquad \delta_O = 1.219, \qquad \delta_{T_N} = 0.00042,$$

$$\delta_{T_h} = 0.231, \qquad \delta_{T_r} = 0.063, \qquad \delta_{T_c} = 0.406, \qquad \delta_{D_N} = 1.664,$$

$$\delta_D = 0.277, \qquad \delta_{I_\gamma} = 33.27, \qquad \delta_{\mu_1} = 487.48, \qquad \delta_{\mu_2} = 5.15,$$

$$\delta_H = 58.7$$

For the proliferation rate of tumor cells, we gathered information on osteosarcoma growth in humans. A study reported that the mean exponential growth constant of primary osteosarcoma tumors was between 0.0054 and 0.02784 [254]. We took the average of these values and chose $\lambda_C = 0.01662$. Then, we made the assumption that the proliferation rate of cancer cells themselves was 20 times larger than the proliferation rate of cancer for the cytokines group μ_2; that is,

$$\lambda_C = 20\lambda_{C\mu_2}\mu_2^{\text{mean}} \tag{A.15}$$

μ_2 consists of IL-6, which is a major pro-tumor cytokine; therefore, we assume that μ_2 is twice as effective at promoting tumor growth compared with μ_1:

$$\lambda_{C\mu_2}\mu_2^{\text{mean}} = 2\lambda_{C\mu_1}\mu_1^{\text{mean}}$$

$$\text{or equivalently, } \lambda_C = 40\lambda_{C\mu_1}\mu_1^{\text{mean}} \tag{A.16}$$

We also assume that cytotoxic cells kill tumor cells twice as fast as IFN-γ, and IFN-γ is 10-times more effective at killing cancer cells compared with the cancer cell natural death rate:

$$\delta_{CI_\gamma} I_\gamma^{\text{mean}} = 10\delta_C \tag{A.17}$$

$$\delta_{CT_c} T_c^{\text{mean}} = 2\delta_{CI_\gamma} I_\gamma^{\text{mean}} \text{ or } \delta_{CT_c} T_c^{\text{mean}} = 20\delta_C \tag{A.18}$$

Since M1 and M2 macrophages are activated solely by IFN-γ and μ_1, respectively, we assume that the ratio of macrophages activated by IFN-γ to macrophages activated by μ_1 equals to the ratio of M1 to M2 macrophages:

$$\frac{\lambda_{MI_\gamma} I_\gamma^{\text{mean}}}{\lambda_{M\mu_1}\mu_1^{\text{mean}}} = \frac{M_1^{\text{mean}}}{M_2^{\text{mean}}} \tag{A.19}$$

We make the assumption that helper T cells are predominantly activated by antigen presenting dendritic cells and that the inhibition of helper T cells by regulatory T cells and by μ_1 are more effective than natural decay:

$$\lambda_{T_h D} D^{\text{mean}} = 200 \lambda_{T_h M} M^{\text{mean}} \tag{A.20}$$

$$\delta_{T_h T_r} T_r^{\text{mean}} = \delta_{T_h \mu_1} \mu_1^{\text{mean}} = 20 \delta_{T_h} \tag{A.21}$$

We also assume that the activation of cytotoxic cells by dendritic cells (both through antigen presentation and IL-12) is twice as effective compared with activation by helper T cells, and four-times as effective compared with activation by macrophages (through IL-12):

$$\lambda_{T_c D} D^{\text{mean}} = 2 \lambda_{T_c T_h} T_h^{\text{mean}} = 4 \lambda_{T_c M} M^{\text{mean}} \tag{A.22}$$

while the inhibition of cytotoxic cells by regulatory T cells and by μ_1 are each 20-times larger than with natural decay:

$$\delta_{T_c T_r} T_r^{\text{mean}} = \delta_{T_c \mu_1} \mu_1^{\text{mean}} = 20 \delta_{T_c} \tag{A.23}$$

For dendritic cells, we make the assumption that activation by HMGB1 is twice as effective compared with activation by cancer cells and that the inhibition by cancer cells is equivalent to the dendritic cells' innate decay rate:

$$\lambda_{DH} H^{\text{mean}} = 2 \lambda_{DC} C^{\text{mean}} \tag{A.24}$$

$$\delta_{DC} C^{\text{mean}} = \delta_D \tag{A.25}$$

Additionally, the following assumptions were used for the production rates of cytokines:

$$\lambda_{I_\gamma T_c} T_c^{\text{mean}} = 4 \lambda_{I_\gamma T_h} T_h^{\text{mean}} \tag{A.26}$$

$$\lambda_{\mu_1 T_h} T_h^{\text{mean}} = \lambda_{\mu_1 M} M^{\text{mean}} = \lambda_{\mu_1 C} C^{\text{mean}} \tag{A.27}$$

$$\lambda_{\mu_2 M} M^{\text{mean}} = \lambda_{\mu_2 C} C^{\text{mean}} = 2\lambda_{\mu_2 T_h} T_h^{\text{mean}} \tag{A.28}$$

$$\lambda_{HN} N^{\text{mean}} = 10\lambda_{HM} M^{\text{mean}} = 20\lambda_{HD} D^{\text{mean}} \tag{A.29}$$

Lastly, we assume that $\alpha_{NC} = 3/4$ and that carrying capacity of cancer is twice the steady state value of cancer, that is $C_0 = 2C^{\infty}$.

A.2.2 Parameter values and sources

Table A.1: Non-dimensional parameter values for each cluster.

Parameter	Cluster 1	Cluster 2	Cluster 3	Source
$\overline{\lambda}_{MI_\gamma}$	4.3649×10^{-3}	8.4234×10^{-4}	2.8083×10^{-3}	Estimated
$\overline{\lambda}_{M\mu_1}$	1.0635×10^{-2}	1.4158×10^{-2}	1.2192×10^{-2}	Estimated
$\overline{\lambda}_{T_h M}$	3.3434×10^{-2}	1.9270×10^{-2}	2.2194×10^{-2}	Estimated
$\overline{\lambda}_{T_h D}$	1.0963×10	7.3778	9.8325	Estimated
$\overline{\lambda}_{T_r \mu_1}$	6.3×10^{-2}	6.3×10^{-2}	6.3×10^{-2}	Estimated
$\overline{\lambda}_{T_c T_h}$	6.1171	2.3846	2.8415	Estimated
$\overline{\lambda}_{T_c M}$	1.7478	1.2263	1.4683	Estimated
$\overline{\lambda}_{T_c D}$	1.1463×10	9.3900	1.3011×10	Estimated
$\overline{\lambda}_{DC}$	4.0114×10^{-1}	4.8942×10^{-1}	5.9472×10^{-1}	Estimated
$\overline{\lambda}_{DH}$	4.1518×10^{-1}	4.2621×10^{-1}	4.1729×10^{-1}	Estimated
λ_C	1.662×10^{-2}	1.662×10^{-2}	1.662×10^{-2}	[254]
$\overline{\lambda}_{C\mu_1}$	3.7101×10^{-4}	3.5910×10^{-4}	4.0692×10^{-4}	Estimated
$\overline{\lambda}_{C\mu_2}$	7.1405×10^{-4}	6.3207×10^{-4}	5.7910×10^{-4}	Estimated
$\overline{\lambda}_{I_\gamma T_h}$	6.3095	1.1946×10	4.1848	Estimated
$\overline{\lambda}_{I_\gamma T_c}$	2.6961×10	2.1324×10	2.9085×10	Estimated

$\overline{\lambda}_{\mu_1 T_h}$	1.8813×10^2	1.1491×10^2	1.0315×10^2	Estimated
$\overline{\lambda}_{\mu_1 M}$	1.0751×10^2	1.1818×10^2	1.0661×10^2	Estimated
$\overline{\lambda}_{\mu_1 C}$	1.9184×10^2	2.5440×10^2	2.7772×10^2	Estimated
$\overline{\lambda}_{\mu_2 T_h}$	1.2313	6.8806×10^{-1}	6.0936×10^{-1}	Estimated
$\overline{\lambda}_{\mu_2 M}$	1.4073	1.4153	1.2595	Estimated
$\overline{\lambda}_{\mu_2 C}$	2.5113	3.0467	3.2811	Estimated
$\overline{\lambda}_{HM}$	4.4046	5.8355	1.8254	Estimated
$\overline{\lambda}_{HD}$	3.6107	5.5856	2.0218	Estimated
$\overline{\lambda}_{HN}$	5.0685×10	4.7279×10	5.4853×10	Estimated
δ_{M_N}	6.93×10^{-1}	6.93×10^{-1}	6.93×10^{-1}	[255, 256, 257]
δ_M	1.5×10^{-2}	1.5×10^{-2}	1.5×10^{-2}	[258, 259]
δ_{T_N}	4.2×10^{-4}	4.2×10^{-4}	4.2×10^{-4}	[260]
$\overline{\delta}_{T_h T_r}$	6.6404	3.1732	5.0991	Estimated
$\overline{\delta}_{T_h \mu_1}$	4.1253	3.9929	4.5246	Estimated
δ_{T_h}	2.31×10^{-1}	2.31×10^{-1}	2.31×10^{-1}	[261]
δ_{T_r}	6.3×10^{-2}	6.3×10^{-2}	6.3×10^{-2}	[262]
$\overline{\delta}_{T_c T_r}$	1.1671×10	5.5771	8.9620	Estimated
$\overline{\delta}_{T_c \mu_1}$	7.2505	7.0179	7.9524	Estimated
δ_{T_c}	4.06×10^{-1}	4.06×10^{-1}	4.06×10^{-1}	[261]
δ_{D_N}	1.664	1.664	1.664	[263]
$\overline{\delta}_{DC}$	5.3932×10^{-1}	6.3864×10^{-1}	7.3501×10^{-1}	Estimated
δ_D	2.77×10^{-1}	2.77×10^{-1}	2.77×10^{-1}	[264]
$\overline{\delta}_{CT_c}$	1.2269×10^{-2}	9.6574×10^{-3}	8.4017×10^{-3}	Estimated
$\overline{\delta}_{CI_\gamma}$	4.5923×10^{-3}	6.4192×10^{-3}	8.4660×10^{-3}	Estimated
δ_C	3.0078×10^{-4}	1.0390×10^{-3}	2.4530×10^{-4}	Estimated

δ_N	4.5935×10^{-1}	4.8360×10^{-1}	1.1137×10^{-1}	Estimated
δ_{I_γ}	3.327×10	3.327×10	3.327×10	[265]
δ_{μ_1}	4.8748×10^2	4.8748×10^2	4.8748×10^2	[266, 267, 268, 269]
δ_{μ_2}	5.15	5.15	5.15	[270, 271]
δ_H	5.87×10	5.87×10	5.87×10	[272]
$\overline{A}_{M_N}$	7.4055×10^{-1}	7.0151×10^{-1}	7.1382×10^{-1}	Estimated
$\overline{A}_{T_N}$	1.0581×10^2	1.7917	3.1561	Estimated
$\overline{A}_{D_N}$	3.3325	2.3958	2.4867	Estimated
$\alpha_{M_N M}$	3.1701	5.6721×10^{-1}	1.3878	Scaling factor
$\alpha_{T_N T_h}$	1.4396	1.8848×10^{-1}	8.8053×10^{-2}	Scaling factor
$\alpha_{T_N T_r}$	7.4588×10^{-1}	8.2864×10^{-2}	1.0267×10^{-1}	Scaling factor
$\alpha_{T_N T_c}$	4.6531	3.0144×10^{-2}	1.3172×10^{-1}	Scaling factor
$\alpha_{D_N D}$	2.0440	7.9922×10^{-1}	8.1299×10^{-1}	Scaling factor

A.3 Non-Dimensionalization

We obtain the following non-dimensional system:

$$\frac{d\left[\overline{M}_N\right]}{dt} = \overline{A}_{M_N} - \alpha_{M_N M}\left(\overline{\lambda}_{M I_\gamma}\left[\overline{I}_\gamma\right] + \overline{\lambda}_{M \mu_1}\left[\overline{\mu}_1\right]\right)\left[\overline{M}_N\right] - \delta_{M_N}[\overline{M}_N] \tag{A.30}$$

$$\frac{d\left[\overline{M}\right]}{dt} = \left(\overline{\lambda}_{M I_\gamma}\left[\overline{I}_\gamma\right] + \overline{\lambda}_{M \mu_1}\left[\overline{\mu}_1\right]\right)\left[\overline{M}_N\right] - \delta_M\left[\overline{M}\right] \tag{A.31}$$

$$\begin{aligned}\frac{d\left[\overline{T}_N\right]}{dt} = {} & \overline{A}_{T_N} - \alpha_{T_N T_h}\left(\overline{\lambda}_{T_h M}\left[\overline{M}\right] + \overline{\lambda}_{T_h D}\left[\overline{D}\right]\right)\left[\overline{T}_N\right] - \alpha_{T_N T_r}\overline{\lambda}_{T_r \mu_1}\left[\overline{\mu}_1\right]\left[\overline{T}_N\right] \\ & - \alpha_{T_N T_c}\left(\overline{\lambda}_{T_c T_h}\left[\overline{T}_h\right] + \overline{\lambda}_{T_c M}\left[\overline{M}\right] + \overline{\lambda}_{T_c D}\left[\overline{D}\right]\right)\left[\overline{T}_N\right] - \delta_{T_N}\left[\overline{T}_N\right]\end{aligned} \tag{A.32}$$

$$\frac{d\left[\overline{T}_h\right]}{dt} = \left(\overline{\lambda}_{T_h M}\left[\overline{M}\right] + \overline{\lambda}_{T_h D}\left[\overline{D}\right]\right)\left[\overline{T}_N\right] - \left(\overline{\delta}_{T_h T_r}\left[\overline{T}_r\right] + \overline{\delta}_{T_h \mu_1}\left[\overline{\mu}_1\right] + \delta_{T_h}\right)\left[\overline{T}_h\right] \tag{A.33}$$

$$\frac{d\left[\overline{T}_r\right]}{dt} = \left(\overline{\lambda}_{T_r\mu_1}\left[\overline{\mu}_1\right]\right)\left[\overline{T}_N\right] - \delta_{T_r}\left[\overline{T}_r\right] \tag{A.34}$$

$$\frac{d\left[\overline{T}_c\right]}{dt} = \left(\overline{\lambda}_{T_cT_h}\left[\overline{T}_h\right] + \overline{\lambda}_{T_cM}\left[\overline{M}\right] + \overline{\lambda}_{T_cD}\left[\overline{D}\right]\right)\left[\overline{T}_N\right] - \left(\overline{\delta}_{T_cT_r}\left[\overline{T}_r\right] + \overline{\delta}_{T_c\mu_1}\left[\overline{\mu}_1\right] + \delta_{T_c}\right)\left[\overline{T}_c\right] \tag{A.35}$$

