UTRECHT UNIVERSITY

Master Thesis

A dose-rate (and dose) dependent model for radiation-induced murine Acute Myeloid Leukaemia (AML), and a possible dose and dose-rate effectiveness function (DDREFu)

written by

Sophie Palmer

under the supervision of

Sjors Stouten (RIVM) and Sjoerd Verduyn Lunel (UU)

A thesis submitted in partial fulfillment for the
degree of Masters in Mathematical Sciences
in combination with an Internship at RIVM

June 2019

UTRECHT UNIVERSITY

Department of Mathematics

Abstract

Due to the increase in medical imaging and clinical radiation use, it is important to be able to quantify the health implications from low doses of radiation delivered at low dose rates. There is a lot of data available on the incidence of different types of cancer as a response to high doses of radiation when delivered at high dose rates, and many mathematical models have been developed from these data, however much less data is available on the incidence of cancer caused by low levels of radiation received at low dose rates. In this thesis we present a model to predict the incidence of cancer, specifically murine AML as a function of dose and dose rate, and give recommendations for DDREF functions (dose and dose rate effectiveness functions).

Contents

Chapter 1

Introduction

People are exposed to ionising radiation through many different sources: for example doctors performing X-rays on patients, patients receiving CT (Computed Tomography) scans or radiotherapy, and workers in the nuclear power industry. Also ionising radiation occurs naturally in the air from radioactive material such as radon, as terrestrial radiation in the soil and as cosmic radiation from space [1]. Radiation can be very damaging to living tissue, particularly at a cellular level, which will be the focus of this thesis. High doses (above 1 Gy) delivered at high dose-rates (above 1 Gy/hour), cause extensive damage that results in almost immediate cell death or many mutations that lead to a high risk of cancer development, and it is quite easy to measure this damage since the effects are usually observed quite soon after radiation exposure. Damage caused by low dose and low dose-rate radiation often takes much longer to show up, since this damage is usually induced by stochastic events, and therefore the health implications are much harder to measure. It is very important to be able to quantify the risk to low dose and low dose-rate radiation so that we can make developments in medical radiation procedures, radiotherapy and radiation protection policies. The effective dose received from X-ray imaging varies from 0.01 to 10 mGy. A CT scan delivers higher doses of between 1 and 20 mGy, depending on the part of the body [2]. As a comparison, the average yearly dose of background radiation people receive is 3 mGy. Radiotherapy usually issues a total dose of between 50 and 70 Gy, divided into small amounts of around 2 Gy per day, five days a week. Figure 1.1 shows the percent of malignant cells killed, defined here as tumor control, along with the healthy tissue damage as a response to dose [3]. Here it is clear that precision in dosage is required to effectively kill as many tumor cells as possible without causing damage to healthy cells. Due to the limited available data and difficulty of carrying out experiments, there is much uncertainty in the risk from low dose and low dose-rate radiation exposure. Mathematical models have been built

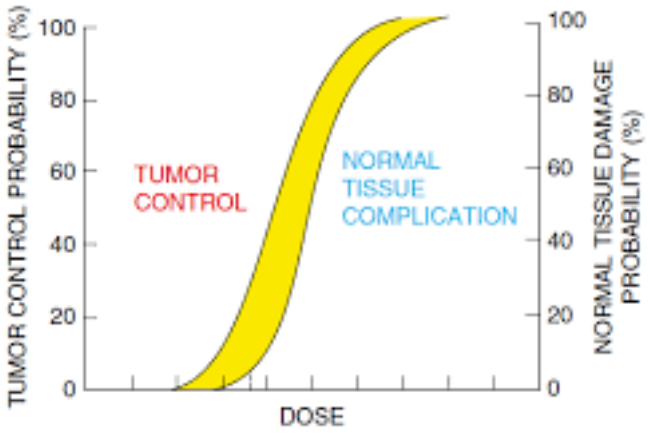


FIGURE 1.1: Tumor control and normal tissue damage responses to dose

from data known about high dose and high dose-rate risk, that enable extrapolation to estimate low dose risk.

1.1 Overview of previous research

There are many different approaches that have been used to model radiation-induced DNA damage as a function of dose, and the surviving fraction of a cell population [4]. One of the first approaches was Target Theory, which assumes that there are sensitive targets in the cell and the modelling is based on the probability of the radiation track hitting these specific targets in cells. Then came the Linear-Quadratic model developed by Kellerer and Rossi in 1972, which suggests that a lethal event, often called a chromosome aberration, is caused by either one particle track, represented by a linear component, or two independent particle tracks, represented by a quadratic component. Equation 1.1 shows this relation, where N_{CA} represents the number of chromosome aberrations per cell, and D represents the dose of radiation received.

$$N_{CA} = \alpha D + \beta D^2 \tag{1.1}$$

Kellerer and Rossi also suggested that the surviving fraction of a cell population is related exponentially to the number of chromosome aberrations per cell, see equation 1.2, where SF represents the surviving fraction.

$$SF = e^{-(N_{CA})} \tag{1.2}$$

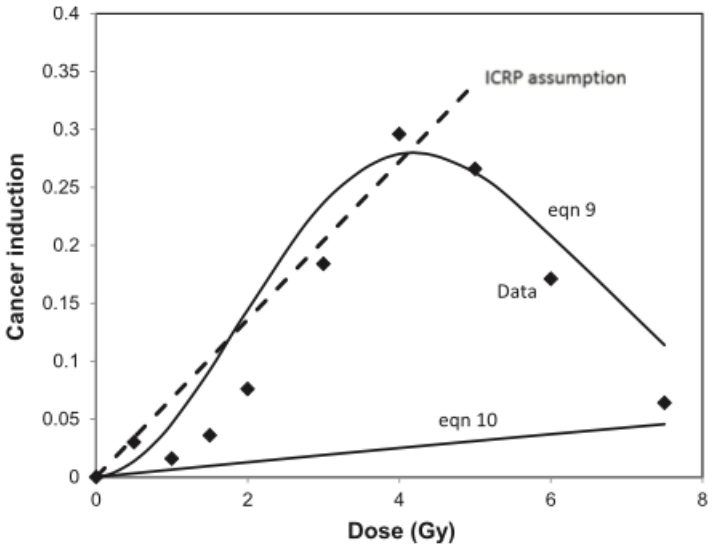


FIGURE 1.2: Taken from Chadwick (2017) [7], **eqn 9** in this figure refers to cancer induction from high dose-rate exposure, and **eqn10** refers to cancer induction from low dose-rate exposure. The diamonds are data points, and the dotted line shows the Linear No-Threshold model.

From understanding about cellular responses to radiation exposure and epidemiological data on cancer incidence from high doses at high dose-rates, for example survivors of the Hiroshima and Nagasaki atomic bombings, models have been developed to predict the risk of cancer from a given dose of radiation, assuming the radiation is delivered all at once. For example, the Linear No-Threshold model describes a linear relationship between dose and cancer risk, and was built for radiation protection purposes to extrapolate the risk to low dose radiation exposure [5]. This model suggests that the risk from a given dose is the same regardless of the rate at which that dose is delivered, so that separate exposure events can be summed to estimate the risk from the total accumulated dose. However, there is much controversy about extrapolation by a linear dependency, as this is thought to lead to over-estimations for low dose and low dose-rate risk [6]. To try and compensate for these over-estimations, the International Commission on Radiation Protection (ICRP) proposed in 2007 a dose and dose-rate effectiveness factor (DDREF) of 2, by which high dose-rate risk for a given dose is divided by this factor to estimate low dose-rate risk. Even though a DDREF of 2 gives improved risk estimates for some doses, this factor is thought to be quite different depending on the dose of radiation received [7], as shown in figure 1.3.

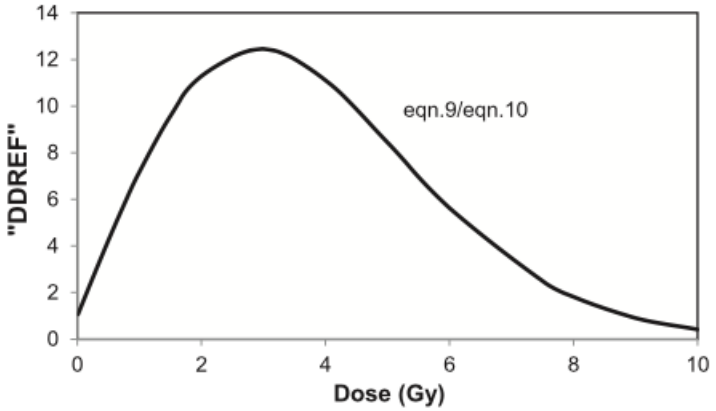


FIGURE 1.3: From figure 1.2, a DDREF for different doses is calculated by dividing **eqn9** by **eqn10**.