$$\frac{d\left[\overline{D}_N\right]}{dt} = \overline{A}_{D_N} - \alpha_{D_ND}\left(\overline{\lambda}_{DC}\left[\overline{C}\right] + \overline{\lambda}_{DH}\left[\overline{H}\right]\right)\left[\overline{D}_N\right] - \delta_{D_N}\left[\overline{D}_N\right] \tag{A.36}$$

$$\frac{d\left[\overline{D}\right]}{dt} = \left(\overline{\lambda}_{DC}\left[\overline{C}\right] + \overline{\lambda}_{DH}\left[\overline{H}\right]\right)\left[\overline{D}_N\right] - \left(\overline{\delta}_{DC}\left[\overline{C}\right] + \delta_D\right)\left[\overline{D}\right] \tag{A.37}$$

$$\frac{d\left[\overline{C}\right]}{dt} = \left(\lambda_C + \overline{\lambda}_{C\mu_1}\left[\overline{\mu}_1\right] + \overline{\lambda}_{C\mu_2}\left[\overline{\mu}_2\right]\right)\left[\overline{C}\right]\left(1 - \frac{\left[\overline{C}\right]}{\overline{C}_0}\right) - \left(\overline{\delta}_{CT_c}\left[\overline{T}_c\right] + \overline{\delta}_{CI_\gamma}\left[\overline{I}_\gamma\right] + \delta_C\right)\left[\overline{C}\right] \tag{A.38}$$

$$\frac{d\left[\overline{N}\right]}{dt} = \overline{\alpha}_{NC}\left(\overline{\delta}_{CT_c}\left[\overline{T}_c\right] + \overline{\delta}_{CI_\gamma}\left[\overline{I}_\gamma\right] + \delta_C\right)\left[\overline{C}\right] - \delta_N\left[\overline{N}\right] \tag{A.39}$$

$$\frac{d\left[\overline{I}_\gamma\right]}{dt} = \overline{\lambda}_{I_\gamma T_h}\left[\overline{T}_h\right] + \overline{\lambda}_{I_\gamma T_c}\left[\overline{T}_c\right] - \delta_{I_\gamma}\left[\overline{I}_\gamma\right] \tag{A.40}$$

$$\frac{d\left[\overline{\mu}_1\right]}{dt} = \overline{\lambda}_{\mu_1T_h}\left[\overline{T}_h\right] + \overline{\lambda}_{\mu_1M}\left[\overline{M}\right] + \overline{\lambda}_{\mu_1C}\left[\overline{C}\right] - \delta_{\mu_1}\left[\overline{\mu}_1\right] \tag{A.41}$$

$$\frac{d\left[\overline{\mu}_2\right]}{dt} = \overline{\lambda}_{\mu_2T_h}\left[\overline{T}_h\right] + \overline{\lambda}_{\mu_2M}\left[\overline{M}\right] + \overline{\lambda}_{\mu_2C}\left[\overline{C}\right] - \delta_{\mu_2}\left[\overline{\mu}_2\right] \tag{A.42}$$

$$\frac{d\left[\overline{H}\right]}{dt} = \overline{\lambda}_{HM}\left[\overline{M}\right] + \overline{\lambda}_{HD}\left[\overline{D}\right] + \overline{\lambda}_{HN}\left[\overline{N}\right] - \delta_H\left[\overline{H}\right] \tag{A.43}$$

The non-dimensional parameters are defined as:

$$\overline{A}_{M_N} = \frac{A_{M_N}}{M_N^\infty}, \quad \overline{A}_{T_N} = \frac{A_{T_N}}{T_N^\infty}, \quad \overline{A}_{D_N} = \frac{A_{D_N}}{D_N^\infty},$$

$$\alpha_{M_NM} = \frac{M^\infty}{M_N^\infty}, \quad \alpha_{T_NT_h} = \frac{T_h^\infty}{T_N^\infty}, \quad \alpha_{T_NT_r} = \frac{T_r^\infty}{T_N^\infty},$$

$$\alpha_{T_NT_c} = \frac{T_c^\infty}{T_N^\infty}, \quad \alpha_{D_ND} = \frac{D^\infty}{D_N^\infty}, \quad \overline{\alpha}_{NC} = \alpha_{NC}\frac{C^\infty}{N^\infty},$$

$$\overline{C}_0 = \frac{C_0}{C^\infty}, \quad \overline{\lambda}_{MI_\gamma} = \frac{\lambda_{MI_\gamma}I_\gamma^\infty M_N^\infty}{M^\infty}, \quad \overline{\lambda}_{M\mu_1} = \frac{\lambda_{M\mu_1}\mu_1^\infty M_N^\infty}{M^\infty},$$

$$\overline{\lambda}_{T_hM} = \frac{\lambda_{T_hM}M^\infty T_N^\infty}{T_h^\infty}, \quad \overline{\lambda}_{T_hD} = \frac{\lambda_{T_hD}D^\infty T_N^\infty}{T_h^\infty}, \quad \overline{\lambda}_{T_r\mu_1} = \frac{\lambda_{T_r\mu_1}\mu_1^\infty T_N^\infty}{T_r^\infty},$$

$$\overline{\lambda}_{T_cT_h} = \frac{\lambda_{T_cT_h}T_h^\infty T_N^\infty}{T_c^\infty}, \quad \overline{\lambda}_{T_cM} = \frac{\lambda_{T_cM}M^\infty T_N^\infty}{T_c^\infty}, \quad \overline{\lambda}_{T_cD} = \frac{\lambda_{T_cD}D^\infty T_N^\infty}{T_c^\infty},$$

$$\overline{\lambda}_{DC} = \frac{\lambda_{DC}C^\infty D_N^\infty}{D^\infty}, \quad \overline{\lambda}_{DH} = \frac{\lambda_{DH}H^\infty D_N^\infty}{D^\infty}, \quad \overline{\lambda}_{C\mu_1} = \lambda_{C\mu_1}\mu_1^\infty,$$

$$
\begin{aligned}
\overline{\lambda}_{C\mu_2} &= \lambda_{C\mu_2}\mu_2^{\infty}, & \overline{\lambda}_{I_\gamma T_h} &= \frac{\lambda_{I_\gamma T_h} T_h^{\infty}}{I_\gamma^{\infty}}, & \overline{\lambda}_{I_\gamma T_c} &= \frac{\lambda_{I_\gamma T_c} T_c^{\infty}}{I_\gamma^{\infty}},\\
\overline{\lambda}_{\mu_1 T_h} &= \frac{\lambda_{\mu_1 T_h} T_h^{\infty}}{\mu_1^{\infty}}, & \overline{\lambda}_{\mu_1 M} &= \frac{\lambda_{\mu_1 M} M^{\infty}}{\mu_1^{\infty}}, & \overline{\lambda}_{\mu_1 C} &= \frac{\lambda_{\mu_1 C} C^{\infty}}{\mu_1^{\infty}},\\
\overline{\lambda}_{\mu_2 T_h} &= \frac{\lambda_{\mu_2 T_h} T_h^{\infty}}{\mu_2^{\infty}}, & \overline{\lambda}_{\mu_2 M} &= \frac{\lambda_{\mu_2 M} M^{\infty}}{\mu_2^{\infty}}, & \overline{\lambda}_{\mu_2 C} &= \frac{\lambda_{\mu_2 C} C^{\infty}}{\mu_2^{\infty}},\\
\overline{\lambda}_{HM} &= \frac{\lambda_{HM} M^{\infty}}{H^{\infty}}, & \overline{\lambda}_{HD} &= \frac{\lambda_{HD} D^{\infty}}{H^{\infty}}, & \overline{\lambda}_{HN} &= \frac{\lambda_{HN} N^{\infty}}{H^{\infty}},\\
\overline{\delta}_{T_h T_r} &= \delta_{T_h T_r} T_r^{\infty}, & \overline{\delta}_{T_h \mu_1} &= \delta_{T_h \mu_1} \mu_1^{\infty}, & \overline{\delta}_{T_c T_r} &= \delta_{T_c T_r} T_r^{\infty},\\
\overline{\delta}_{T_c \mu_1} &= \delta_{T_c \mu_1} \mu_1^{\infty}, & \overline{\delta}_{DC} &= \delta_{DC} C^{\infty}, & \overline{\delta}_{CT_c} &= \delta_{CT_c} T_c^{\infty},\\
\overline{\delta}_{CI_\gamma} &= \delta_{CI_\gamma} I_\gamma^{\infty}.
\end{aligned}
$$

The assumptions (Equations (A.15)–(A.29)) in non-dimensional form are:

$$
\begin{aligned}
\lambda_C &= 20\overline{\lambda}_{C\mu_2}\frac{\mu_2^{\text{mean}}}{\mu_2^{\infty}} = 40\overline{\lambda}_{C\mu_1}\frac{\mu_1^{\text{mean}}}{\mu_1^{\infty}},\\
\overline{\delta}_{CT_c}\frac{T_c^{\text{mean}}}{T_c^{\infty}} &= 2\overline{\delta}_{CI_\gamma}\frac{I_\gamma^{\text{mean}}}{I_\gamma^{\infty}} = 20\delta_C,\\
\frac{\overline{\lambda}_{MI_\gamma}\frac{I_\gamma^{\text{mean}}}{I_\gamma^{\infty}}}{\overline{\lambda}_{M\mu_1}\frac{\mu_1^{\text{mean}}}{\mu_1^{\infty}}} &= \frac{M_1^{\text{mean}}}{M_2^{\text{mean}}},\\
\overline{\lambda}_{T_h D}\frac{D^{\text{mean}}}{D^{\infty}} &= 200\overline{\lambda}_{T_h M}\frac{M^{\text{mean}}}{M^{\infty}},\\
\overline{\delta}_{T_h T_r}\frac{T_r^{\text{mean}}}{T_r^{\infty}} &= \overline{\delta}_{T_h \mu_1}\frac{\mu_1^{\text{mean}}}{\mu_1^{\infty}} = 20\delta_{T_h},\\
\overline{\lambda}_{T_c D}\frac{D^{\text{mean}}}{D^{\infty}} &= 2\overline{\lambda}_{T_c T_h}\frac{T_h^{\text{mean}}}{T_h^{\infty}} = 4\overline{\lambda}_{T_c M}\frac{M^{\text{mean}}}{M^{\infty}},\\
\overline{\delta}_{T_c T_r}\frac{T_r^{\text{mean}}}{T_r^{\infty}} &= \overline{\delta}_{T_c \mu_1}\frac{\mu_1^{\text{mean}}}{\mu_1^{\infty}} = 20\delta_{T_c},\\
\overline{\lambda}_{DH}\frac{H^{\text{mean}}}{H^{\infty}} &= 2\overline{\lambda}_{DC}\frac{C^{\text{mean}}}{C^{\infty}},\\
\overline{\delta}_{DC}\frac{C^{\text{mean}}}{C^{\infty}} &= \delta_D,\\
\overline{\lambda}_{I_\gamma T_c}\frac{T_c^{\text{mean}}}{T_c^{\infty}} &= 4\overline{\lambda}_{I_\gamma T_h}\frac{T_h^{\text{mean}}}{T_h^{\infty}},\\
\overline{\lambda}_{\mu_1 T_h}\frac{T_h^{\text{mean}}}{T_h^{\infty}} &= \overline{\lambda}_{\mu_1 M}\frac{M^{\text{mean}}}{M^{\infty}} = \overline{\lambda}_{\mu_1 C}\frac{C^{\text{mean}}}{C^{\infty}},\\
\overline{\lambda}_{\mu_2 M}\frac{M^{\text{mean}}}{M^{\infty}} &= \overline{\lambda}_{\mu_2 C}\frac{C^{\text{mean}}}{C^{\infty}} = 2\overline{\lambda}_{\mu_2 T_h}\frac{T_h^{\text{mean}}}{T_h^{\infty}},
\end{aligned}
$$

$$\overline{\lambda}_{HN}\frac{N^{\text{mean}}}{N^{\infty}} = \ 10\overline{\lambda}_{HM}\frac{M^{\text{mean}}}{M^{\infty}} = \ 20\overline{\lambda}_{HD}\frac{D^{\text{mean}}}{D^{\infty}}.$$

A P P E N D I X B

SUPPORTING INFORMATION FOR CHAPTER 5

B.1 Full dimensional system of ODEs

Combining Equations (4.1)-(4.4) and (5.1)-(5.13) we obtain the following system of ODEs:

$$\begin{aligned}
\frac{d\left[M_N\right]}{dt} = \; & A_{M_N} - \left(\lambda_{MI_\gamma}\left[I_\gamma\right] + \lambda_{M\mu_1}\left[\mu_1\right]\right)\left[M_N\right] - \delta_{M_N}\left[M_N\right] \\
& - \left(\boldsymbol{K_{M_N}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_{M_N}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right. \\
& \left. + \boldsymbol{K_{M_N}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{M_N}\right] \qquad \text{(B.1)} \\
\frac{d\left[M\right]}{dt} = \; & \left(\lambda_{MI_\gamma}\left[I_\gamma\right] + \lambda_{M\mu_1}\left[\mu_1\right]\right)\left[M_N\right] - \delta_M\left[M\right] - \left(\boldsymbol{K_M}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}}\right)\right. \\
& \left.\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_M}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right) + \boldsymbol{K_M}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{M}\right] \qquad \text{(B.2)} \\
\frac{d\left[T_N\right]}{dt} = \; & A_{T_N} - \left(\lambda_{T_hM}\left[M\right] + \lambda_{T_hD}\left[D\right]\right)\left[T_N\right] - \lambda_{T_r\mu_1}\left[\mu_1\right]\left[T_N\right] \\
& - \left(\lambda_{T_cT_h}\left[T_h\right] + \lambda_{T_cM}\left[M\right] + \lambda_{T_cD}\left[D\right]\right)\left[T_N\right] - \delta_{T_N}\left[T_N\right] \\
& - \left(\boldsymbol{K_{T_N}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_{T_N}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right. \\
& \left. + \boldsymbol{K_{T_N}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{T_N}\right] \qquad \text{(B.3)}
\end{aligned}$$

$$
\begin{aligned}
\frac{d\left[T_h\right]}{dt} = & \left(\lambda_{T_hM}\left[M\right] + \lambda_{T_hD}\left[D\right]\right)\left[T_N\right] - \left(\delta_{T_hT_r}\left[T_r\right] + \delta_{T_h\mu_1}\left[\mu_1\right] + \delta_{T_h}\right)\left[T_h\right] \\
& - \left(\boldsymbol{K_{T_h}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_{T_h}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right. \\
& \left. + \boldsymbol{K_{T_h}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{T_h}\right] \qquad \text{(B.4)} \\
\frac{d\left[T_r\right]}{dt} = & \left(\lambda_{T_r\mu_1}\left[\mu_1\right]\right)\left[T_N\right] - \delta_{T_r}\left[T_r\right] - \left(\boldsymbol{K_{T_r}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right)\right. \\
& \left. + \boldsymbol{K_{T_r}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right) + \boldsymbol{K_{T_r}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{T_r}\right] \qquad \text{(B.5)} \\
\frac{d\left[T_c\right]}{dt} = & \left(\lambda_{T_cT_h}\left[T_h\right] + \lambda_{T_cM}\left[M\right] + \lambda_{T_cD}\left[D\right]\right)\left[T_N\right] - \left(\delta_{T_cT_r}\left[T_r\right] + \delta_{T_c\mu_1}\left[\mu_1\right] + \delta_{T_c}\right)\left[T_c\right] \\
& - \left(\boldsymbol{K_{T_c}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_{T_c}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right. \\
& \left. + \boldsymbol{K_{T_c}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{T_c}\right] \qquad \text{(B.6)} \\
\frac{d\left[D_N\right]}{dt} = & A_{D_N} - \left(\lambda_{DC}\left[C\right] + \lambda_{DH}\left[H\right]\right)\left[D_N\right] - \delta_{D_N}\left[D_N\right] \\
& - \left(\boldsymbol{K_{D_N}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_{D_N}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right. \\
& \left. + \boldsymbol{K_{D_N}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{D_N}\right] \qquad \text{(B.7)} \\
\frac{d\left[D\right]}{dt} = & \left(\lambda_{DC}\left[C\right] + \lambda_{DH}\left[H\right]\right)\left[D_N\right] - \left(\delta_{DC}\left[C\right] + \delta_D\right)\left[D\right] \\
& - \left(\boldsymbol{K_D}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_D}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right. \\
& \left. + \boldsymbol{K_D}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[\boldsymbol{D}\right] \qquad \text{(B.8)} \\
\frac{d\left[C\right]}{dt} = & \left(\lambda_C + \lambda_{C\mu_1}\left[\mu_1\right] + \lambda_{C\mu_2}\left[\mu_2\right]\right)\left[C\right]\left(1 - \frac{\left[C\right]}{C_0}\right) \\
& - \left(\delta_{CI_\gamma}\left[I_\gamma\right] + \delta_C + \delta_{CT_c}\left(1 + \boldsymbol{\delta_{CT_cA_3}}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}}\right)\right)\left[T_c\right]\right)\left[C\right] \\
& - \left(\boldsymbol{K_C}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}}\right)\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}\right) + \boldsymbol{K_C}\left(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}\right)\right.
\end{aligned}
$$

$$+ \boldsymbol{K_C}(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}})\Bigg)[\boldsymbol{C}] \tag{B.9}$$

$$\begin{aligned}\frac{d[N]}{dt} = {} & \alpha_{NC}\Bigg(\delta_{CI_\gamma}[I_\gamma] + \delta_C + \delta_{CT_c}\left(1 + \boldsymbol{\delta_{CT_cA_3}}(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}})\right)[T_c]\Bigg)[C] \\ & + \boldsymbol{\alpha_{NCA}}\Bigg(\boldsymbol{K_C}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_1 A_1}}) + \boldsymbol{K_C}(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_2 A_2}}) \\ & + \boldsymbol{K_C}(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\beta_3 A_3}})\Bigg)[\boldsymbol{C}] - \delta_N[N]\end{aligned} \tag{B.10}$$

$$\frac{d[I_\gamma]}{dt} = \lambda_{I_\gamma T_h}[T_h] + \lambda_{I_\gamma T_c}[T_c] - \delta_{I_\gamma}[I_\gamma] \tag{B.11}$$

$$\frac{d[\mu_1]}{dt} = \lambda_{\mu_1 T_h}[T_h] + \lambda_{\mu_1 M}[M] + \lambda_{\mu_1 C}[C] - \delta_{\mu_1}[\mu_1] \tag{B.12}$$

$$\frac{d[\mu_2]}{dt} = \lambda_{\mu_2 T_h}[T_h] + \lambda_{\mu_2 M}[M] + \lambda_{\mu_2 C}[C] - \delta_{\mu_2}[\mu_2] \tag{B.13}$$

$$\frac{d[H]}{dt} = \lambda_{HM}[M] + \lambda_{HD}[D] + \lambda_{HN}[N] - \delta_H[H] \tag{B.14}$$

$$\frac{\boldsymbol{d[A_1]}}{\boldsymbol{dt}} = \boldsymbol{v_{A_1}(t) - \delta_{A_1}[A_1]} \tag{B.15}$$

$$\frac{\boldsymbol{d[A_2]}}{\boldsymbol{dt}} = \boldsymbol{v_{A_2}(t) - \delta_{A_2}[A_2]} \tag{B.16}$$

$$\frac{\boldsymbol{d[A_3]}}{\boldsymbol{dt}} = \boldsymbol{v_{A_3}(t) - \delta_{A_3}[A_3]} \tag{B.17}$$

B.2 Non-Dimensionalization

We introduce the non-dimensional drug variables as follows:

$$\overline{A}_1 = \frac{A_1\delta_{A_1}}{v^*_{A_1}}, \qquad \overline{A}_2 = \frac{A_2\delta_{A_2}}{v^*_{A_2}}, \qquad \overline{A}_3 = \frac{A_3\delta_{A_3}}{v^*_{A_3}}$$

Non-dimensional system (only equations with changes):

$$\begin{aligned}\frac{d\left[\overline{M}_N\right]}{dt} = {} & \overline{A}_{M_N} - \alpha_{M_N M}\left(\overline{\lambda}_{MI_\gamma}\left[\overline{I}_\gamma\right] + \overline{\lambda}_{M\mu_1}[\overline{\mu}_1]\right)\left[\overline{M}_N\right] - \delta_{M_N}[\overline{M}_N] \\ & - \Bigg(\boldsymbol{K_{M_N}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\overline{\beta}_1 \overline{A}_1}}) + \boldsymbol{K_{M_N}}(\mathbf{1} - \boldsymbol{e}^{-\boldsymbol{\overline{\beta}_2 \overline{A}_2}})\end{aligned}$$

$$+ \boldsymbol{K_{M_N}}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{3}}\overline{\boldsymbol{A}}_{\mathbf{3}}})\Bigg) \left[\overline{\boldsymbol{M}}_{\boldsymbol{N}}\right] \tag{B.18}$$

$$\begin{aligned}\frac{d\left[\overline{M}\right]}{dt} = & \left(\overline{\lambda}_{MI_\gamma}\left[\overline{I}_\gamma\right] + \overline{\lambda}_{M\mu_1}\left[\overline{\mu}_1\right]\right)\left[\overline{M}_N\right] - \delta_M\left[\overline{M}\right] - \Bigg(\boldsymbol{K_M}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right) \\ & (\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{1}}\overline{\boldsymbol{A}}_{\mathbf{1}}}) + \boldsymbol{K_M}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{2}}\overline{\boldsymbol{A}}_{\mathbf{2}}}) + \boldsymbol{K_M}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{3}}\overline{\boldsymbol{A}}_{\mathbf{3}}})\Bigg)\left[\overline{\boldsymbol{M}}\right]\end{aligned} \tag{B.19}$$

$$\begin{aligned}\frac{d\left[\overline{T}_N\right]}{dt} = & \overline{A}_{T_N} - \alpha_{T_NT_h}\left(\overline{\lambda}_{T_hM}\left[\overline{M}\right] + \overline{\lambda}_{T_hD}\left[\overline{D}\right]\right)\left[\overline{T}_N\right] - \alpha_{T_NT_r}\overline{\lambda}_{T_r\mu_1}\left[\overline{\mu}_1\right]\left[\overline{T}_N\right] \\ & - \alpha_{T_NT_c}\left(\overline{\lambda}_{T_cT_h}\left[\overline{T}_h\right] + \overline{\lambda}_{T_cM}\left[\overline{M}\right] + \overline{\lambda}_{T_cD}\left[\overline{D}\right]\right)\left[\overline{T}_N\right] - \delta_{T_N}\left[\overline{T}_N\right] \\ & - \Bigg(\boldsymbol{K_{T_N}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{1}}\overline{\boldsymbol{A}}_{\mathbf{1}}}) + \boldsymbol{K_{T_N}}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{2}}\overline{\boldsymbol{A}}_{\mathbf{2}}}) \\ & + \boldsymbol{K_{T_N}}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{3}}\overline{\boldsymbol{A}}_{\mathbf{3}}})\Bigg)\left[\overline{\boldsymbol{T}}_{\boldsymbol{N}}\right]\end{aligned} \tag{B.20}$$

$$\begin{aligned}\frac{d\left[\overline{T}_h\right]}{dt} = & \left(\overline{\lambda}_{T_hM}\left[\overline{M}\right] + \overline{\lambda}_{T_hD}\left[\overline{D}\right]\right)\left[\overline{T}_N\right] - \left(\overline{\delta}_{T_hT_r}\left[\overline{T}_r\right] + \overline{\delta}_{T_h\mu_1}\left[\overline{\mu}_1\right] + \delta_{T_h}\right)\left[\overline{T}_h\right] \\ & - \Bigg(\boldsymbol{K_{T_h}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{1}}\overline{\boldsymbol{A}}_{\mathbf{1}}}) + \boldsymbol{K_{T_h}}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{2}}\overline{\boldsymbol{A}}_{\mathbf{2}}}) \\ & + \boldsymbol{K_{T_h}}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{3}}\overline{\boldsymbol{A}}_{\mathbf{3}}})\Bigg)\left[\overline{\boldsymbol{T}}_{\boldsymbol{h}}\right]\end{aligned} \tag{B.21}$$

$$\begin{aligned}\frac{d\left[\overline{T}_r\right]}{dt} = & \left(\overline{\lambda}_{T_r\mu_1}\left[\overline{\mu}_1\right]\right)\left[\overline{T}_N\right] - \delta_{T_r}\left[\overline{T}_r\right] - \Bigg(\boldsymbol{K_{T_r}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{1}}\overline{\boldsymbol{A}}_{\mathbf{1}}}) \\ & + \boldsymbol{K_{T_r}}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{2}}\overline{\boldsymbol{A}}_{\mathbf{2}}}) + \boldsymbol{K_{T_r}}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{3}}\overline{\boldsymbol{A}}_{\mathbf{3}}})\Bigg)\left[\overline{\boldsymbol{T}}_{\boldsymbol{r}}\right]\end{aligned} \tag{B.22}$$

$$\begin{aligned}\frac{d\left[\overline{T}_c\right]}{dt} = & \left(\overline{\lambda}_{T_cT_h}\left[\overline{T}_h\right] + \overline{\lambda}_{T_cM}\left[\overline{M}\right] + \overline{\lambda}_{T_cD}\left[\overline{D}\right]\right)\left[\overline{T}_N\right] - (\overline{\delta}_{T_cT_r}\left[\overline{T}_r\right] + \overline{\delta}_{T_c\mu_1}\left[\overline{\mu}_1\right] \\ & + \delta_{T_c})\left[\overline{T}_c\right] - \Bigg(\boldsymbol{K_{T_c}}\left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\mathbf{24}\boldsymbol{a}}\right)(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{1}}\overline{\boldsymbol{A}}_{\mathbf{1}}}) + \boldsymbol{K_{T_c}}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{2}}\overline{\boldsymbol{A}}_{\mathbf{2}}}) \\ & + \boldsymbol{K_{T_c}}(\mathbf{1} - \boldsymbol{e}^{-\overline{\boldsymbol{\beta}}_{\mathbf{3}}\overline{\boldsymbol{A}}_{\mathbf{3}}})\Bigg)\left[\overline{\boldsymbol{T}}_{\boldsymbol{c}}\right]\end{aligned} \tag{B.23}$$

$$\frac{d\left[\overline{D}_N\right]}{dt} = \overline{A}_{D_N} - \alpha_{D_ND}\left(\overline{\lambda}_{DC}\left[\overline{C}\right] + \overline{\lambda}_{DH}\left[\overline{H}\right]\right)\left[\overline{D}_N\right] - \delta_{D_N}\left[\overline{D}_N\right]$$

$$
\begin{aligned}
&- \left(\boldsymbol{K_{D_N}} \left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}} \right) (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_1 \overline{A}_1}) + \boldsymbol{K_{D_N}} (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_2 \overline{A}_2}) \right. \\
&\left. + \boldsymbol{K_{D_N}} (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_3 \overline{A}_3}) \right) \left[\overline{\boldsymbol{D}}_{\boldsymbol{N}} \right] \qquad \text{(B.24)}
\end{aligned}
$$

$$
\begin{aligned}
\frac{d\left[\overline{D}\right]}{dt} = \; & \left(\overline{\lambda}_{DC} \left[\overline{C} \right] + \overline{\lambda}_{DH} \left[\overline{H} \right] \right) \left[\overline{D}_N \right] - \left(\overline{\delta}_{DC} \left[\overline{C} \right] + \delta_D \right) \left[\overline{D} \right] \\
& - \left(\boldsymbol{K_D} \left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}} \right) (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_1 \overline{A}_1}) + \boldsymbol{K_D} (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_2 \overline{A}_2}) \right. \\
& \left. + \boldsymbol{K_D} (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_3 \overline{A}_3}) \right) \left[\overline{\boldsymbol{D}} \right] \qquad \text{(B.25)}
\end{aligned}
$$

$$
\begin{aligned}
\frac{d\left[\overline{C}\right]}{dt} = \; & \left(\lambda_C + \overline{\lambda}_{C\mu_1} \left[\overline{\mu}_1 \right] + \overline{\lambda}_{C\mu_2} \left[\overline{\mu}_2 \right] \right) \left[\overline{C} \right] \left(1 - \frac{\left[\overline{C} \right]}{\overline{C}_0} \right) \\
& - \left(\overline{\delta}_{CI_\gamma} \left[\overline{I}_\gamma \right] + \delta_C + \overline{\delta}_{CT_c} \left(1 + \overline{\boldsymbol{\delta}}_{\boldsymbol{CT_cA_3}} (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_3 \overline{A}_3}) \right) \left[\overline{T}_c \right] \right) \left[\overline{C} \right] \\
& - \left(\boldsymbol{K_C} \left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}} \right) (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_1 \overline{A}_1}) + \boldsymbol{K_C} (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_2 \overline{A}_2}) \right. \\
& \left. + \boldsymbol{K_C} (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_3 \overline{A}_3}) \right) \left[\overline{\boldsymbol{C}} \right] \qquad \text{(B.26)}
\end{aligned}
$$