1.2 Motivation and Research Focus

The uncertainty in the health implications caused by low dose and low dose-rate radiation exposure sparks the motivation for this thesis, which will try to understand more about low dose and low dose-rate radiation exposure and make better risk estimates in the consequences of this type of radiation. With a specific focus on the risk of murine Acute Myeloid Leukaemia (AML), this thesis will investigate the dose-rate effects on cellular damage and the formation of malignant cells, to present a dose-rate dependent model for AML. As a starting point the dose dependent model developed by Dekkers et al (2011) [8] will be considered, which predicts the incidence of AML for a given dose, as shown in figure 1.4.

From the dose-rate dependent model for AML, a DDREF will be calculated which is expected to vary for different doses.

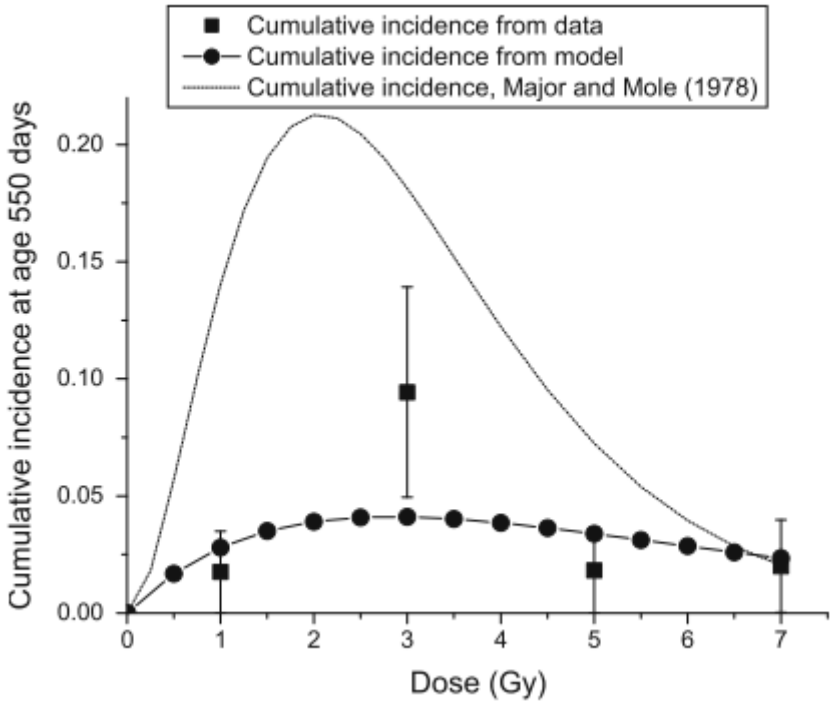


FIGURE 1.4: Predicted incidence of Acute Myeloid Leukemia by Dekkers et al (2011) [8].

Chapter 2

Biological Background

This chapter will provide a background of biology in order to make the modelling section clearer. It is fascinating that cells are able to identify significant damage to the DNA, pause the cell cycle while repair takes place, and then continue the cell cycle once the damage is sufficiently repaired. Repair of the damage occurs at a very high accuracy, but since this is a natural process sometimes things can go wrong and the mis-repair of DNA can sometimes lead to gene mutations, which can cause the onset of cancer.

2.1 What is Ionising Radiation

Radiation can be ionising or non-ionising. Ionising radiation consists of radiation waves at the high frequency end of the electromagnetic spectrum: from γ-rays and X-rays down to extreme UV light, whose wavelengths range from 10^{-5}nm to 10nm. Heavy particles such as alpha and beta particles are also a source of ionising radiation. These particles have a very unstable outer electron shell, and therefore decay easily by losing electrons and becoming ionised. The process of losing electrons generates waves of ionising radiation. It is ionising radiation that causes damage to cells, because there is enough energy to displace electrons from atoms, damaging the molecule. If the radiation track hits a molecule in the DNA of a cell, this can have serious effects on the DNA structure which can be fatal to the cell [9].

2.2 How Ionising Radiation causes damage to DNA

DNA is made up of nitrogenous bases and sugar-phosphate molecules, held together by hydrogen bonds, wound tightly in a helix. Chromosomes are made of this tightly wound

DNA, and can be thought of as an arm that becomes replicated during the process of cell division. When an ionising track hits a sugar molecule of a DNA base, the displacement of an electron from this sugar molecule disrupts the structure and causes damage to the backbone of DNA, resulting in a single-stranded DNA break (SSB). If two SSBs occur opposite each other within 10 base pairs, this causes a double-stranded break (DSB) [9]. If a cell nucleus receives a high number of DSBs in its DNA, this can lead to apoptosis (programmed cell death). Experimental results show that on average, a human cell will receive 35 DSBs per Gy of X-ray irradiation [10], irrespective of the dose-rate. The DNA of a human diploid cell is known to contain 6.1 Giga base pairs (Gbp), so this gives a yield of approximately **5.7 DSBs/Gy/Gbp** [10], which is now a widely used constant for all species.

2.3 Repair Pathways of DSBs

Cells cycle through different phases during their lifetime in order to carry out DNA replication and cell division, as shown in figure 2.1. DNA replication happens throughout the synthesis (S) phase, and in the mitosis (M) phase the cell divides so both new cells have non-replicated DNA again. There are a number of different repair pathways, some which can only take place when the DNA is replicated and these have a higher accuracy of repair. Therefore the current phase of the cell cycle upon radiation induced damage can have an impact on the accuracy of repair. DNA that is replicated consists of chromosomes that have two identical copies, called sister chromatids, which are held together by a centromere. The next sections will describe pathways of repair of radiation induced double stranded breaks.

2.3.1 Non-Homologous End Joining (NHEJ)

In this repair pathway, the broken ends are identified, bound to by a protein and fixed back together. This pathway is used when the DNA is not replicated, since there is no template DNA available [9].

2.3.2 Homologous Recombination (HR)

This repair pathway can only take place when DNA is replicated. It uses the intact sister chromatid as a template to repair the broken chromatid, meaning that this pathway seldom makes errors. Even if DNA is replicated however, the NHEJ pathway can still sometimes used [9].

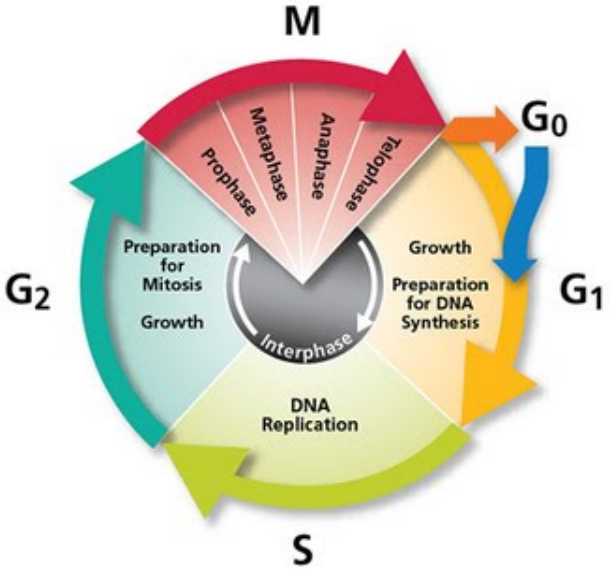


FIGURE 2.1: Cell Cycle phases [11]

2.4 Mutations

Despite the high accuracy of these repair processes, especially HR, errors can sometimes be made. For example, the wrong base could be used, or the wrong ends joined back together. This leads to different severities of repair errors. Those involving just a few DNA bases are known as point mutations, see figure 2.2.

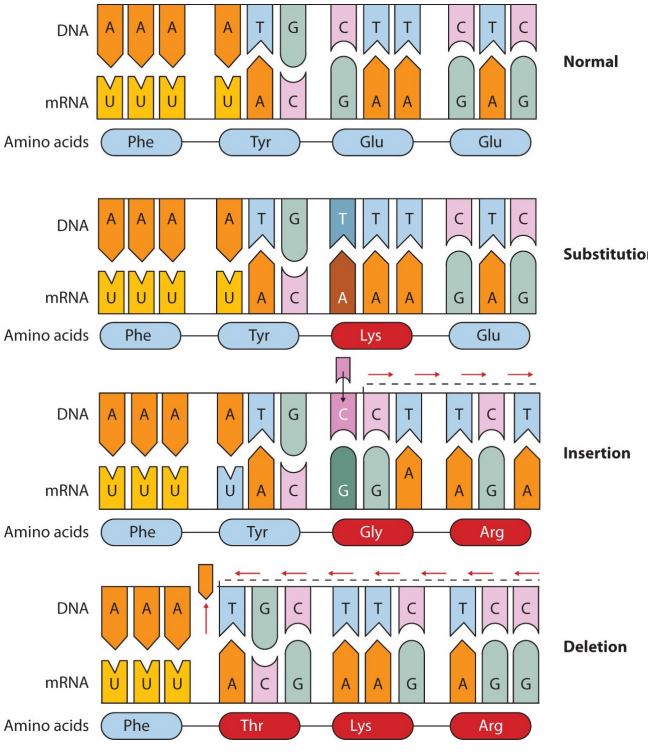


FIGURE 2.2: Genetic Mutations [12]

Some examples of more severe mis-repair events that can occur to a whole chromosome are:

- Dicentric Chromosome: Two chromatids both with a DSB, forming a dicentric chromatid with two centromeres and the acentric part is lost.
- Ring: A chromatid has a break at both ends i.e. either side of the centromere, and it repairs itself as a ring.
- Anaphase Bridge: A chromosome gets a break on each chromatid and repairs itself as a bow-like structure, which upon division creates one long chromatid.
- Deletion: A chromatid receives two DSBs along a length of DNA and the ends are repaired incorrectly, causing a chunk of a DNA to be lost, see figure 2.3. The lost chunk of DNA could code for one or several genes.

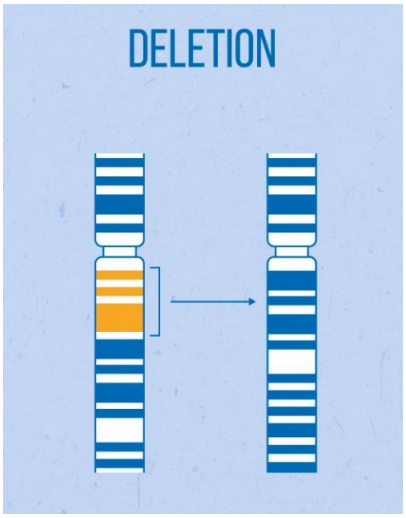


FIGURE 2.3: Gene Deletion

These damages to chromosomes are known as **chromosome aberrations**, and mathematical models such as the Linear-Quadratic model as discussed in chapter 1, have been developed to describe the number of chromosome aberrations formed per cell as a function of the dose of radiation received.

All of these mutations result in that section of DNA code not being transcribed correctly, therefore the gene that is coded for in this section will not be expressed fully, if at all. Genes code for proteins, a process called the central dogma of biology, and proteins have many important roles in the functioning of a cell. If some are not expressed, this can lead to the malfunctioning of large cellular components. As described below, experimental evidence shows that certain mutations are a likely pathway to the diagnosis of cancer [13].

Mutations can also occur spontaneously during DNA replication. Although DNA replication is extremely accurate most of the time, due to the large number of cell divisions occurring throughout the lifetime of an organism, there is a small probability that replication sometimes induces some errors. Since this is a random event, these spontaneous mutations are often modelled as a stochastic process, to estimate the chance of one occurring in the lifetime of an organism.

2.5 Measurement of Cell Survival

A lot of damage in a cell can lead to apoptosis, a process of programmed cell death. Therefore the higher the dose of radiation, the more DSBs and the more likely that the cell will die.

Despite more recent models for cell survival that include DNA damage repair, the *Linear-Quadratic model* developed by Kellerer and Rossi in 1972, as mentioned in Chapter 1, remains the most widely used as it is seen to be the best fitting model to cell survival data [4]. As described before, this model proposes that a lethal event (mis-repair of DSBs leading to mutations or chromosome aberrations) occurs either by a hit from one single ionising radiation track, or as the consequence of two ionising tracks. The number of lethal events per cell, N, is hence proportional to $\alpha D + \beta D^2$, where D represents the dose of radiation, and hence β incorporates some dose-rate dependency as two ionising tracks are more likely to lead to a lethal event if they happen within a shorter period of each other, i.e. if the dose-rate is higher.

Kellerer and Rossi then proposed that the surviving fraction of cells decreases exponentially with the number of lethal events, in the following relation:

$$\text{Surviving Fraction } (D) = e^{-(\alpha D + \beta D^2)} \tag{2.1}$$

Experiments on cells in petri dishes have been done in order to measure the surviving fraction of a cell population exposed to radiation. These experiments involve exposing a parallel dish seeded with cells to a dose of x-rays, incubating the dish for 1 to 2 weeks, then counting the number of colonies that have grown. If an original cell has managed to grow into a colony it still has its reproductive integrity intact. [9].

The fraction of surviving cells is then calculated by equation 2.2:

$$\text{Surviving Fraction} = \frac{\text{Colonies counted}}{\text{Cells seeded} * PE} \tag{2.2}$$

where PE is the *Plating Efficiency*: the number of colonies counted as a fraction of the number of seeds planted in a control experiment. The parameters α and β of equation 2.1 can then be inferred from the results of experiments like this.

Cell death can be defined as the loss of reproductive integrity, in the case when cells cannot give rise to the differentiated cells they are meant to produce. A surviving cell that is still able to produce the differentiated cell types indefinitely, is called clonogenic. Since the model in this thesis focuses on stem cells of the blood known as hematopoietic stem cells (HSCs), as the target cells of radiation absorption, the cell survival fractions used in the model will relate to HSCs that still have their reproductive integrity, and the fraction that has ”died” could still be present but they are not able to ”give birth” to different cell types.

2.6 Genomic Instability

Genomic instability is central to carcinogenesis. It is well known that radiation can cause an instability in cells that is inherited over many generations through DNA replication and cell division. This instability can mean cells are more prone to errors during cell division, and one consequence is an increased rate of chromosome aberrations [14].

2.7 Acute Myeloid Leukemia (AML)

Hematopoietic stem cells differentiate into myeloid or lymphatic types. Red blood cells are formed from the myeloid lineage and make up the bone marrow tissue, and white blood cells are formed from both the myeloid and lymphatic lineage and are produced as a response to foreign bodies. Figure 2.4 shows the cell differentiation in the hematopoietic system, and literature suggests that the target cells of radiation are either the HSCs or the common myeloid progenitor cells [8].

AML is a cancer of the myeloid blood cells. Some symptoms of AML include unusual bleeding, such as bleeding gums or nosebleeds, weight-loss and tiredness [16].

Exposures to levels of ionising radiation between 1 and 10 Gy primarily affect the hematopoietic system (the blood system), one of the most radio-sensitive tissues in the human body due to its highly proliferate activity [17]. One consequential effect of radiation on the human hematopoietic system is the increased risk of AML, and experiments on mice show approximately 25% incidence of AML after exposure to 3 Gy of X-rays delivered at a high dose-rate, [8]. Since the mechanism that causes AML in

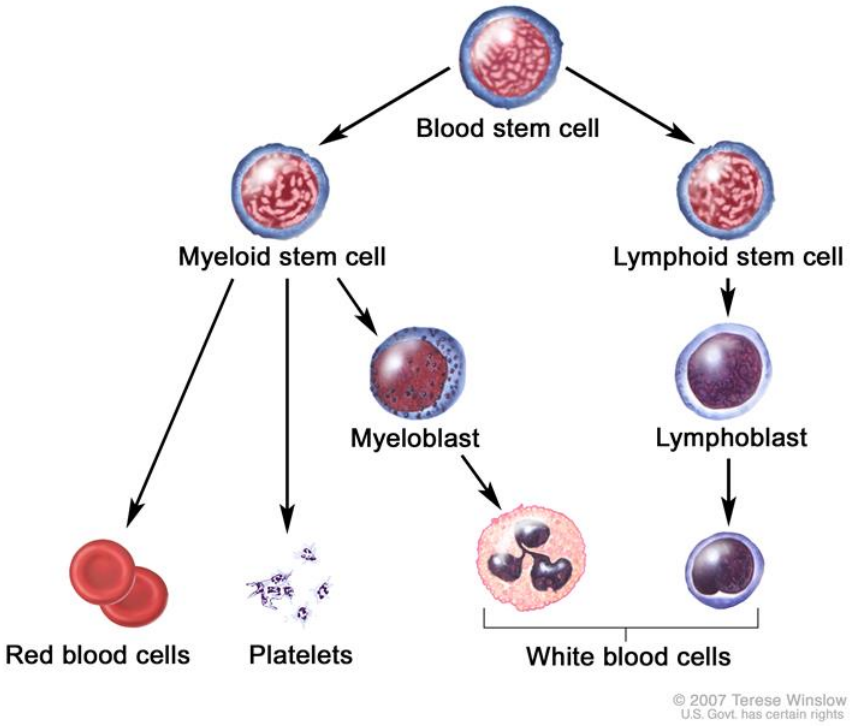


FIGURE 2.4: HSC Lineages [15]

humans is not understood exactly and because of ethical limitations, mice are often used in AML research because 99% of their protein-coding genes are the same as in humans. Specifically, the CBA strain of mice are most commonly used since they show a very low spontaneous rate of AML.