$$
\begin{aligned}
\frac{d\left[\overline{N}\right]}{dt} = \; & \overline{\alpha}_{NC} \left(\overline{\delta}_{CI_\gamma} \left[\overline{I}_\gamma \right] + \delta_C + \overline{\delta}_{CT_c} \left(1 + \overline{\boldsymbol{\delta}}_{\boldsymbol{CT_cA_3}} (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_3 \overline{A}_3}) \right) \left[\overline{T}_c \right] \right) \left[\overline{C} \right] \\
& + \overline{\boldsymbol{\alpha}}_{\boldsymbol{NCA}} \left(\boldsymbol{K_C} \left(\boldsymbol{f} - \frac{\boldsymbol{\tau}}{\boldsymbol{a}} + \frac{\mathbf{1}}{\boldsymbol{24a}} \right) (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_1 \overline{A}_1}) + \boldsymbol{K_C} (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_2 \overline{A}_2}) \right. \\
& \left. + \boldsymbol{K_C} (\mathbf{1} - \boldsymbol{e}^{-\overline{\beta}_3 \overline{A}_3}) \right) \left[\overline{\boldsymbol{C}} \right] - \delta_N \left[\overline{N} \right] \qquad \text{(B.27)}
\end{aligned}
$$

$$
\frac{\boldsymbol{d[\overline{A}_1]}}{\boldsymbol{dt}} = \boldsymbol{\overline{v}_{A_1}(t) - \delta_{A_1}[\overline{A}_1]} \qquad \text{(B.28)}
$$

$$
\frac{\boldsymbol{d[\overline{A}_2]}}{\boldsymbol{dt}} = \boldsymbol{\overline{v}_{A_2}(t) - \delta_{A_2}[\overline{A}_2]} \qquad \text{(B.29)}
$$

$$
\frac{\boldsymbol{d[\overline{A}_3]}}{\boldsymbol{dt}} = \boldsymbol{\overline{v}_{A_3}(t) - \delta_{A_3}[\overline{A}_3]} \qquad \text{(B.30)}
$$

where the new non-dimensional parameters are:

$$
\overline{\beta}_1 = \frac{\beta_1 v^*_{A_1}}{\delta_{A_1}}, \qquad \overline{\beta}_2 = \frac{\beta_2 v^*_{A_2}}{\delta_{A_2}}, \qquad \overline{\beta}_3 = \frac{\beta_3 v^*_{A_3}}{\delta_{A_3}},
$$

$$\overline{\delta}_{CT_cA_3} = \delta_{CT_cA_3} T_c^{\infty}, \qquad \overline{\alpha}_{NCA} = \alpha_{NCA} \frac{C^{\infty}}{N^{\infty}}, \qquad \overline{v}_{A_1}(t) = \frac{v_{A_1}(t)\delta_{A_1}}{v_{A_1}^*},$$

$$\overline{v}_{A_2}(t) = \frac{v_{A_2}(t)\delta_{A_2}}{v_{A_2}^*}, \qquad \overline{v}_{A_3}(t) = \frac{v_{A_3}(t)\delta_{A_3}}{v_{A_3}^*},$$

BIBLIOGRAPHY

[1] National Cancer Institute, What Is Cancer?, 2021. Available from: https://www.cancer.gov/about-cancer/understanding/what-is-cancer.

[2] Memorial Sloan Kettering Cancer Center, What Is Tumor Heterogeneity?, 2021. Available from: https://www.mskcc.org/news/what-tumor-heterogeneity.

[3] Grivennikov, S., Greten, F. & Karin, M. Immunity, Inflammation, and Cancer. *Cell*. **140**, 883-899 (2010).

[4] Kitamura, T., Qian, B. & Pollard, J. Immune cell promotion of metastasis. *Nature Reviews Immunology*. **15**, 73-86 (2015).

[5] Swann, J. & Smyth, M. Immune surveillance of tumors. *Journal Of Clinical Investigation*. **117**, 1137-1146 (2007).

[6] Woo, S., Corrales, L. & Gajewski, T. Innate Immune Recognition of Cancer. *Annual Review Of Immunology*. **33**, 445-474 (2015).

[7] Banchereau, J. & Steinman, R. Dendritic cells and the control of immunity. *Nature*. **392**, 245-252 (1998).

[8] Kroemer, G., Galluzzi, L., Kepp, O. & Zitvogel, L. Immunogenic Cell Death in Cancer Therapy. *Annual Review Of Immunology*. **31**, 51-72 (2013).

[9] Miwa, S., Shirai, T., Yamamoto, N., Hayashi, K., Takeuchi, A., Igarashi, K. & Tsuchiya, H. Current and Emerging Targets in Immunotherapy for Osteosarcoma. *Journal Of Oncology*. **2019** (2019).

[10] Wang, Z., Wang, Z., Li, B., Wang, S., Chen, T. & Ye, Z. Innate immune cells: A potential and promising cell population for treating osteosarcoma. *Frontiers In Immunology*. **10**, 1-19 (2019).

[11] Castro, F., Cardoso, A., Gonçalves, R., Serre, K. & Oliveira, M. Interferon-gamma at the crossroads of tumor immune surveillance or evasion. *Frontiers In Immunology*. **9** pp. 847 (2018).

[12] Wang, K. & Vella, A. Regulatory T Cells and Cancer: A Two-Sided Story. *Immunological Investigations*. **45**, 797-812 (2016).

[13] Wang, Z., Li, B., Ren, Y. & Ye, Z. T-cell-based immunotherapy for osteosarcoma: Challenges and opportunities. *Frontiers In Immunology*. **7**, 1-13 (2016).

[14] Corthay, A. How do regulatory t cells work?. *Scandinavian Journal Of Immunology*. **70**, 326-336 (2009).

[15] Kansara, M., Teng, M., Smyth, M. & Thomas, D. Translational biology of osteosarcoma. *Nature Reviews Cancer*. **14**, 722-735 (2014).

[16] Lewis, C. & Pollard, J. Distinct role of macrophages in different tumor microenvironments. *Cancer Research*. **66**, 605-612 (2006).

[17] Heymann, M., Lézot, F. & Heymann, D. The contribution of immune infiltrates and the local microenvironment in the pathogenesis of osteosarcoma. *Cellular Immunology*. **343**, 103711 (2019).

[18] Cersosimo, F., Lonardi, S., Bernardini, G., Telfer, B., Mandelli, G., Santucci, A., Vermi, W. & Giurisato, E. Tumor-associated macrophages in osteosarcoma: From mechanisms to therapy. *International Journal Of Molecular Sciences*. **21**, 1-20 (2020).

[19] Zheng, Y., Wang, G., Chen, R., Hua, Y. & Cai, Z. Mesenchymal stem cells in the osteosarcoma microenvironment: Their biological properties, influence on tumor growth, and therapeutic implications. *Stem Cell Research And Therapy*. **9**, 1-9 (2018).

[20] Corre, I., Verrecchia, F., Crenn, V., Redini, F. & Trichet, V. The Osteosarcoma Microenvironment: A Complex But Targetable Ecosystem. *Cells*. **9**, 1-25 (2020).

[21] Petitprez, F., Meylan, M., Reyniès, A., Sautès-Fridman, C. & Fridman, W. The tumor microenvironment in the response to immune checkpoint blockade therapies. *Frontiers In Immunology*. **11** pp. 784 (2020).

[22] Liu, R., Liao, Y., Zhang, W. & Zhou, H. Relevance of Immune Infiltration and clinical outcomes in pancreatic ductal adenocarcinoma subtypes. *Frontiers In Oncology*. **10** (2020).

[23] Palmerini, E., Agostinelli, C., Picci, P., Pileri, S., Marafioti, T., Lollini, P., Scotlandi, K., Longhi, A., Benassi, M. & Ferrari, S. Tumoral immune-infiltrate (IF), PD-L1 expression and role of CD8/TIA-1 lymphocytes in localized osteosarcoma patients treated within protocol ISG-OS1. *Oncotarget*. **8**, 111836 (2017).

[24] Wu, T. & Dai, Y. Tumor microenvironment and therapeutic response. *Cancer Letters*. **387** pp. 61-68 (2017).

[25] Golden, E., Frances, D., Pellicciotta, I., Demaria, S., Helen Barcellos-Hoff, M. & Formenti, S. Radiation fosters dose-dependent and chemotherapy-induced immunogenic cell death. *Oncoimmunology*. **3**, e28518 (2014).

[26] Schildkopf, P., Frey, B., Mantel, F., Ott, O., Weiss, E., Sieber, R., Janko, C., Sauer, R., Fietkau, R. & Gaipl, U. Application of hyperthermia in addition to ionizing irradiation fosters necrotic cell death and HMGB1 release of colorectal tumor cells. *Biochemical And Biophysical Research Communications*. **391**, 1014-1020 (2010).

[27] Apetoh, L., Ghiringhelli, F., Tesniere, A., Criollo, A., Ortiz, C., Lidereau, R., Mariette, C., Chaput, N., Mira, J., Delaloge, S. & Others The interaction between HMGB1 and TLR4 dictates the outcome of anticancer chemotherapy and radiotherapy. *Immunological Reviews*. **220**, 47-59 (2007).

[28] Liu, L., Yang, M., Kang, R., Wang, Z., Zhao, Y., Yu, Y., Xie, M., Yin, X., Livesey, K., Lotze, M. & Others HMGB1-induced autophagy promotes chemotherapy resistance in leukemia cells. *Leukemia*. **25**, 23-31 (2011).

[29] Parker, K., Sinha, P., Horn, L., Clements, V., Yang, H., Li, J., Tracey, K. & Ostrand-Rosenberg, S. HMGB1 enhances immune suppression by facilitating the differentiation and suppressive activity of myeloid-derived suppressor cells. *Cancer Research*. **74**, 5723-5733 (2014).

[30] Dumitriu, I., Baruah, P., Manfredi, A., Bianchi, M. & Rovere-Querini, P. HMGB1: Guiding immunity from within. *Trends In Immunology*. **26**, 381-387 (2005).

[31] Ranzato, E., Martinotti, S. & Patrone, M. Emerging roles for HMGB1 protein in immunity, inflammation, and cancer. *ImmunoTargets And Therapy*. pp. 101 (2015).

[32] Klune, J., Dhupar, R., Cardinal, J., Billiar, T. & Tsung, A. HMGB1: Endogenous danger signaling. *Molecular Medicine*. **14**, 476-484 (2008).

[33] Dagogo-Jack, I. & Shaw, A. Tumour heterogeneity and resistance to cancer therapies. *Nature Reviews Clinical Oncology*. **15**, 81-94 (2018).

[34] Werner, H., Mills, G. & Ram, P. Cancer Systems Biology: a peek into the future of patient care?. *Nature Reviews Clinical Oncology*. **11**, 167-176 (2014).

[35] Shahriyari, L. & Komarova, N. Symmetric vs. asymmetric stem cell divisions: an adaptation against cancer?. *PloS One*. **8**, e76195 (2013).

[36] Shahriyari, L. & Komarova, N. The role of the bi-compartmental stem cell niche in delaying cancer. *Physical Biology*. **12**, 055001 (2015).

[37] Shahriyari, L. & Mahdipour-Shirayeh, A. Modeling dynamics of mutants in heterogeneous stem cell niche. *Physical Biology*. **14**, 016004 (2017).

[38] Bollas, A. & Shahriyari, L. The role of backward cell migration in two-hit mutants' production in the stem cell niche. *Plos One*. **12**, e0184651 (2017).

[39] Brady, R. & Enderling, H. Mathematical models of cancer: when to predict novel therapies, and when not to. *Bulletin Of Mathematical Biology*. **81**, 3722-3731 (2019).

[40] Chamseddine, I. & Rejniak, K. Hybrid modeling frameworks of tumor development and treatment. *Wiley Interdisciplinary Reviews: Systems Biology And Medicine*. **12**, e1461 (2020).

[41] Moreira, J. & Deutsch, A. Cellular automaton models of tumor development: a critical review. *Advances In Complex Systems*. **5**, 247-267 (2002).

[42] Lowengrub, J., Frieboes, H., Jin, F., Chuang, Y., Li, X., Macklin, P., Wise, S. & Cristini, V. Nonlinear modelling of cancer: bridging the gap between cells and tumours. *Nonlinearity*. **23**, R1 (2009).

[43] Shahriyari, L. Cell dynamics in tumour environment after treatments. *Journal Of The Royal Society Interface*. **14**, 20160977 (2017).

[44] Frei, C., Hillen, T. & Rhodes, A. A stochastic model for cancer metastasis: branching stochastic process with settlement. *Mathematical Medicine And Biology: A Journal Of The IMA*. **37**, 153-182 (2020).

[45] Rhodes, A. & Hillen, T. Implications of immune-mediated metastatic growth on metastatic dormancy, blow-up, early detection, and treatment. *Journal Of Mathematical Biology*. **81**, 799-843 (2020).

[46] Hoffman, F., Gavaghan, D., Osborne, J., Barrett, I., You, T., Ghadially, H., Sainson, R., Wilkinson, R. & Byrne, H. A mathematical model of antibody-dependent cellular cytotoxicity (ADCC). *Journal Of Theoretical Biology*. **436** pp. 39-50 (2018).

[47] Mahasa, K., Ouifki, R., Eladdadi, A. & Pillis, L. Mathematical model of tumor–immune surveillance. *Journal Of Theoretical Biology*. **404** pp. 312-330 (2016).

[48] Breems, N. & Eftimie, R. The re-polarisation of M2 and M1 macrophages and its role on cancer outcomes. *Journal Of Theoretical Biology*. **390** pp. 23-39 (2016).

[49] Frascoli, F., Kim, P., Hughes, B. & Landman, K. A dynamical model of tumour immunotherapy. *Mathematical Biosciences*. **253** pp. 50-62 (2014).

[50] Chappell, M., Chelliah, V., Cherkaoui, M., Derks, G., Dumortier, T., Evans, N., Ferrarini, M., Fornari, C., Ghazal, P., Guerriero, M. & Others Mathematical modelling for combinations of immuno-oncology and anti-cancer therapies. *Rep. QSP UK Meet.* pp. 1-15 (2015).

[51] Kaur, G. & Ahmad, N. On study of immune response to tumor cells in prey-predator system. *International Scholarly Research Notices*. **2014** (2014).

[52] Pillis, L., Gallegos, A. & Radunskaya, A. A model of dendritic cell therapy for melanoma. *Frontiers In Oncology*. **3** (2013).

[53] López, Á., Seoane, J. & Sanjuán, M. A validated mathematical model of tumor growth including tumor–host interaction, cell-mediated immune response and chemotherapy. *Bulletin Of Mathematical Biology*. **76**, 2884-2906 (2014).

[54] Wilkie, K. & Hahnfeldt, P. Modeling the dichotomy of the immune response to cancer: cytotoxic effects and tumor-promoting inflammation. *Bulletin Of Mathematical Biology*. **79**, 1426-1448 (2017).