It has been found that in cases of AML in mice, more than 95% of these have a deletion of one copy of the *Sfpi1* gene on chromosome 2, [18]. Also, approximately 85% of these cases have been found to hold a point mutation, typically a base substitution at codon 235, in the remaining copy of the *Sfpi1* gene on the second allele of chromosome 2, ([18], [13], [19], [20]). Due to these observations it is assumed that this is the likely molecular mechanism that causes the onset of AML.

The *Sfpi1* gene codes for the protein PU.1 which is a transcription factor that controls the differentiation of hematopoietic cells into their different cell lineages [20]. It also plays a role in maintaining internal functional properties of the HSCs. Furthermore, the disruption of PU1 in myeloid progenitors is seen to inhibit their maturation, which could result in a depleted population of red blood cells, thus causing the onset of leukaemia [20].

2.8 Dose-Rate Effects

As mentioned in the introduction, there are observations of differences in cellular responses to radiation depending on whether it is received at a high dose-rate or a low dose-rate. When DNA DSBs are formed, part of the repair process involves the phosphorylation of a histone protein called $H2AX$ which is contained in the DNA chromatin, causing $H2AX$ to become $\gamma H2AX$. The phosphorylation of $H2AX$ can be measured by fluorescent markers, enabling the measurement of DSBs. Experiments such as those discussed in paper [21], show an increase of $\gamma H2AX$ foci with dose of radiation, but the amount of $\gamma H2AX$ is reduced if the radiation is delivered at a lower dose-rate. Figure 2.5 is taken from this paper and shows graphical results of this experiment. This motivates the further research into investigation of dose-rate effects on the incidence of cancer.

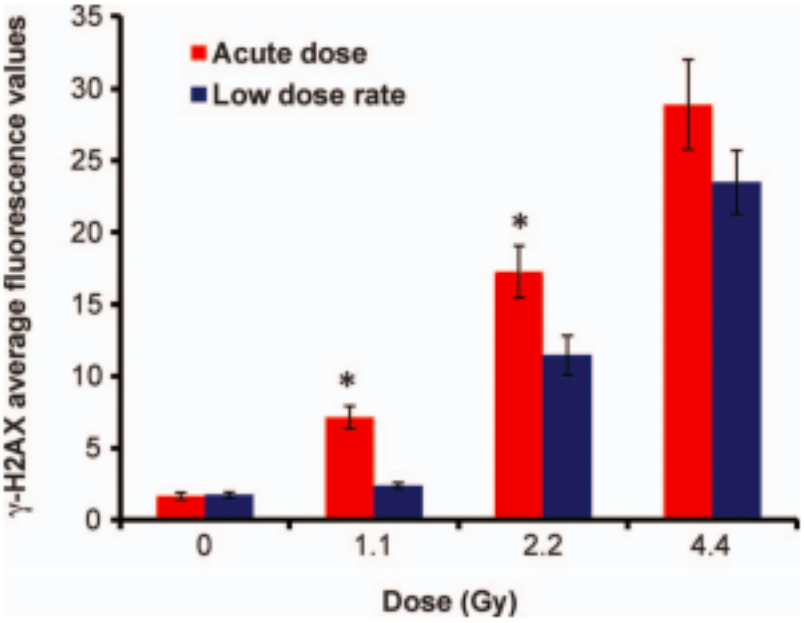


FIGURE 2.5: Dose-rate effects on γH2AX levels [21].

The number of DSBs formed have been found to be proportional to the number of chromosome aberrations that are induced [7], since if more DSBs are formed, the number of mis-repaired breaks is likely to be higher, leading to more chromosome aberrations. Therefore there are differences in the number of chromosome aberrations that are formed as a result of radiation delivered at different dose-rates. This observation leads us to believe there are dose-rate effects on the incidence rates of Acute Myeloid Leukemia.

Chapter 3

Simulating Stochastic Processes and Parameter Estimation

In this chapter, a framework of some of the modelling and statistical methods used in this thesis are defined.

3.1 Gillespie Algorithm

There are many algorithms that can be used to simulate a stochastic process. One that is used a lot in chemical and biological systems of reactions is the *Gillespie algorithm*. This algorithm simulates a feasible trajectory based on the stochastic dynamics of the system and is ideal for systems that have only a few possible reactions.

For example, consider a simple reversible chemical reaction between two molecules A and B that can form a compound AB, and call K_1 the forward rate and K_2 the backward rate, as defined in equations 3.1 and 3.2.

$$A + B \underset{K_1}{\rightarrow} AB \tag{3.1}$$

$$AB \underset{K_2}{\rightarrow} A + B \tag{3.2}$$

Let N_A, N_B and N_{AB} be the numbers of molecules of A, B and AB respectively, and define the current state of the system as $\mathbf{x}(t) = (N_A(t), N_B(t), N_{AB}(t))$. The total reaction rate is calculated as $R_{total} = K_1 N_A N_B + K_2 N_{AB}$. See algorithm 1 for a formal description of the algorithm.

Algorithm 1 Gillespie Algorithm

Input: Initial conditions: $\mathbf{x}(0)$, rates of reactions: K_1, K_2.
Output: The resulting state of the system at desired end time T: $\mathbf{x}(T)$

while $t < T$ **do**
  Monte Carlo step:
  Pick u_1 where $u_1 \sim$ unif(0,1)
  $r_1 = K_1 N_A(t) N_B(t)$, $r_2 = K_2 N_{AB}$
  $R_{total} = r_1 + r_2$
  $dt = -log(u_1)/R_{total}$
  Reaction threshold: $z = r_1/R_{total}$
  Pick u_2 where $u_2 \sim$ unif(0,1)
  Update step:
  $t = t + dt$
  if $u_2 < z$ **then**
    $\mathbf{x}(t) = (N_A - 1, N_B - 1, N_{AB} + 1)$
  else
    $\mathbf{x}(t) = (N_A + 1, N_B + 1, N_{AB} - 1)$
  end if
end while

This algorithm can be quite inefficient when simulating for long time periods if there are many reactions. The *Tau-leaping algorithm* approximates the *Gillespie algorithm* by taking a chosen time step τ and simulating all reactions a certain number of times, proportional to their reaction rates. For example, given reactions p_i and rates r_i for $i \in [1, n]$, then define $M_i \sim Poisson(r_i \tau)$ as the number of times reaction p_i occurs during the time interval $[t, t + \tau)$. Let the reactants in the system be defined as x_j, for $j \in [1, m]$. Then the state of the system is updated accordingly based on the numbers of each reaction that took place. Let $u_{i,j}$ be the effect on reactant x_j by reaction p_i occurring, then the new state of the system at time $t + \tau$ is defined as $\mathbf{x}(t + \tau) = \mathbf{x}(t) + \sum_i M_i u_{i,j}$. The value of τ is chosen dependent on the size of the reaction rates at the given time; for example during radiation exposure the rate of the *Sfpi1* deletion occurring is much higher than the rate of the *Sfpi1* point mutation occurring post radiation exposure, therefore the size of τ chosen in the post radiation exposure period is much larger than the size of τ chosen for the radiation exposure period.

3.2 Bayesian Inference

Bayesian Inference is a statistical method used to search for values of variables or parameters that are most likely, given some observed data, by making use of Bayes' rule. This method gives a probability distribution of the variables or parameters under consideration, which is a function describing the probability of the parameter(s) taking certain values, given the data that is observed. Some formal definitions are now introduced.

- Let $\boldsymbol{X} \in \mathbb{R}^n$ be a set of data points,
- $\boldsymbol{\theta} \in \mathbb{R}^n$ a set of parameters,
- $\boldsymbol{\alpha} \in \mathbb{R}^n$ a set of hyper-parameters of the parameter $\boldsymbol{\theta}$'s distribution.

From these definitions, the following distributions are defined. Firstly, let $p(a|b)$ denote the conditional probability of an event a happening, given that an event b has happened. For $\theta \in \boldsymbol{\theta}$,

- **Prior distribution:** $p(\theta \mid \alpha)$, is the probability distribution of the parameter θ given its hyper-parameter α before any data is observed, therefore it represents any prior knowledge about the value of θ.
- **Marginal likelihood:** $p(\boldsymbol{X} \mid \alpha) = \int p(\boldsymbol{X} \mid \theta) \cdot p(\theta \mid \alpha)\, d\theta$, is the distribution of the sample data set $\boldsymbol{X}$, given parameter α, integrated out over the range of θ.
- **Sampling distribution:** $p(\boldsymbol{X} \mid \theta)$, is the probability distribution of the sample data set $\boldsymbol{X}$, given the parameter θ.
- **Posterior distribution:** $p(\theta \mid \boldsymbol{X}, \alpha) = \frac{p(\boldsymbol{X}|\theta,\alpha)\cdot p(\theta|\alpha)}{p(\boldsymbol{X}|\alpha)}$, the probability distribution of the parameter θ given the sample data set $\boldsymbol{X}$ and the hyper-parameter α of θ, which is calculated using Bayes' rule.

Bayesian Inference generates posterior distributions of the parameters $\boldsymbol{\theta}$, which could be parameters of a model, to improve the accuracy of the model. Any prior knowledge known about the parameters can be entered in the prior distributions. The parameters $\boldsymbol{\theta}$ to be estimated need a defined starting value and a range of possible values that the parameter can take. Samples are taken from the prior distribution to determine the likelihood of the data set given the sampled values, which defines the sampling distribution. **Markov chain Monte Carlo** (MCMC) methods are a family of algorithms that sample from a probability distribution, therefore many Bayesian Inference methods use MCMC algorithms.

The package *Stan* is a package for *R Studio* that performs Bayesian Inference using the *No-U-Turn Sampler (NUTS)* algorithm, which is an adaptation of the *Metropolis-Hastings* algorithm. The steps of the *Metropolis-Hastings* algorithm are described below. A starting value for the parameter is chosen, θ_0, and an arbitrary symmetric probability density $Q(x \mid y)$ used for the proposal distribution. Usually $Q(x \mid y)$ is taken as the Gaussian distribution. Let $f(x)$ be a function that is proportional to the desired posterior distribution $p(\theta \mid \boldsymbol{X}, \alpha)$.

At each iteration t:

1. Pick a value θ' from the proposal distribution $Q(\theta' \mid \theta_t)$
2. Calculate the acceptance ratio $\alpha = f(\theta')/f(\theta_t)$, which is also equal to $p(\theta' \mid \boldsymbol{X}, \alpha)/p(\theta_t \mid \boldsymbol{X}, \alpha)$
3. Generate a random number u from the uniform distribution, $u \sim [0, 1]$.
4. If $u \leq \alpha$, then the value θ' is accepted and $\theta_{t+1} = \theta'$, otherwise θ_{t+1} keeps the current value θ_t.

The sampling creates a trajectory, and the idea behind *NUTS* is that once the trajectory begins to turn back on itself, the simulation is stopped.

3.3 Simulated Annealing

Simulated Annealing is a meta-heuristic of local search used for optimisation problems, specifically to try and converge to the global optimum in a large search space. It is a method derived from a physical problem of slowly cooling a material such that the size of the crystals formed when the material solidifies are maximised and form a solid structure with minimum free energy. If the material is cooled quickly, the crystals can form an irregular structure with high energy, which is not favourable.

The method starts with an initial positive temperature $T_0 \in \mathbb{R}^+$, and an initial starting solution to the problem. An important concept of this method is that during the search process, solutions that are worse than the current solution are accepted with a small probability, to ensure that the search does not get stuck in a local optimum. The temperature T is slowly decreased to zero with a fixed incremental factor α, while accepting worse solutions with a decreasing probability. The resulting solution should be the global optimum if the correct parameters T_0 and α are chosen.

Chapter 4

Modelling

This chapter will introduce some previous research on the dose-dependent incidence of AML developed by *Dekkers et al*, and follow on to discuss the dose-rate dependent model developed in this thesis to predict the incidence of AML.

It is important to note that other factors can also increase the risk of AML, such as genetic disorders, blood conditions and smoking. The model in this thesis is only concerned with the increased risk of AML due to radiation exposure, so if any of these other factors are present the risk could be increased even more.

4.1 Previous research: A dose-dependent two-mutation model of radiation-induced AML

Dekkers et al (2011) [8] developed a model that predicts the incidence of radiation-induced AML in mice as a function of dose, from acute radiation exposure, i.e. radiation delivered all at once. This model uses the two-mutational step molecular mechanism as described in Chapter 2, and summarised below.

1. Deletion of the *Sfpi1* gene produces an intermediate cell, denoted by I, which has a growth advantage but is not yet fully malignant.
2. An I cell then receives a point mutation in the remaining copy of the *Sfpi1* gene (on the other copy of chromosome 2), leading to a malignant cell, denoted by M.

The model is built on the idea that radiation is issued instantaneously and causes an initial number of healthy cells to become intermediate cells, dependent on the dose,

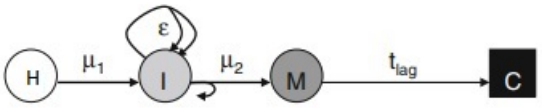


FIGURE 4.1: Two-step molecular mechanism. Healthy cells (H) are transformed into intermediate cells (I), which then transform into malignant cells (M). AML is diagnosed some t_{lag} after the appearance of the first malignant cell.

then over the rest of the mouse's lifetime a background mutation rate can induce a point mutation in the intermediate cells, causing them to become malignant. There is assumed to be a time lag between the appearance of the first malignant cell and the diagnosis of AML, which allows for significant growing of the malignant cell population. The intermediate cells also grow with rate ϵ which is **independent** of the amount of radiation. See figure 4.1 for a diagram of the mechanism used in this model.

The mutation rates take the form:

- Healthy cell to intermediate cell : $\mu_1 = \mu_{0,1}[1 + a_1 D]e^{-p_1 D}$
- Intermediate cell to malignant cell: $\mu_2 = \mu_{0,2}[1 + a_2 D]e^{-p_2 D}$

where $\mu_{0,1}, \mu_{0,2}, a_1, a_2, p_1, p_2 \in \mathbb{R}^+$ are all parameters to be estimated from fitting the model to data, and $D \in \mathbb{R}^+$ represents the dose of radiation. $\mu_{0,1}$ and $\mu_{0,2}$ are background mutation rates of the deletion and the point mutation occurring, respectively. μ_1 is deterministic and describes that the rate of the *Sfpi1* deletion is linear with dose. The exponential function accounts for the chance of cell killing, which increases with higher doses. Since the point mutation is assumed to occur as a spontaneous mutation at some point after radiation exposure, the dose D is then equal to zero so $\mu_2 = \mu_{0,2}$ and is modelled as a stochastic process. The target cells, H, are taken to be the hematopoietic stem cells, with an initial population size of 10,000.

Using Di Majo 1986 data from experiments on CBA/H mice, the optimum parameter values were found with maximum likelihood. The ϵ, μ_0 and t_{lag} parameters were found to have values with maximum likelihood (and 95% confidence intervals) of:

- Growth rate of intermediate cells: $\epsilon = 4 * 10^{-3}$ per day, $(3 * 10^{-3} - 4 * 10^{-3}$ per day)
- Background mutation rate: $\mu_0 \leq 10^{-7}$ per day
- Time lag: t_{lag} = 153 days, (139 - 154 days)

Implementing the model using the *Gillespie algorithm* as defined in chapter 3 gives a dose-response incidence curve seen in figure 4.2. The peak incidence rate is seen between 2 and 3 Gy, and for higher doses the incidence rate drops.