[55] Pillis, L., Caldwell, T., Sarapata, E. & Williams, H. Mathematical modeling of regulatory T cell effects on renal cell carcinoma treatment. *Discrete Continuous Dynamical Systems-B*. **18**, 915 (2013).

[56] Robertson-Tessi, M., El-Kareh, A. & Goriely, A. A model for effects of adaptive immunity on tumor response to chemotherapy and chemoimmunotherapy. *Journal Of Theoretical Biology*. **380** pp. 569-584 (2015).

[57] Robertson-Tessi, M., El-Kareh, A. & Goriely, A. A mathematical model of tumor–immune interactions. *Journal Of Theoretical Biology*. **294** pp. 56-73 (2012).

[58] American Cancer Society, Key Statistics for Osteosarcoma, 2021. Available from: https://www.cancer.org/cancer/osteosarcoma/about/key-statistics.html.

[59] Ottaviani, G. & Jaffe, N. The Epidemiology of Osteosarcoma. *Cancer Treat Res.* . pp. 3-13 (2009).

[60] Yang, Y., Han, L., He, Z., Li, X., Yang, S., Yang, J., Zhang, Y., Li, D., Yang, Y. & Yang, Z. Advances in limb salvage treatment of osteosarcoma. *Journal Of Bone Oncology*. **10** pp. 36-40 (2018).

[61] PDQ Pediatric Treatment Editorial Board Osteosarcoma and Malignant Fibrous Histiocytoma of Bone Treatment (PDQ®): Patient Version. *PDQ Cancer Information Summaries*. (2002), http://www.ncbi.nlm.nih.gov/pubmed/26389380.

[62] American Cancer Society, What Cause Osteosarcoma? 2021. Available from: https://www.cancer.org/cancer/osteosarcoma/causes-risks-prevention/what-causes.html.

[63] Tsukamoto, S., Errani, C., Angelini, A. & Mavrogenis, A. Current Treatment Considerations for Osteosarcoma Metastatic at Presentation. *Orthopedics*. **43** (2020).

[64] Marchandet, L., Lallier, M., Charrier, C., Baud'huin, M., Ory, B. & Lamoureux, F. Mechanisms of Resistance to Conventional Therapies for Osteosarcoma. *Cancers*. **13**, 683 (2021).

[65] He, X., Gao, Z., Xu, H., Zhang, Z. & Fu, P. A meta-analysis of randomized control trials of surgical methods with osteosarcoma outcomes. *Journal Of Orthopaedic Surgery And Research*. **12**, 5 (2017).

[66] Meyers, P., Schwartz, C., Krailo, M., Healey, J., Bernstein, M., Betcher, D., Ferguson, W., Gebhardt, M., Goorin, A., Harris, M., Kleinerman, E., Link, M., Nadel, H., Nieder, M., Siegal, G., Weiner, M., Wells, R., Womer, R. & Grier, H. Osteosarcoma: The Addition of Muramyl Tripeptide to Chemotherapy Improves Overall Survival?A Report From the Children's Oncology Group. *Journal Of Clinical Oncology*. **26**, 633-638 (2008).

[67] Conforti, F., Pala, L., Bagnardi, V., De Pas, T., Martinetti, M., Viale, G., Gelber, R. & Goldhirsch, A. Cancer immunotherapy efficacy and patients' sex: a systematic review and meta-analysis. *The Lancet Oncology*. **19**, 737-746 (2018).

[68] Lee, Y., Tan, Y. & Oon, C. Molecular targeted therapy: Treating cancer with specificity. *European Journal Of Pharmacology*. **834** pp. 188-196 (2018).

[69] Davis, K., Fox, E., Merchant, M., Reid, J., Kudgus, R., Liu, X., Minard, C., Voss, S., Berg, S., Weigel, B. & Mackall, C. Nivolumab in children and young adults with relapsed or refractory solid tumours or lymphoma (ADVL1412): a multicentre, open-label, single-arm, phase 1–2 trial. *The Lancet Oncology*. **21**, 541-550 (2020).

[70] Schwarz, R., Bruland, O., Cassoni, A., Schomberg, P. & Bielack, S. The Role of Radiotherapy in Oseosarcoma. *Cancer Treat Res.* . pp. 147-164 (2009).

[71] Ranjit K Sahu, Alok K Sharma, Surendra Patel, Prakash Kala, Amit Goyal, S. Sternal Mass with Respiratory Compromise in a 10-year-old Child. *Journal Of Bone And Soft Tissue Tumors*. **2**, 2-3 (2019).

[72] Burova, I., Peticone, C., De Silva Thompson, D., Knowles, J., Wall, I. & Shipley, R. A parameterised mathematical model to elucidate osteoblast cell growth in a phosphate-glass microcarrier culture. *Journal Of Tissue Engineering*. **10** (2019).

[73] Fernández-Cervantes, I., Morales, M., Agustín-Serrano, R., Cardenas-García, M., Pérez-Luna, P., Arroyo-Reyes, B. & Maldonado-García, A. Polylactic acid/sodium alginate/hydroxyapatite composite scaffolds with trabecular tissue morphology designed by a bone remodeling model using 3D printing. *Journal Of Materials Science*. pp. 9478-9496 (2019).

[74] Ji, B., Chen, J., Zhen, C., Yang, Q. & Yu, N. Mathematical modelling of the role of Endo180 network in the development of metastatic bone disease in prostate cancer. *Computers In Biology And Medicine*. **117**, 103619 (2020).

[75] Haghiralsadat, F., Amoabediny, G., Naderinezhad, S., Nazmi, K., De Boer, J., Zandieh-Doulabi, B., Forouzanfar, T. & Helder, M. EphA2 Targeted Doxorubicin-Nanoliposomes for Osteosarcoma Treatment. *Pharmaceutical Research*. **34**, 2891-2900 (2017).

[76] Hui, K., Chu, H., Fong, P., Cheng, W. & Lam, T. Population Pharmacokinetic Study and Individual Dose Adjustments of High-Dose Methotrexate in Chinese Pediatric Patients With Acute Lymphoblastic Leukemia or Osteosarcoma. *Journal Of Clinical Pharmacology*. **59**, 566-577 (2019).

[77] Lui, G., Treluyer, J., Fresneau, B., Piperno-Neumann, S., Gaspar, N., Corradini, N., Gentet, J., Marec Berard, P., Laurence, V., Schneider, P., Entz-Werle, N., Pacquement, H., Millot, F., Taque, S., Freycon, C., Lervat, C., Le Deley, M., Mahier Ait Oukhatar, C., Brugieres, L., Le Teuff, G. & Bouazza, N. A Pharmacokinetic and Pharmacogenetic Analysis of Osteosarcoma Patients Treated With High-Dose Methotrexate: Data From the OS2006/Sarcoma-09 Trial. *Journal Of Clinical Pharmacology*. **58**, 1541-1549 (2018).

[78] Tsukahara, T., Kawaguchi, S., Torigoe, T., Asanuma, H., Nakazawa, E., Shimozawa, K., Nabeta, Y., Kimura, S., Kaya, M., Nagoya, S., Wada, T., Yamashita, T. & Sato, N. Prognostic significance of HLA class I expression in osteosarcoma defined by anti-pan HLA class I monoclonal antibody, EMR8-5. *Cancer Science*. **97**, 1374-1380 (2006).

[79] Song, Y., Xu, Y., Zhu, X., Fu, J., Deng, C., Chen, H., Xu, H., Song, G., Lu, J., Tang, Q. & Others Immune Landscape of the Tumor Microenvironment Identifies Prognostic Gene Signature CD4/CD68/CSF1R in Osteosarcoma. *Frontiers In Oncology*. **10** pp. 1198 (2020).

[80] Khader, A., Jia-Wen, T. & Li, J. Construction of immune-related gene pairs signature to predict the overall survival of osteosarcoma patients. *Aging*. **12** (2020).

[81] Biasi, A., Villena-Vargas, J. & Adusumilli, P. Cisplatin-induced antitumor immunomodulation: a review of preclinical and clinical evidence. *Clinical Cancer Research*. **20**, 5384-5391 (2014).

[82] Nejad, E., Sluis, T., Duikeren, S., Yagita, H., Janssen, G., Veelen, P., Melief, C., Burg, S. & Arens, R. Tumor eradication by cisplatin is sustained by CD80/86-mediated costimulation of CD8+ T cells. *Cancer Research*. **76**, 6017-6029 (2016).

[83] Tran, L., Allen, C., Xiao, R., Moore, E., Davis, R., Park, S., Spielbauer, K., Van Waes, C. & Schmitt, N. Cisplatin alters antitumor immunity and synergizes with PD-1/PD-L1 inhibition in head and neck squamous cell carcinoma. *Cancer Immunology Research*. **5**, 1141-1151 (2017).

[84] Merritt, R., Mahtabifard, A., Yamada, R., Crystal, R. & Korst, R. Cisplatin augments cytotoxic T-lymphocyte–mediated antitumor immunity in poorly immunogenic murine lung cancer. *The Journal Of Thoracic And Cardiovascular Surgery*. **126**, 1609-1617 (2003).

[85] Zhang, C., Zheng, J., Lin, Z., Lv, H., Ye, Z., Chen, Y. & Zhang, X. Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of osteosarcoma. *Aging (Albany NY)*. **12**, 3486 (2020).

[86] Tang, Y., Gu, Z., Fu, Y. & Wang, J. CXCR3 from chemokine receptor family correlates with immune infiltration and predicts poor survival in osteosarcoma. *Bioscience Reports*. **39** (2019).

[87] Heath, J., Ribas, A. & Mischel, P. Single-cell analysis tools for drug discovery and development. *Nature Reviews Drug Discovery*. **15**, 204-216 (2016).

[88] Kukurba, K. & Montgomery, S. RNA sequencing and analysis. *Cold Spring Harbor Protocols*. **2015**, pdb-top084970 (2015).

[89] Newman, A., Steen, C., Liu, C., Gentles, A., Chaudhuri, A., Scherer, F., Khodadoust, M., Esfahani, M., Luca, B., Steiner, D., Diehn, M. & Alizadeh, A. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nature Biotechnology*. **37**, 773-782 (2019).

[90] Le, T., Aronow, R., Kirshtein, A. & Shahriyari, L. A review of digital cytometry methods: Estimating the relative abundance of cell types in a bulk of cells. *Briefings In Bioinformatics*. (2020)

[91] Le, T., Su, S. & Shahriyari, L. Immune classification of osteosarcoma. *Mathematical Biosciences And Engineering: MBE*. **18**, 1879 (2021).

[92] Le, T., Su, S., Kirshtein, A. & Shahriyari, L. Data-Driven Mathematical Model of Osteosarcoma. *Cancers*. **13**, 2367 (2021).

[93] Le, T., Su, S. & Shahriyari, L. Investigating Optimal Chemotherapy Options for Osteosarcoma Patients through a Mathematical Model. *Cells*. **10**, 2009 (2021).

[94] Gong, T., Hartmann, N., Kohane, I., Brinkmann, V., Staedtler, F., Letzkus, M., Bongiovanni, S. & Szustakowski, J. Optimal deconvolution of transcriptional profiling data using quadratic programming with application to complex clinical blood samples. *PLoS ONE*. **6** (2011).

[95] Newman, A., Liu, C., Green, M., Gentles, A., Feng, W., Xu, Y., Hoang, C., Diehn, M. & Alizadeh, A. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods*. **12**, 453-457 (2015).

[96] Zyla, J., Marczyk, M., Domaszewska, T., Kaufmann, S., Polanska, J., Weiner, J. & Wren, J. Gene set enrichment for reproducible science: Comparison of CERNO and eight other algorithms. *Bioinformatics*. **35**, 5146-5154 (2019).

[97] Şenbabaoğlu, Y., Gejman, R., Winer, A., Liu, M., Van Allen, E., Velasco, G., Miao, D., Ostrovnaya, I., Drill, E., Luna, A., Weinhold, N., Lee, W., Manley, B., Khalil, D., Kaffenberger, S., Chen, Y., Danilova, L., Voss, M., Coleman, J., Russo, P., Reuter, V., Chan, T., Cheng, E., Scheinberg, D., Li, M., Choueiri, T., Hsieh, J., Sander, C. & Hakimi, A. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures.. *Genome Biology*. **17**, 231 (2016).

[98] Foroutan, M., Bhuva, D., Lyu, R., Horan, K., Cursons, J. & Davis, M. Single sample scoring of molecular phenotypes. *BMC Bioinformatics*. **19**, 404 (2018).

[99] Lawson, C. & Hanson, R. Solving Least Squares Problems. (Society for Industrial, 1995).

[100] Bertsekas, D. Nonlinear Programming. (Athena Scientific, 1999).

[101] Mackey, M., Mackey, D., Higgins, H. & Wright, S. CHEMTAX - A program for estimating class abundances from chemical markers: Application to HPLC measurements of phytoplankton. *Marine Ecology Progress Series*. **144**, 265-283 (1996).

[102] Gustafsson, J., Held, F., Robinson, J., Björnson, E., Jörnsten, R. & Nielsen, J. Sources of variation in cell-type RNA-Seq profiles. *PloS One*. **15**, e0239495 (2020).

[103] Johnson, W., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. **8**, 118-127 (2007).

[104] Barbie, D., Tamayo, P., Boehm, J., Kim, S., Moody, S., Dunn, I., Schinzel, A., Sandy, P., Meylan, E., Scholl, C., Fröhling, S., Chan, E., Sos, M., Michel, K., Mermel, C., Silver, S., Weir, B., Reiling, J., Sheng, Q., Gupta, P., Wadlow, R., Le, H., Hoersch, S., Wittner, B., Ramaswamy, S., Livingston, D., Sabatini, D., Meyerson, M., Thomas, R., Lander, E., Mesirov, J., Root, D., Gilliland, D., Jacks, T. & Hahn, W. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*. **462**, 108-112 (2009).

[105] Subramanian, A., Tamayo, P., Mootha, V., Mukherjee, S., Ebert, B., Gillette, M., Paulovich, A., Pomeroy, S., Golub, T., Lander, E. & Mesirov, J. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings Of The National Academy Of Sciences Of The United States Of America*. **102**, 15545-15550 (2005).

[106] Chen, B., Khodadoust, M., Liu, C., Newman, A. & Alizadeh, A. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. (2018).

[107] Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A., Angell, H., Fredriksen, T., Lafontaine, L., Berger, A., Bruneval, P., Fridman, W., Becker, C., Pagès, F., Speicher, M., Trajanoski, Z. & Galon, J. Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity*. **39**, 782-795 (2013).

[108] Aran, D., Hu, Z. & Butte, A. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biology*. **18**, 1-14 (2017).

[109] Ali, H., Chlon, L., Pharoah, P., Markowetz, F. & Caldas, C. Patterns of immune infiltration in breast cancer and their clinical implications: a gene-expression-based retrospective study. *PLoS Medicine*. **13**, e1002194 (2016).