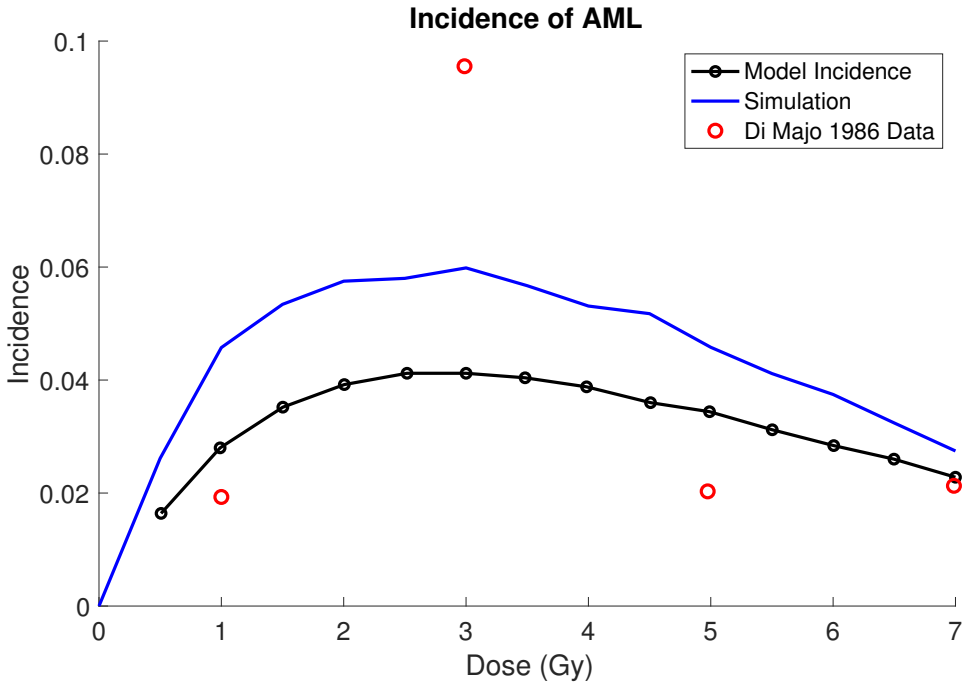


FIGURE 4.2: Incidence of AML as a function of dose. Blue line shows incidence from the simulation using the Gillespie algorithm, black dotted line shows model incidence from the paper and the red dots are data points from Di Majo 1986.

4.2 Dose-Rate (and dose) Dependent Model

In designing a model that can take as input the dose and dose-rate and predict the chance of AML, there are many possible routes to take. A discussion of the biological aspects that could impact the development of AML are described below, and their outcomes are described in the next chapter. For ease of notation, let Id cells denote HSCs that have a deletion of the *Sfpi1* gene, and let D and R denote the dose and dose-rate, respectively, where $D, R \in \mathbb{R}^+$. As a basis of the model, the induction of Id cells is modelled over a period of radiation exposure so that time $T = D/R$ is the time at the end of radiation exposure, then the induction of the point mutation in these Id cells is modelled over the rest of the lifetime, in the absence of radiation. During the radiation exposure period, there is the possibility of radiation-induced cell death.

4.2.1 Assumptions

1. A rate at which the *Sfpi1* deletion occurs per cell is an approximation to the true biology, since this specific deletion is either induced in a cell or not during one time interval. However, there is not enough data to produce a probability distribution of this deletion, therefore as an approximation of the number of Id cells induced in one time interval, one can take the rate of the *Sfpi1* deletion per cell multiplied by the size of the cell population of interest.

2. It is assumed that only the Id cell population has an increased net growth rate, and that H cells have such a low rate of cell division that their growth rate can be taken as zero. Therefore ”healthy” cells can only die due to radiation exposure, and post irradiation the population of ”healthy” cells remains fixed.

3. Since CBA mice show a very low spontaneous rate of AML [22], it is assumed that the induction of an *Sfpi1* deletion in a cell can only occur during radiation exposure, since it is quite a major aberration and would be very unlikely to occur spontaneously.

4. Even though surviving cell populations are usually experimentally measured a few days after radiation exposure, the values used for the surviving fractions in this model can be considered the surviving population immediately after radiation exposure has ended, since it can be assumed that the radiation damage immediately causes the cell to lose its ability to function normally and replicate, even if the cell is still technically alive.

5. Mice are irradiated at 100 days old which is when they are a fully grown adult, and since this is near the beginning of their lifetime it is assumed that they are unlikely to have acquired any spontaneous mutations at this age, i.e. all target cells are free of the *Sfpi1* deletion and the point mutation. On average mice live for 2 years (730 days), and since there is an observed time-lag of 153 days from the appearance of the first malignant cell to AML diagnosis, the model will be simulated for 477 days from the start of radiation exposure.

6. In steady state conditions, healthy HSCs are seen to divide at a very low rate and the cell population size remains at an almost fixed size [23], so for simplification of the model the net growth rate of healthy HSCs is taken to be zero.

4.2.2 Biological aspects that could influence the model

This section details the biological aspects gathered from literature that could influence cell dynamics and activity during the development of radiation-induced AML.

1. Influence on net growth rate of the Id cell population: Experiments [20],[19] show a positive feedback loop between levels of the PU.1 transcription factor and lengthening of the cell cycle in hematopoietic cells, but it is unclear which is the initial cause of this loop. Even so, if little or no PU.1 is present in HSCs, which could be the case if the *Sfpi1* gene is mutated or deleted, the cell is unlikely to maintain a lengthened cell cycle, and these cells would likely have shorter cell cycles than those with fully functioning *Sfpi1* genes. A shorter cell cycle means a higher rate of cell division. It is hard to conclude whether the net growth rate of the Id cell population really increases from "healthy" HSCs, since the effect on the death rate is not known, however experiments [24] do show a general increase in the *percent* of Id cells during the month after irradiation, although for 3 Gy the percent increases and then drops 7 days post irradiation.

   Also, it could be the case that other DNA damage in cells affects the length of the cell cycle. Experiments [25] show a general increase in the percent of HSCs that are cycling after exposure to 3 Gy of radiation, increasing from around 30% in the absence of radiation, to 60%. Overall this means the division rate of HSCs is higher after radiation exposure. This rate of division also increases as time post radiation increases. This could mean that HSCs that didn't receive the *Sfpi1* deletion but still gained some radiation-induced damage maybe also have increased net growth rates.

2. Influence on the radiation-induced death rate: Since the Linear-Quadratic model defined in equation 2.1 describes the surviving fractions of cell populations after radiation absorption, introducing a dose-rate dependency and a time dependency allows the Linear-Quadratic model to be used to describe the time-dependent number of cells dying due to radiation, which changes with dose and dose-rate. As mentioned before, the Linear-Quadratic model suggests that the number of chromosome aberrations N_{CA} formed per cell after a dose D of radiation can be described by the relation 4.1, where $\alpha, \beta \in \mathbb{R}^+$ so that N_{CA} is always non-negative.

$$N_{CA} = \alpha D + \beta D^2 \tag{4.1}$$

   The surviving fraction is stated again as in equation 4.2.

$$SF = e^{-(\alpha D + \beta D^2)} \tag{4.2}$$

   Figure 4.3 shows experimental data of surviving fractions from Matsuya et al (2018) [26], for different dose-rates.

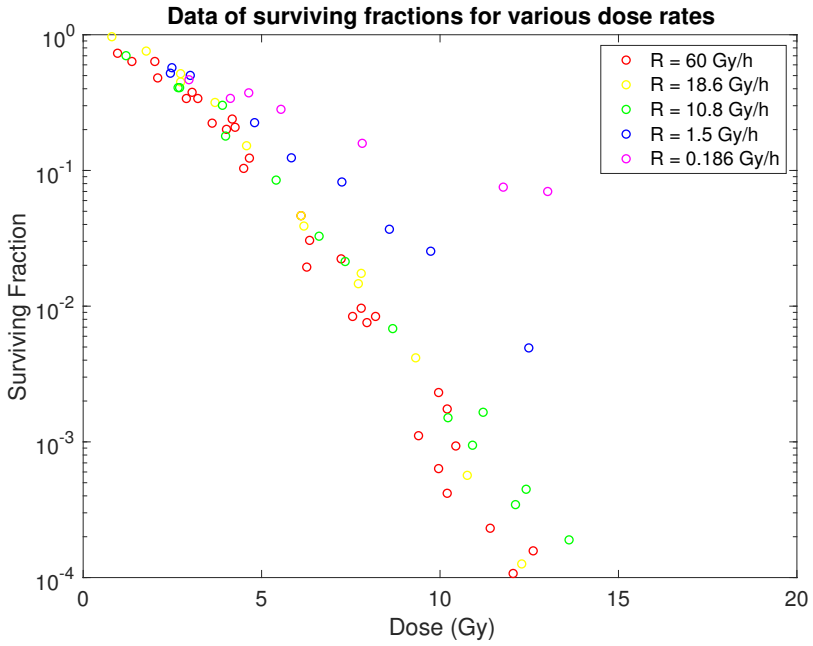


FIGURE 4.3: Surviving Fraction data for different dose rates

Since the quadratic term (D^2) in 4.1 and 4.2 represents the number of chromosome aberrations formed from two ionising tracks, this term is affected by the dose-rate, since when the dose-rate is higher two ionising tracks are more likely to be closer together and a chromosome aberration is more likely to be formed. Therefore β needs to have some dose-rate dependency. Defining β as in 4.3 where $B, C \in \mathbb{R}^+$, introduces a dose-rate dependency so that β eventually saturates with increasing dose-rate.

$$\beta(R) = \frac{B \cdot R}{R + C} \tag{4.3}$$

Additionally, letting $D = R \cdot t$ where $t \in \mathbb{R}^+$ defines the time, gives the Linear-Quadratic model a time dependency so that at time T, the end of radiation exposure, the original form of the Linear-Quadratic model is recovered. Therefore, the number of chromosome aberrations formed per cell during the radiation exposure period, with respect to time, can be defined as in equation 4.4.

$$N_{CA}(t) = \alpha R t + \beta R^2 t^2, \text{ for } t \leq T \tag{4.4}$$

From this, the surviving fraction with respect to time takes the form in equation 4.5.

$$SF(t) = e^{-N_{CA}(t)}, \text{ for } t \leq T \tag{4.5}$$

For time $t > T$, the constant time-independent values for N_{CA} and SF are used, as in 4.1 and 4.2.

3. Influence on the rate of spontaneous point mutations: Much literature [14] [27], talks of a genomic instability which is discussed in chapter 2, however there is no dose or dose-rate specific response known so far as it is very hard to quantify this instability experimentally. It is likely that the more damage induced in cells, the more instability caused and the more cells are prone to errors in DNA replication, meaning a point mutation could be more likely to occur. Also since spontaneous point mutations occur during cell division, an increased cell division rate could lead to a higher rate of the point mutation [25].

4. Influence on the rate of the *Sfpi1* deletion: Literature suggests that the number of chromosome aberrations is proportional to the number of DSBs, [7]. Since the *Sfpi1* deletion is a type of chromosome aberration, the rate at which an Id cell is formed is likely to be proportional to the rate at which chromosome aberrations occur per cell, which is suggested to be linear-quadratic by the Linear-Quadratic model. However, McMahon et al [10] suggest a dose relationship between the rate of large deletions per cell, which is linear for low dose-rates. Peng et al [24] show experimental results of the percentage of Id cells after radiation exposure to a high dose-rate of 0.5-1 Gy/minute. This relationship is almost linear with dose in the dose range 0-2 Gy, showing a slight attenuation as dose increases, and it is likely that the percentage would saturate for higher doses. It is also likely that the percentage would increase with dose-rate, but also saturate. It is not explicitly mentioned in the methodology whether the percentages are measured using only surviving cells or including dead cells as well. However, under the assumption that the measurements are of only the surviving cells, this data could be used to fit the model at high dose-rates.

4.2.3 Structure of the model

Cells will be defined to 3 states. Let H be a healthy HSC, Id be a HSC with the *Sfpi1* deletion, and M be an Id cell that received a *Sfpi1* point mutation: a malignant cell. See figures 4.4 and 4.10 for a diagram of the transition rates between states. From now on let β be defined as in equation 4.3. $H(t)$, $Id(t)$ and $M(t)$ will be used to represent the size of the population of each type of cell at time t, so that $H(t), Id(t), M(t) \in \mathbb{Z}^+$. The initial conditions are taken to be $Id(0) = 0$ by assumption 5, and $H(0) = 10000$ from literature [8].

4.2.4 Radiation Exposure Period

Based on the assumptions above, two rates are to be defined during the period of radiation exposure, as shown in figure 4.4. Firstly, the radiation-induced death rate of HSCs, and secondly the rate of induction of Id cells from healthy HSCs.

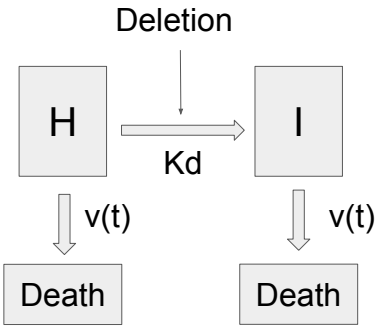


FIGURE 4.4: Cellular states and rates of transition during radiation period, where K_d is the rate of induction of Id cells and $\nu(t)$ is the radiation-induced death rate.

Radiation-induced death rate

Equation 4.5 describes the time-dependent surviving fraction of the original cell population, so assuming no cell growth and death only due to radiation, the radiation-induced death rate can be approximated by taking the derivative of equation 4.5, multiplied by the size of the initial population of healthy cells. Since the derivative of equation 4.5 is negative, defining $\nu(t)$ as in equation 4.7 gives a positive death rate.

$$\dot{SF}(t) = -(\alpha R + 2\beta R^2 t)e^{-(\alpha R t + \beta R^2 t^2)} \tag{4.6}$$

$$\nu(t) = -\dot{SF} \cdot H(0) \tag{4.7}$$

Now let $h(t) = \frac{H(t)}{H(t)+Id(t)}$ and $id(t) = \frac{Id(t)}{H(t)+Id(t)}$ be the fractions of healthy and intermediate cells respectively, out of the total current cell population. Then $\nu(t) \cdot h(t)$ gives the new number of H cells dying at time t, and $\nu(t) \cdot id(t)$ is the new number of Id cells dying at time t.

Using data on surviving fractions dependent on dose and dose-rate as shown in figure 4.3 [26], we can determine best-fit values for the parameters α, B and C, see section on **Bayesian Inference** in chapter 3. The results are displayed fully in chapter 5, but figure 4.5 shows predicted surviving fractions for different dose-rates using the optimum parameter values. These results show that the surviving fractions are very similar for dose-rates of 10 Gy/hour or higher, and also for dose-rates of 0.1 Gy/hour and lower,

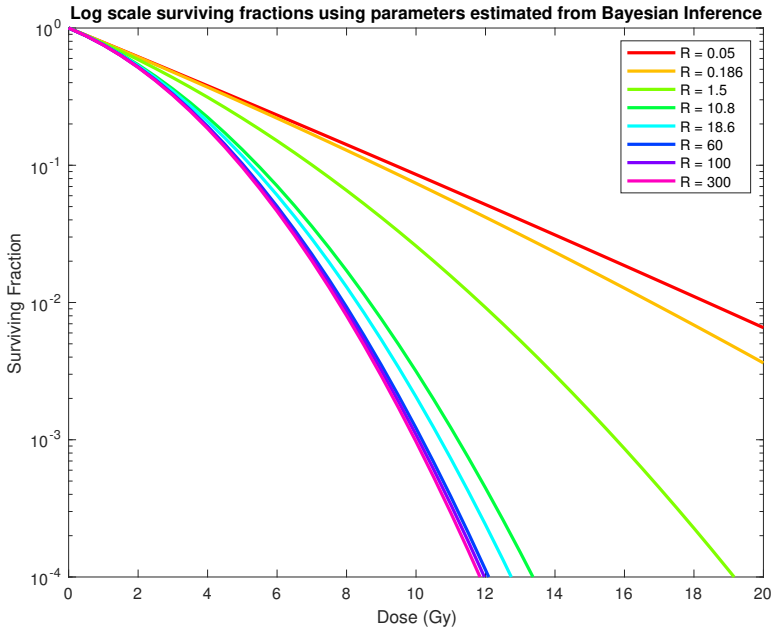


FIGURE 4.5: Surviving Fractions estimated by the Linear-Quadratic model

since at some point increasing the dose-rate does not have an effect on the rate at which chromosome aberrations are formed.

Defining the death rate in this way only gives the numbers of cells dying from the original population size, and therefore does not allow for cell growth during radiation exposure. Using the derivative of the number of chromosome aberrations provides an approximation to the derivative of the surviving fractions of cell populations, and since the rate of chromosome aberrations is defined per cell this allows the death rate to be defined per the current population size, and hence allows for cell growth to occur.

Therefore, the death rate is defined as in equation 4.8, where α and β are the same values determined from the Bayesian Inference, and $\phi \in \mathbb{R}^+$. After radiation exposure, so for time $t > T$, the death rate is set to zero since there are no more radiation-induced chromosome aberrations formed.

$$\begin{aligned} \nu(t) &= \phi \cdot \dot{N_{CA}}(t) \\ &= \phi \cdot (\alpha R + 2\beta R^2 t) \end{aligned} \tag{4.8}$$

Rate of induction of Id cells

To define the rate of induction of Id cells, denoted by K_d, a number of hypotheses are tested to see the effect on the number of Id cells formed.