[110] Zeng, D., Li, M., Zhou, R., Zhang, J., Sun, H., Shi, M., Bin, J., Liao, Y., Rao, J. & Liao, W. Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunology Research*. **7**, 737-750 (2019).

[111] Gentles, A., Newman, A., Liu, C., Bratman, S., Feng, W., Kim, D., Nair, V., Xu, Y., Khuong, A., Hoang, C. & Others The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nature Medicine*. **21**, 938-945 (2015).

[112] Zhou, R., Zhang, J., Zeng, D., Sun, H., Rong, X., Shi, M., Bin, J., Liao, Y. & Liao, W. Immune cell infiltration as a biomarker for the diagnosis and

prognosis of stage I–III colon cancer. *Cancer Immunology, Immunotherapy*. **68**, 433-442 (2019).

[113] Zeng, D., Zhou, R., Yu, Y., Luo, Y., Zhang, J., Sun, H., Bin, J., Liao, Y., Rao, J., Zhang, Y. & Others Gene expression profiles for a prognostic immunoscore in gastric cancer. *The British Journal Of Surgery*. **105**, 1338 (2018).

[114] Jia, Q., Wu, W., Wang, Y., Alexander, P., Sun, C., Gong, Z., Cheng, J., Sun, H., Guan, Y., Xia, X. & Others Local mutational diversity drives intratumoral immune heterogeneity in non-small cell lung cancer. *Nature Communications*. **9**, 1-10 (2018).

[115] Wang, S., Zhang, Q., Yu, C., Cao, Y., Zuo, Y. & Yang, L. Immune cell infiltration-based signature for prognosis and immunogenomic analysis in breast cancer. *Briefings In Bioinformatics*. (2020).

[116] Yoshihara, K., Shahmoradgoli, M., Martınez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P., Levine, D. & Others Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications*. **4**, 1-11 (2013).

[117] Tian, X., Zhu, X., Yan, T., Yu, C., Shen, C., Hong, J., Chen, H. & Fang, J. Differentially expressed lncRNAs in gastric cancer patients: a potential biomarker for gastric cancer prognosis. *Journal Of Cancer*. **8**, 2575 (2017).

[118] Newman, A., Nakao, A., Li, K., Liu, C., Mathi, K., Sigal, N., Maecker, H., Diehn, M. & Alizadeh, A. Analytical validation of digital cytometry (iSort) for leukocyte enumeration using stored blood.. (American Society of Clinical Oncology, 2020).

[119] Newman, A., Nakao, A., Li, K., Wilson, D., Liu, C., Diehn, M. & Alizadeh, A. Analytical validation of iSort digital cytometry for leukocyte enumeration in clinical tumor specimens.. (American Society of Clinical Oncology, 2020).

[120] Hong, W., Yuan, H., Gu, Y., Liu, M., Ji, Y., Huang, Z., Yang, J. & Ma, L. Immune-related prognosis biomarkers associated with osteosarcoma microenvironment. *Cancer Cell International*. **20**, 1-12 (2020).

[121] Yu, Y., Zhang, H., Ren, T., Huang, Y., Liang, X., Wang, W., Niu, J., Han, Y. & Guo, W. Development of a prognostic gene signature based on an immunogenomic infiltration analysis of osteosarcoma. *Journal Of Cellular And Molecular Medicine*. **24**, 11230-11242 (2020).

[122] Hu, C., Liu, C., Tian, S., Wang, Y., Shen, R., Rao, H., Li, J., Yang, X., Chen, B. & Ye, L. Comprehensive analysis of prognostic tumor microenvironment-related genes in osteosarcoma patients. *BMC Cancer*. **20**, 1-11 (2020).

[123] Niu, J., Yan, T., Guo, W., Wang, W., Zhao, Z., Ren, T., Huang, Y., Zhang, H., Yu, Y. & Liang, X. Identification of Potential Therapeutic Targets and Immune Cell Infiltration Characteristics in Osteosarcoma Using Bioinformatics Strategy. *Frontiers In Oncology*. **10** pp. 1628 (2020).

[124] Zhang, T., Nie, Y., Xia, H., Zhang, Y., Cai, K., Chen, X., Li, H. & Wang, J. Identification of Immune-Related Prognostic Genes and LncRNAs Biomarkers Associated With Osteosarcoma Microenvironment. *Frontiers In Oncology*. **10** pp. 1109 (2020).

[125] Yuan, W., Deng, Y., Ren, E., Zhang, G., Wu, Z. & Xie, Q. Analysis of Immune Infiltration Pattern in Osteosarcoma and Its Clinical Significance. *Research Square* . (2020).

[126] Chen, T. & Zhao, L. Patrolling monocytes inhibit osteosarcoma metastasis to the lung. *Aging*. **12** (2020).

[127] Deng, C., Xu, Y., Fu, J., Zhu, X., Chen, H., Xu, H., Wang, G., Song, Y., Song, G., Lu, J. & Others Reprograming the tumor immunologic microenvironment using neoadjuvant chemotherapy in osteosarcoma. *Cancer Science*. **111**, 1899 (2020).

[128] Yang, X., Zhang, W. & Xu, P. NK cell and macrophages confer prognosis and reflect immune status in osteosarcoma. *Journal Of Cellular Biochemistry*. **120**, 8792-8797 (2019).

[129] Wu, C., Beird, H., Livingston, J., Advani, S., Mitra, A., Cao, S., Reuben, A., Ingram, D., Wang, W., Ju, Z. & Others Immuno-genomic landscape of osteosarcoma. *Nature Communications*. **11**, 1-11 (2020).

[130] Yoshihara, K., Shahmoradgoli, M., Martınez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P., Levine, D. & Others Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications*. **4**, 1-11 (2013).

[131] Qiao, G., Miao, H., Yi, Y., Wang, D., Liu, B., Zhang, Y., Chen, X. & Yin, J. Genetic association between CTLA-4 variations and osteosarcoma risk: Case-control study. *International Journal Of Clinical And Experimental Medicine*. **9**, 9598-9602 (2016).

[132] Zhang, C., Hou, W., Ding, X., Wang, X., Zhao, H., Han, X. & Wang, W. Association of cytotoxic T-lymphocyte antigen-4 polymorphisms with malignant bone tumors risk: A meta-analysis. *Asian Pacific Journal Of Cancer Prevention*. **17**, 3783-3789 (2016).

[133] Schroder, K., Hertzog, P., Ravasi, T. & Hume, D. Interferon-: an overview of signals, mechanisms and functions. *Journal Of Leukocyte Biology*. **75**, 163-189 (2004).

[134] Wajant, H. The role of TNF in cancer. *Death Receptors And Cognate Ligands In Cancer*. pp. 1-15 (2009).

[135] National Center for Biotechnology Information, IL1B interleukin 1 beta, 2021. Available from: https://www.ncbi.nlm.nih.gov/gene/3553.

[136] Li, Y., Liu, Q., He, H. & Luo, W. The possible role of insulin-like growth factor-1 in osteosarcoma. *Current Problems In Cancer*. **43**, 228-235 (2019).
https://linkinghub.elsevier.com/retrieve/pii/S0147027218302289

[137] Jentzsch, T., Robl, B., Husmann, M., Bode-Lesniewska, B. & Fuchs, B. Worse prognosis of osteosarcoma patients expressing IGF-1 on a tissue microarray. *Anticancer Research*. **34**, 3881-3890 (2014).

[138] Martin, J., Zielenska, M., Stein, G., Wijnen, A. & Squire, J. The Role of RUNX2 in Osteosarcoma Oncogenesis. *Sarcoma*. **2011** pp. 1-13 (2011).

[139] Roos, A., Satterfield, L., Zhao, S., Fuja, D., Shuck, R., Hicks, M., Donehower, L. & Yustein, J. Loss of Runx2 sensitises osteosarcoma to chemotherapy-induced apoptosis. *British Journal Of Cancer*. **113**, 1289-1297 (2015).

[140] Heymann, M. & Heymann, D. Immune environment and osteosarcoma. *Osteosarcoma-Biology, Behavior And Mechanisms; InTech: London, UK*. pp. 105-120 (2017).

[141] Maciel, T., Moura, I. & Hermine, O. The role of mast cells in cancers. *F1000Prime Reports*. **7**, 5-10 (2015).

[142] Zhao, Y., Niu, C. & Cui, J. Gamma-delta () T Cells: Friend or Foe in Cancer Development. *Journal Of Translational Medicine*. **16**, 1-13 (2018).

[143] Lamora, A., Talbot, J., Mullard, M., Brounais-Le Royer, B., Redini, F. & Verrecchia, F. TGF- Signaling in Bone Remodeling and Osteosarcoma Progression. *Journal Of Clinical Medicine*. **5**, 96 (2016).

[144] Chen, D. & Mellman, I. Oncology Meets Immunology: The Cancer-Immunity Cycle. *Immunity*. **39**, 1-10 (2013).

[145] Motzer, R., Escudier, B., McDermott, D., George, S., Hammers, H., Srinivas, S., Tykodi, S., Sosman, J., Procopio, G., Plimack, E., Castellano, D., Choueiri, T., Gurney, H., Donskov, F., Bono, P., Wagstaff, J., Gauler, T., Ueda, T., Tomita, Y., Schutz, F., Kollmannsberger, C., Larkin, J., Ravaud, A., Simon, J., Xu, L., Waxman, I. & Sharma, P. Nivolumab versus Everolimus

in Advanced Renal-Cell Carcinoma. *New England Journal Of Medicine*. **373**, 1803-1813 (2015).

[146] Dine, J., Gordon, R., Shames, Y., Kasler, M. & Barton-Burke, M. Immune checkpoint inhibitors: An innovation in immunotherapy for the treatment and management of patients with cancer. *Asia-Pacific Journal Of Oncology Nursing*. **4**, 127 (2017).

[147] Topalian, S., Hodi, F., Brahmer, J., Gettinger, S., Smith, D., McDermott, D., Powderly, J., Carvajal, R., Sosman, J., Atkins, M., Leming, P., Spigel, D., Antonia, S., Horn, L., Drake, C., Pardoll, D., Chen, L., Sharfman, W., Anders, R., Taube, J., McMiller, T., Xu, H., Korman, A., Jure-Kunkel, M., Agrawal, S., McDonald, D., Kollia, G., Gupta, A., Wigginton, J. & Sznol, M. Safety, Activity, and Immune Correlates of Anti?PD-1 Antibody in Cancer. *New England Journal Of Medicine*. **366**, 2443-2454 (2012).

[148] Koyama, S., Akbay, E., Li, Y., Herter-Sprie, G., Buczkowski, K., Richards, W., Gandhi, L., Redig, A., Rodig, S., Asahina, H., Jones, R., Kulkarni, M., Kuraguchi, M., Palakurthi, S., Fecci, P., Johnson, B., Janne, P., Engelman, J., Gangadharan, S., Costa, D., Freeman, G., Bueno, R., Hodi, F., Dranoff, G., Wong, K. & Hammerman, P. Adaptive resistance to therapeutic PD-1 blockade is associated with upregulation of alternative immune checkpoints. *Nature Communications*. **7**, 10501 (2016).

[149] Hodi, F., O'Day, S., McDermott, D., Weber, R., Sosman, J., Haanen, J., Gonzalez, R., Robert, C., Schadendorf, D., Hassel, J., Akerley, W., Eertwegh, A., Lutzky, J., Lorigan, P., Vaubel, J., Linette, G., Hogg, D., Ottensmeier, C., Lebbé, C., Peschel, C., Quirt, I., Clark, J., Wolchok, J., Weber, J., Tian, J., Yellin, M., Nichol, G., Hoos, A. & Urba, W. Improved Survival with Ipilimumab in Patients with Metastatic Melanoma. *New England Journal Of Medicine*. **363**, 711-723 (2010).

[150] Tawbi, H., Burgess, M., Bolejack, V., Van Tine, B., Schuetze, S., Hu, J., D'Angelo, S., Attia, S., Riedel, R., Priebat, D., Movva, S., Davis, L., Okuno, S., Reed, D., Crowley, J., Butterfield, L., Salazar, R., Rodriguez-Canales, J., Lazar, A., Wistuba, I., Baker, L., Maki, R., Reinke, D. & Patel, S. Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARC028): a multicentre, two-cohort, single-arm, open-label, phase 2 trial. *The Lancet Oncology*. **18**, 1493-1501 (2017).

[151] Thanindratarn, P., Dean, D., Nelson, S., Hornicek, F. & Duan, Z. Advances in immune checkpoint inhibitors for bone sarcoma therapy. *Journal Of Bone Oncology*. **15** pp. 100221 (2019).

[152] Sercan, Ö., Hämmerling, G., Arnold, B. & Schüler, T. Cutting Edge: Innate Immune Cells Contribute to the IFN--Dependent Regulation of Antigen-

Specific CD8 + T Cell Homeostasis. *The Journal Of Immunology.* **176**, 735-739 (2006).

[153] Lascelles, B., Dernell, W., Correa, M., Lafferty, M., Devitt, C., Kuntz, C., Straw, R. & Withrow, S. Improved Survival Associated With Postoperative Wound Infection in Dogs Treated With Limb-Salvage Surgery for Osteosarcoma. *Annals Of Surgical Oncology.* **12**, 1073-1083 (2005).

[154] Chen, Y., Xu, S., Xu, M. & Yu, X. Postoperative infection and survival in osteosarcoma patients: Reconsideration of immunotherapy for osteosarcoma. *Molecular And Clinical Oncology.* **3**, 495-500 (2015).

[155] Karbach, J., Neumann, A., Brand, K., Wahle, C., Siegel, E., Maeurer, M., Ritter, E., Tsuji, T., Gnjatic, S., Old, L., Ritter, G. & Jäger, E. Phase I clinical trial of mixed bacterial vaccine (Coley's toxins) in patients with NY-ESO-1 expressing cancers: Immunological effects and clinical activity. *Clinical Cancer Research.* **18**, 5449-5459 (2012).

[156] Ling, Z., Fan, G., Yao, D., Zhao, J., Zhou, Y., Feng, J., Zhou, G. & Chen, Y. MicroRNA?150 functions as a tumor suppressor and sensitizes osteosarcoma to doxorubicin?induced apoptosis by targeting RUNX2. *Experimental And Therapeutic Medicine.* (2019).

[157] Zhou, Q., Xian, M., Xiang, S., Xiang, D., Shao, X., Wang, J., Cao, J., Yang, X., Yang, B., Ying, M. & Others All-trans retinoic acid prevents osteosarcoma metastasis by inhibiting M2 polarization of tumor-associated macrophages. *Cancer Immunology Research.* **5**, 547-559 (2017).

[158] Kimura, Y. & Sumiyoshi, M. Resveratrol Prevents Tumor Growth and Metastasis by Inhibiting Lymphangiogenesis and M2 Macrophage Activation and Differentiation in Tumor-associated Macrophages. *Nutrition And Cancer.* **68**, 667-678 (2016).

[159] Kimura, Y. & Sumiyoshi, M. Antitumor and antimetastatic actions of dihydroxycoumarins (esculetin or fraxetin) through the inhibition of M2 macrophage differentiation in tumor-associated macrophages and/or G1 arrest in tumor cells. *European Journal Of Pharmacology.* **746** pp. 115-125 (2015).

[160] Byrne, H., Cox, S. & Kelly, C. Macrophage-tumour interactions: in vivo dynamics. *Discrete Continuous Dynamical Systems-B.* **4**, 81 (2004).

[161] Pillis, L., Renee Fister, K., Gu, W., Collins, C., Daub, M., Gross, D., Moore, J. & Preskill, B. Mathematical model creation for cancer chemo-immunotherapy. *Computational And Mathematical Methods In Medicine.* **10**, 165-184 (2009).

[162] Kelleher, F. & O'Sullivan, H. Monocytes, Macrophages, and Osteoclasts in Osteosarcoma. *Journal Of Adolescent And Young Adult Oncology*. **6**, 396-405 (2017).

[163] Aras, S. & Zaidi, M. TAMeless traitors: macrophages in cancer progression and metastasis. *British Journal Of Cancer*. **117**, 1583-1591 (2017).

[164] Boyman, O. & Sprent, J. The role of interleukin-2 during homeostasis and activation of the immune system. *Nature Reviews Immunology*. **12**, 180-190 (2012).

[165] Fisher, D., Appenheimer, M. & Evans, S. The two faces of IL-6 in the tumor microenvironment. *Seminars In Immunology*. **26**, 38-47 (2014).

[166] Whelan, J., Patterson, D., Perisoglou, M., Bielack, S., Marina, N., Smeland, S. & Bernstein, M. The role of interferons in the treatment of osteosarcoma. *Pediatric Blood Cancer*. **54**, 350-354 (2010).

[167] Rovere-Querini, P., Capobianco, A., Scaffidi, P., Valentinis, B., Catalanotti, F., Giazzon, M., Dumitriu, I., Müller, S., Iannacone, M., Traversari, C., Bianchi, M. & Manfredi, A. HMGB1 is an endogenous immune adjuvant released by necrotic cells. *EMBO Reports*. **5**, 825-830 (2004).

[168] Yang, J., Ma, Z., Wang, Y., Wang, Z., Tian, Y., Du, Y., Bian, W., Duan, Y. & Liu, J. Necrosis of osteosarcoma cells induces the production and release of high-mobility group box 1 protein. *Experimental And Therapeutic Medicine*. **15**, 461-466 (2018).

[169] Kang, R., Zhang, Q., Zeh, H., Lotze, M. & Tang, D. HMGB1 in cancer: Good, bad, or both?. *Clinical Cancer Research*. **19**, 4046-4057 (2013).

[170] Ma, Y., Shurin, G., Peiyuan, Z. & Shurin, M. Dendritic cells in the cancer microenvironment. *Journal Of Cancer*. **4**, 36-44 (2013).

[171] Pahl, J., Kwappenberg, K., Varypataki, E., Santos, S., Kuijjer, M., Mohamed, S., Wijnen, J., Tol, M., Cleton-Jansen, A., Egeler, R. & Others Macrophages inhibit human osteosarcoma cell growth after activation with the bacterial cell wall derivative liposomal muramyl tripeptide in combination with interferon-. *Journal Of Experimental Clinical Cancer Research*. **33**, 1-13 (2014).

[172] Jacobson, N., Szabo, S., Weber-Nordt, R., Zhong, Z., Schreiber, R., Darnell Jr, J. & Murphy, K. Interleukin 12 signaling in T helper type 1 (Th1) cells involves tyrosine phosphorylation of signal transducer and activator of transcription (Stat) 3 and Stat4.. *The Journal Of Experimental Medicine*. **181**, 1755-1762 (1995).

[173] Couper, K., Blount, D. & Riley, E. IL-10: the master regulator of immunity to infection. *The Journal Of Immunology*. **180**, 5771-5777 (2008).

[174] Oh, S. & Li, M. TGF-: guardian of T cell function. *The Journal Of Immunology*. **191**, 3973-3979 (2013).

[175] Lafont, V., Sanchez, F., Laprevotte, E., Michaud, H., Gros, L., Eliaou, J. & Bonnefoy, N. Plasticity of T cells: Impact on the anti-tumor response. *Frontiers In Immunology*. **5**, 1-13 (2014).

[176] Henry, C., Ornelles, D., Mitchell, L., Brzoza-Lewis, K. & Hiltbold, E. IL-12 produced by dendritic cells augments CD8+ T cell activation through the production of the chemokines CCL1 and CCL17. *The Journal Of Immunology*. **181**, 8576-8584 (2008).

[177] Li, L., Jay, S., Wang, Y., Wu, S. & Xiao, Z. IL-12 stimulates CTLs to secrete exosomes capable of activating bystander CD8+ T cells. *Scientific Reports*. **7**, 1-10 (2017).

[178] Dyson, K., Stover, B., Grippin, A., Mendez-Gomez, H., Lagmay, J., Mitchell, D. & Sayour, E. Emerging trends in immunotherapy for pediatric sarcomas. *Journal Of Hematology And Oncology*. **12**, 1-10 (2019).

[179] Tsukumo, S. & Yasutomo, K. Regulation of CD8+ T cells and antitumor immunity by Notch signaling. *Frontiers In Immunology*. **9** pp. 101 (2018).

[180] Durgeau, A., Virk, Y., Corgnac, S. & Mami-Chouaib, F. Recent advances in targeting CD8 T-cell immunity for more effective cancer immunotherapy. *Frontiers In Immunology*. **9** pp. 14 (2018).

[181] Folkman, J. & Hochberg, M. Self-regulation of growth in three dimensions. *The Journal Of Experimental Medicine*. **138**, 745-753 (1973).

[182] Enderling, H., Sunassee, E. & Caudell, J. Predicting patient-specific radiotherapy responses in head and neck cancer to personalize radiation dose fractionation. *BioRxiv*. pp. 630806 (2019).

[183] Kasalak, Ö., Overbosch, J., Glaudemans, A., Boellaard, R., Jutte, P. & Kwee, T. Primary tumor volume measurements in Ewing sarcoma: MRI inter-and intraobserver variability and comparison with FDG-PET. *Acta Oncologica*. **57**, 534-540 (2018).

[184] Grimer, R. Size matters for sarcomas!. *The Annals Of The Royal College Of Surgeons Of England*. **88**, 519-524 (2006).

[185] Qiu, Z., Cui, Y. & Wang, X. Natural bone tissue and its biomimetic. *Mineralized Collagen Bone Graft Substitutes*. pp. 1-22 (2019).

[186] Hao, W., Crouser, E. & Friedman, A. Mathematical model of sarcoidosis. *Proceedings Of The National Academy Of Sciences*. **111**, 16065-16070 (2014).

[187] Hao, W. & Friedman, A. Mathematical model on Alzheimer's disease. *BMC Systems Biology*. **10**, 108 (2016).

[188] Virtanen, P., Gommers, R., Oliphant, T., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J. & Others SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*. **17**, 261-272 (2020).

[189] Zi, Z. Sensitivity analysis approaches applied to systems biology models. *IET Systems Biology*. **5**, 336-346 (2011).

[190] Heiss, F. & Winschel, V. Likelihood approximation by numerical integration on sparse grids. *Journal Of Econometrics*. **144**, 62-80 (2008).

[191] Gerstner, T. & Griebel, M. Numerical integration using sparse grids. *Numerical Algorithms*. **18**, 209-232 (1998).

[192] Fu, C. & Jiang, A. Dendritic cells and CD8 T cell immunity in tumor microenvironment. *Frontiers In Immunology*. **9** pp. 3059 (2018).

[193] Kim, R. Cancer immunoediting: from immune surveillance to immune escape. *Cancer Immunotherapy*. pp. 9-27 (2007).

[194] Heymann, M. & Heymann, D. Immune environment and osteosarcoma. *Osteosarcoma-Biology, Behavior And Mechanisms; InTech: London, UK*. pp. 105-120 (2017).

[195] Tawbi, H., Burgess, M., Bolejack, V., Van Tine, B., Schuetze, S., Hu, J., D'Angelo, S., Attia, S., Riedel, R., Priebat, D. & Others Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARC028): a multicentre, two-cohort, single-arm, open-label, phase 2 trial. *The Lancet Oncology*. **18**, 1493-1501 (2017).

[196] Thanindratarn, P., Dean, D., Nelson, S., Hornicek, F. & Duan, Z. Advances in immune checkpoint inhibitors for bone sarcoma therapy. *Journal Of Bone Oncology*. **15** pp. 100221 (2019).

[197] Fritzsching, B., Fellenberg, J., Moskovszky, L., Sápi, Z., Krenacs, T., Machado, I., Poeschl, J., Lehner, B., Szendroi, M., Bosch, A. & Others CD8+/FOXP3+-ratio in osteosarcoma microenvironment separates survivors from non-survivors: a multicenter validated retrospective study. *Oncoimmunology*. **4**, e990800 (2015).

[198] Asano, Y., Kashiwagi, S., Goto, W., Kurata, K., Noda, S., Takashima, T., Onoda, N., Tanaka, S., Ohsawa, M. & Hirakawa, K. Tumour-infiltrating CD8 to FOXP3 lymphocyte ratio in predicting treatment responses to neoadjuvant chemotherapy of aggressive breast cancer. *British Journal Of Surgery*. **103**, 845-854 (2016).

[199] Yasuda, K., Nirei, T., Sunami, E., Nagawa, H. & Kitayama, J. Density of CD4 (+) and CD8 (+) T lymphocytes in biopsy samples can be a predictor of pathological response to chemoradiotherapy (CRT) for rectal cancer. *Radiation Oncology*. **6**, 1-6 (2011).

[200] Riemann, D., Hase, S., Fischer, K. & Seliger, B. Granulocyte-to-dendritic cell-ratio as marker for the immune monitoring in patients with renal cell carcinoma. *Clinical And Translational Medicine*. **3**, 1-6 (2014).

[201] Riemann, D., Cwikowski, M., Turzer, S., Giese, T., Grallert, M., Schütte, W. & Seliger, B. Blood immune cell biomarkers in lung cancer. *Clinical Experimental Immunology*. **195**, 179-189 (2019).

[202] Owen, M. & SHERRATT, J. Modelling the macrophage invasion of tumours: Effects on growth and composition. *Mathematical Medicine And Biology: A Journal Of The IMA*. **15**, 165-185 (1998).

[203] Pillis, L. & Radunskaya, A. A mathematical tumor model with immune resistance and drug therapy: an optimal control approach. *Computational And Mathematical Methods In Medicine*. **3**, 79-100 (2001).

[204] Pillis, L. & Radunskaya, A. The dynamics of an optimally controlled tumor model: A case study. *Mathematical And Computer Modelling*. **37**, 1221-1244 (2003).

[205] Voss, A. & Voss, J. A fast numerical algorithm for the estimation of diffusion model parameters. *Journal Of Mathematical Psychology*. **52**, 1-9 (2008).

[206] Parra-Rojas, C. & Hernandez-Vargas, E. PDEparams: parameter fitting toolbox for partial differential equations in python. *Bioinformatics*. **36**, 2618-2619 (2020).

[207] Vyshemirsky, V. & Girolami, M. BioBayes: a software package for Bayesian inference in systems biology. *Bioinformatics*. **24**, 1933-1934 (2008).

[208] Xun, X., Cao, J., Mallick, B., Maity, A. & Carroll, R. Parameter estimation of partial differential equation models. *Journal Of The American Statistical Association*. **108**, 1009-1020 (2013).

[209] Casali, P., Bielack, S., Abecassis, N., Aro, H., Bauer, S., Biagini, R., Bonvalot, S., Boukovinas, I., Bovee, J., Brennan, B. & Others Bone sarcomas: ESMO–PaedCan–EURACAN Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals Of Oncology*. **29** pp. iv79-iv95 (2018).

[210] American Cancer Society, Chemotherapy and Other Drugs for Osteosarcoma, 2021. Available from: https://www.cancer.org/cancer/osteosarcoma/treating/chemotherapy.html.

[211] Gardner, S. A mechanistic, predictive model of dose-response curves for cell cycle phase-specific and -nonspecific drugs. *Cancer Research*. **60**, 1417-1425 (2000).

[212] Frohman, E., Cruz, R., Longmuir, R., Steinman, L., Zamvil, S., Villemarette-Pittman, N., Frohman, T. & Parsons, M. Part II. High-dose methotrexate with leucovorin rescue for severe COVID-19: an immune stabilization strategy for SARS-CoV-2 induced 'PANIC'attack. *Journal Of The Neurological Sciences*. pp. 116935 (2020).

[213] Tacar, O., Sriamornsak, P. & Dass, C. Doxorubicin: An update on anticancer molecular action, toxicity and novel drug delivery systems. *Journal Of Pharmacy And Pharmacology*. **65**, 157-170 (2013).

[214] Raudenska, M., Balvan, J., Fojtu, M., Gumulec, J. & Masarik, M. Unexpected therapeutic effects of cisplatin. *Metallomics*. **11**, 1182-1199 (2019).

[215] Spanos, W., Nowicki, P., Lee, D., Hoover, A., Hostager, B., Gupta, A., Anderson, M. & Lee, J. Immune response during therapy with cisplatin or radiation for human papillomavirus-related head and neck cancer. *Archives Of Otolaryngology - Head And Neck Surgery*. **135**, 1137-1146 (2009).

[216] Pillis, L., Fister, K., Gu, W., Collins, C., Daub, M., Gross, D., Moore, J. & Preskill, B. Mathematical model creation for cancer chemo-immunotherapy. *Computational And Mathematical Methods In Medicine*. **10**, 165-184 (2009).

[217] Tseng, C., Hung, C., Alvarez, R., Trimble, C., Huh, W., Kim, D., Chuang, C., Lin, C., Tsai, Y., He, L. & Others Pretreatment with cisplatin enhances E7-specific CD8+ T-cell–mediated antitumor immunity induced by DNA vaccination. *Clinical Cancer Research*. **14**, 3185-3192 (2008).

[218] Grabosch, S., Bulatovic, M., Zeng, F., Ma, T., Zhang, L., Ross, M., Brozick, J., Fang, Y., Tseng, G., Kim, E. & Others Cisplatin-induced immune modulation in ovarian cancer mouse models with distinct inflammation profiles. *Oncogene*. **38**, 2380-2393 (2019).

[219] Mosteller, R. Simplified calculation of body-surface area.. *The New England Journal Of Medicine*. **317**, 1098-1098 (1987).

[220] Sendroy Jr, J. & Collison, H. Determination of human body volume from height and weight.. *Journal Of Applied Physiology*. **21**, 167-172 (1966).

[221] National Center for Biotechnology Information, PubChem Compound Summary for CID 126941, Methotrexate., 2021. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/Methotrexate.

[222] National Center for Biotechnology Information, PubChem Compound Summary for CID 5702198, Cisplatin., 2021. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/trans-Dichlorodiamineplatinum_II.

[223] National Center for Biotechnology Information, PubChem Compound Summary for CID 443939, Doxorubicin Hydrochloride., 2021. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/Doxorubicin-Hydrochloride.

[224] De Pillis, L., Gu, W. & Radunskaya, A. Mixed immunotherapy and chemotherapy of tumors: Modeling, applications and biological interpretations. *Journal Of Theoretical Biology*. **238**, 841-862 (2006).

[225] Perry, M. The chemotherapy source book. (Lippincott Williams & Wilkins, 2008).

[226] Medscape, Drugs & Diseases, doxorubicin (Rx), 2021. Available from: https://reference.medscape.com/drug/doxorubicin-342120showall.

[227] Drugbank Online, Cisplatin DrugBank Accession Number DB00515, 2021. Available from: https://go.drugbank.com/drugs/DB00515.

[228] Drugbank Online, Methotrexate DrugBank Accession Number DB00563, 2021. Available from: https://go.drugbank.com/drugs/DB00563.

[229] Marina, N., Smeland, S., Bielack, S., Bernstein, M., Jovic, G., Krailo, M., Hook, J., Arndt, C., Berg, H., Brennan, B., Brichard, B., Brown, K., Butterfass-Bahloul, T., Calaminus, G., Daldrup-Link, H., Eriksson, M., Gebhardt, M., Gelderblom, H., Gerss, J., Goldsby, R., Goorin, A., Gorlick, R., Grier, H., Hale, J., Hall, K., Hardes, J., Hawkins, D., Helmke, K., Hogendoorn, P., Isakoff, M., Janeway, K., Jürgens, H., Kager, L., Kühne, T., Lau, C., Leavey, P., Lessnick, S., Mascarenhas, L., Meyers, P., Mottl, H., Nathrath, M., Papai, Z., Randall, R., Reichardt, P., Renard, M., Safwat, A., Schwartz, C., Stevens, M., Strauss, S., Teot, L., Werner, M., Sydes, M. & Whelan, J. Comparison of MAPIE versus MAP in patients with a poor response to preoperative chemotherapy for newly diagnosed high-grade osteosarcoma (EURAMOS-1): an open-label, international, randomised controlled trial. *The Lancet Oncology*. **17**, 1396-1408 (2016).

[230] NSW Government, Osteosarcoma MAP (methotrexate, DOXOrubicin, cISplatin), 2021. Available from: https://www.eviq.org.au/medical-oncology/sarcoma/bone-sarcoma/1901-osteosarcoma-map-methotrexate-doxorubicin.

[231] Yuan, G., Chen, J., Wu, D. & Gao, C. Neoadjuvant chemotherapy combined with limb salvage surgery in patients with limb osteosarcoma of Enneking stage II: a retrospective study. *OncoTargets And Therapy*. **10** pp. 2745 (2017).

[232] Yin, Y., Hu, Q., Xu, C., Qiao, Q., Qin, X., Song, Q., Peng, Y., Zhao, Y. & Zhang, Z. Co-delivery of doxorubicin and interferon- by thermosensitive nanoparticles for cancer immunochemotherapy. *Molecular Pharmaceutics*. **15**, 4161-4172 (2018).

[233] Zitvogel, L., Apetoh, L., Ghiringhelli, F. & Kroemer, G. Immunological aspects of cancer chemotherapy. *Nature Reviews Immunology*. **8**, 59-73 (2008).

[234] Tongu, M., Harashima, N., Yamada, T., Harada, T. & Harada, M. Immunogenic chemotherapy with cyclophosphamide and doxorubicin against established murine carcinoma. *Cancer Immunology, Immunotherapy*. **59**, 769-777 (2010).

[235] Kawano, M., Tanaka, K., Itonaga, I., Iwasaki, T., Miyazaki, M., Ikeda, S. & Tsumura, H. Dendritic cells combined with doxorubicin induces immunogenic cell death and exhibits antitumor effects for osteosarcoma. *Oncology Letters*. **11**, 2169-2175 (2016).

[236] Apetoh, L., Mignot, G., Panaretakis, T., Kroemer, G. & Zitvogel, L. Immunogenicity of anthracyclines: moving towards more personalized medicine. *Trends In Molecular Medicine*. **14**, 141-151 (2008).

[237] Zhu, S., Waguespack, M., Barker, S. & Li, S. Doxorubicin Directs the Accumulation of Interleukin-12–Induced IFN into Tumors for Enhancing STAT1–Dependent Antitumor Effect. *Clinical Cancer Research*. **13**, 4252-4260 (2007).

[238] Shi, Y., Moon, M., Dawood, S., McManus, B. & Liu, P. Mechanisms and management of doxorubicin cardiotoxicity. *Herz*. **36**, 296-305 (2011).

[239] Amini, M., Abbasi, A., Cai, P., Lip, H., Gordijo, C., Li, J., Chen, B., Zhang, L., Rauth, A. & Wu, X. Combining tumor microenvironment modulating nanoparticles with doxorubicin to enhance chemotherapeutic efficacy and boost antitumor immunity. *JNCI: Journal Of The National Cancer Institute*. **111**, 399-408 (2019).

[240] Hannesdóttir, L., Tymoszuk, P., Parajuli, N., Wasmer, M., Philipp, S., Daschil, N., Datta, S., Koller, J., Tripp, C., Stoitzner, P. & Others Lapatinib and doxorubicin enhance the S tat1-dependent antitumor immune response. *European Journal Of Immunology*. **43**, 2718-2729 (2013).

[241] Wakita, D., Iwai, T., Harada, S., Suzuki, M., Yamamoto, K. & Sugimoto, M. Cisplatin augments antitumor T-cell responses leading to a potent therapeutic effect in combination with PD-L1 blockade. *Anticancer Research*. **39**, 1749-1760 (2019).

[242] Cronstein, B. The mechanism of action of methotrexate. *Rheumatic Disease Clinics Of North America*. **23**, 739-755 (1997).

[243] Casares, N., Pequignot, M., Tesniere, A., Ghiringhelli, F., Roux, S., Chaput, N., Schmitt, E., Hamai, A., Hervas-Stubbs, S., Obeid, M. & Others Caspase-dependent immunogenicity of doxorubicin-induced tumor cell death. *The Journal Of Experimental Medicine*. **202**, 1691-1701 (2005).

[244] Ujhazy, P., Zaleskis, G., Mihich, E., Ehrke, M. & Berleth, E. Doxorubicin induces specific immune functions and cytokine expression in peritoneal cells. *Cancer Immunology, Immunotherapy*. **52**, 463-472 (2003).

[245] Safavi, F. & Nath, A. Silencing of immune activation with methotrexate in patients with COVID-19. *Journal Of The Neurological Sciences*. **415** (2020).

[246] Cutolo, M., Sulli, A., Pizzorni, C., Seriolo, B. & Straub, R. Anti-inflammatory mechanisms of methotrexate in rheumatoid arthritis. *Annals Of The Rheumatic Diseases*. **60**, 729-735 (2001).

[247] Souhami, R., Craft, A., Eijken, J., Nooij, M., Spooner, D., Bramwell, V., Wierzbicki, R., Malcolm, A., Kirkpatrick, A., Uscinska, B. & Others Randomised trial of two regimens of chemotherapy in operable osteosarcoma: a study of the European Osteosarcoma Intergroup. *The Lancet*. **350**, 911-917 (1997).

[248] Cancer Therapy Advisor, Bone Cancer Treatment Regimens, 2021. Available from: https://www.cancertherapyadvisor.com/home/cancer-topics/bone-cancer/bone-cancer-treatment-regimens/bone-cancer-treatment-regimens.

[249] Saeter, G., Alveg, T., Elomaa, I., Stenwig, A., Holmström, T. & Solheim, O. Treatment of osteosarcoma of the extremities with the T-10 protocol, with emphasis on the effects of preoperative chemotherapy with single-agent high-dose methotrexate: a Scandinavian Sarcoma Group study.. *Journal Of Clinical Oncology*. **9**, 1766-1775 (1991).

[250] Zhang, B., Zhang, Y., Li, R., Li, J., Lu, X. & Zhang, Y. The efficacy and safety comparison of first-line chemotherapeutic agents (high-dose methotrexate, doxorubicin, cisplatin, and ifosfamide) for osteosarcoma: A network meta-analysis. *Journal Of Orthopaedic Surgery And Research*. **15**, 1-10 (2020).

[251] Yu, D., Zhang, S., Feng, A., Xu, D., Zhu, Q., Mao, Y., Zhao, Y., Lv, Y., Han, C., Liu, R. & Tian, Y. Methotrexate, doxorubicin, and cisplatinum regimen is still the preferred option for osteosarcoma chemotherapy: A meta-analysis and clinical observation. *Medicine (United States)*. **98**, 1-8 (2019).

[252] Cardinale, D., Colombo, A., Sandri, M., Lamantia, G., Colombo, N., Civelli, M., Martinelli, G., Veglia, F., Fiorentini, C. & Cipolla, C. Prevention of High-Dose Chemotherapy–Induced Cardiotoxicity in High-Risk Patients by Angiotensin-Converting Enzyme Inhibition. *Circulation*. **114**, 2474-2481 (2006).

[253] Blijham, G. Prevention and treatment of organ toxicity during high-dose chemotherapy: an overview.. *Anti-cancer Drugs*. **4**, 527-533 (1993).

[254] Spratt JR, J. The rates of growth of skeletal sarcomas. *Cancer*. **18**, 14-24 (1965).

[255] Patel, A., Zhang, Y., Fullerton, J., Boelen, L., Rongvaux, A., Maini, A., Bigley, V., Flavell, R., Gilroy, D., Asquith, B. & Others The fate and lifespan of human monocyte subsets in steady state and systemic inflammation. *Journal Of Experimental Medicine*. **214**, 1913-1923 (2017).

[256] Italiani, P. & Boraschi, D. From monocytes to M1/M2 macrophages: phenotypical vs. functional differentiation. *Frontiers In Immunology*. **5** pp. 514 (2014).

[257] He, Z., Allers, C., Sugimoto, C., Ahmed, N., Fujioka, H., Kim, W., Didier, E. & Kuroda, M. Rapid turnover and high production rate of myeloid cells in adult rhesus macaques with compensations during aging. *The Journal Of Immunology*. **200**, 4059-4067 (2018).

[258] Ginhoux, F. & Guilliams, M. Tissue-resident macrophage ontogeny and homeostasis. *Immunity*. **44**, 439-449 (2016).

[259] Hao, W. & Friedman, A. Serum upar as biomarker in breast cancer recurrence: A mathematical model. *PLoS One*. **11**, e0153508 (2016).

[260] Farber, D., Yudanin, N. & Restifo, N. Human memory T cells: generation, compartmentalization and homeostasis. *Nature Reviews Immunology*. **14**, 24-35 (2014).

[261] De Boer, R., Homann, D. & Perelson, A. Different dynamics of CD4+ and CD8+ T cell responses during and after acute lymphocytic choriomeningitis virus infection. *The Journal Of Immunology*. **171**, 3928-3935 (2003).

[262] Vukmanovic-Stejic, M., Zhang, Y., Cook, J., Fletcher, J., McQuaid, A., Masters, J., Rustin, M., Taams, L., Beverley, P., Macallan, D. & Others Human CD4+ CD25 hi Foxp3+ regulatory T cells are derived by rapid turnover of memory populations in vivo. *The Journal Of Clinical Investigation*. **116**, 2423-2433 (2006).

[263] Cella, M., Engering, A., Pinet, V., Pieters, J. & Lanzavecchia, A. Inflammatory stimuli induce accumulation of MHC class II complexes on dendritic cells. *Nature*. **388**, 782-787 (1997).

[264] Diao, J., Winter, E., Cantin, C., Chen, W., Xu, L., Kelvin, D., Phillips, J. & Cattral, M. In situ replication of immediate dendritic cell (DC) precursors contributes to conventional DC homeostasis in lymphoid tissue. *The Journal Of Immunology*. **176**, 7196-7206 (2006).

[265] Foon, K., Sherwin, S., Abrams, P., Stevenson, H., Holmes, P., Maluish, A., Oldham, R. & Herberman, R. A phase I trial of recombinant gamma interferon in patients with cancer. *Cancer Immunology, Immunotherapy*. **20**, 193-197 (1985).

[266] Fuentes-Calvo, I. & Martınez-Salgado, C. TGFB1 (transforming growth factor, beta 1). *Atlas Of Genetics And Cytogenetics In Oncology And Haematology*. (2013).

[267] Saxena, A., Khosraviani, S., Noel, S., Mohan, D., Donner, T. & Hamad, A. Interleukin-10 paradox: A potent immunoregulatory cytokine that has been difficult to harness for immunotherapy. *Cytokine*. **74**, 27-34 (2015).

[268] Conlon, P., Tyler, S., Grabstein, K. & Morrissey, P. Interleukin-4 (B-cell stimulatory factor-1) augments the in vivo generation of cytotoxic cells in immunosuppressed animals.. *Biotechnology Therapeutics*. **1**, 31-41 (1989).

[269] Khodoun, M., Lewis, C., Yang, J., Orekov, T., Potter, C., Wynn, T., Mentink-Kane, M., Hershey, G., Wills-Karp, M. & Finkelman, F. Differences in expression, affinity, and function of soluble (s) IL-4R and sIL-13R2 suggest opposite effects on allergic responses. *The Journal Of Immunology*. **179**, 6429-6438 (2007).

[270] Mehra, R., Storfer-Isser, A., Kirchner, H., Johnson, N., Jenny, N., Tracy, R. & Redline, S. Soluble interleukin 6 receptor: a novel marker of moderate to severe sleep-related breathing disorder. *Archives Of Internal Medicine*. **166**, 1725-1731 (2006).

[271] Balestrino, M. Cytokine imbalances in multiple sclerosis: a computer simulation. (2009).

[272] Zandarashvili, L., Sahu, D., Lee, K., Lee, Y., Singh, P., Rajarathnam, K. & Iwahara, J. Real-time kinetics of high-mobility group box 1 (HMGB1) oxidation in extracellular fluids studied by in situ protein NMR spectroscopy. *Journal Of Biological Chemistry*. **288**, 11621-11627 (2013).