Hypothesis 1

One plausible form of K_d, see biological aspect 4, is that the rate at which the *Sfpi1* deletion is induced per cell is proportional to the total number of chromosome aberrations formed per cell at the end of radiation exposure, with respect to dose and dose-rate. Taking the derivative of equation 4.4 multiplied by a proportionality constant, $\mu_d \in \mathbb{R}^+$, gives the rate of the *Sfpi1* deletion occurring per cell at time t, see equation 4.9.

$$K_d(t) = \mu_d \cdot (\alpha R + 2\beta R^2 t) \tag{4.9}$$

Multiplying 4.9 by the current size of the healthy cell population, gives the new number of healthy cells gaining the deletion at time t. This leads to the ODE system:

$$\dot{H}(t) = -K_d(t) \cdot H(t) - \nu(t) \cdot H(t) \tag{4.10}$$

$$\dot{Id}(t) = K_d(t) \cdot H(t) - \nu(t) \cdot Id(t) \tag{4.11}$$

Obtaining the solution $Id(T)$ to this system gives the number of cells with the *Sfpi1* deletion at the end of the radiation period.

The absolute numbers and the percent of Id cells obtained using equations 4.7 and 4.9 in the above ODE system are shown in figures 4.6 and 4.7 respectively. In figure 4.7 the black dotted line shows the data from Peng et al [24], in which a dose-rate of 30-60 Gy/hour was used, and μ_d was approximated from this data. These figures show that the peak numbers of Id cells shift to the higher end of the dose range as dose-rate decreases, but also that as the dose-rate increases the number of Id cells drops. The percent of Id cells increases as the dose-rate increases from 0.5 Gy/hour to 30 Gy/hour, then decreases again for higher dose-rates. As more cells are killed for higher dose-rates these results could be biologically reasonable, although the observation of increasing numbers of Id cells for doses above 4 Gy as dose-rate decreases seems unlikely.

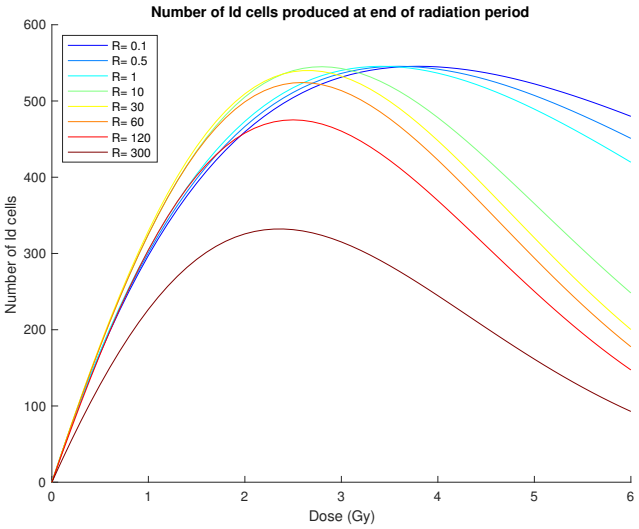


FIGURE 4.6: Number of Id cells present at the end of radiation exposure, for various dose-rates, with K_d as in equation 4.9

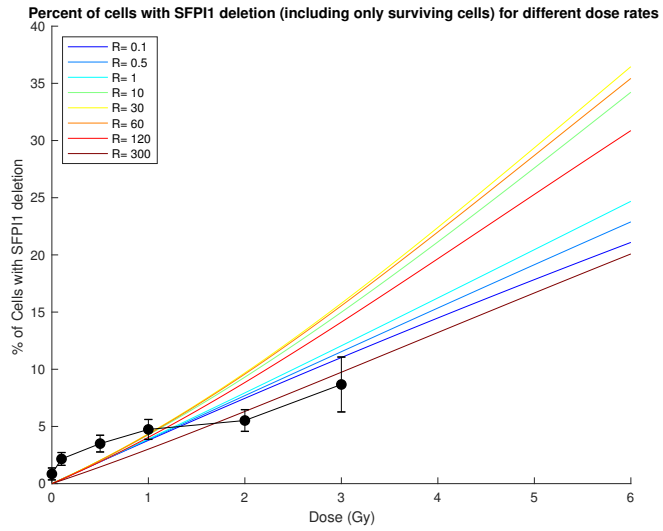


FIGURE 4.7: Percent of Id cells at the end of radiation exposure, for various dose-rates, with K_d as in equation 4.9. Black dotted line shows data from Peng et al [24].

Hypothesis 2 Another possibility is that the rate of induction of Id cells is proportional to the dose-rate at time t, (which is at a constant value R during radiation exposure and zero after), due to the observation of a linear relationship between large deletions and dose [10]. To include a saturation with increasing dose-rate, the term $R/(R+\gamma)$ is used, and to include the small chance of the deletion being caused by every ionising track of radiation, a constant δ is added. The parameter $\mu_d \in \mathbb{R}^+$ is used to scale the rate when fitting the model to data [22]. Equation 4.12 shows the rate of K_d in this case.

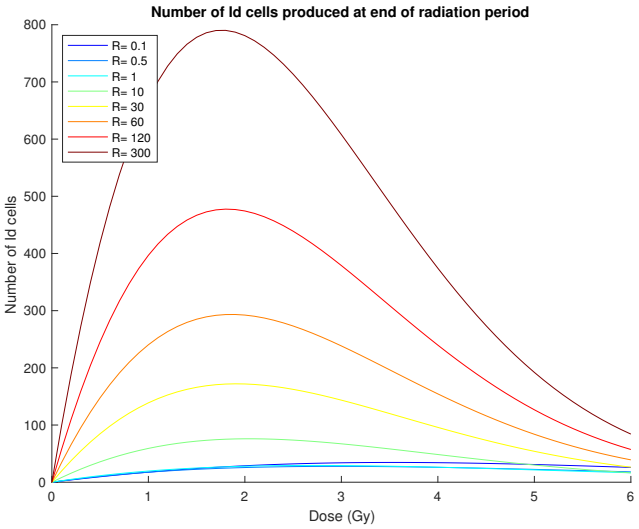


FIGURE 4.8: Number of Id cells present at the end of radiation exposure, for various dose-rates, with K_d as in 4.12

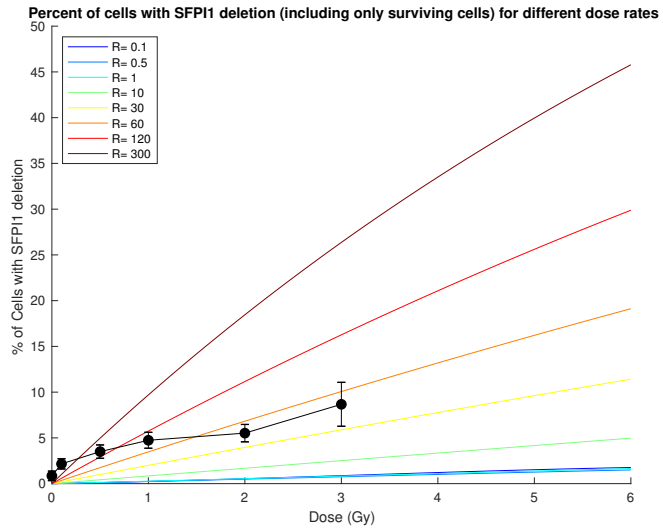


FIGURE 4.9: Percent of Id cells at the end of radiation exposure, for various dose-rates, with K_d as in 4.12,

$$K_d = \mu_d \cdot R \cdot (\delta + \frac{R}{R+\gamma}) \tag{4.12}$$

Figures 4.8 and 4.9 show the absolute numbers and percent of Id cells respectively, obtained when defining the death rate as in equation 4.7 and the rate of induction of Id cells as in equation 4.12.

This case shows absolute numbers of Id cells increasing with dose-rate, and the percent of Id cells also increases with dose-rate. Biologically speaking it seems likely that the absolute number of Id cells would start to drop before the dose-rate reaches as high

as 300 Gy/hour, since the Linear-Quadratic model for cell survival shows a saturation for dose-rates above 10 Gy/hour as shown in figure 4.5, so increasing the dose-rate above 10 Gy/hour would be expected to not have much of an affect on the radiation-induced death rate of cells. However the trends shown in this hypothesis still seem more biologically reasonable than hypothesis 1, hence hypothesis 2 is used in the model.

4.2.5 Post Radiation Exposure Period

The chance of the point mutation occurring in Id cells is modelled until death of the organism, and is allowed to occur during radiation exposure, even though very unlikely since the radiation exposure period is small in comparison to the life time of the mouse. The H cell population is ignored after the end of the radiation exposure period. The cellular transitions are shown in figure 4.10.

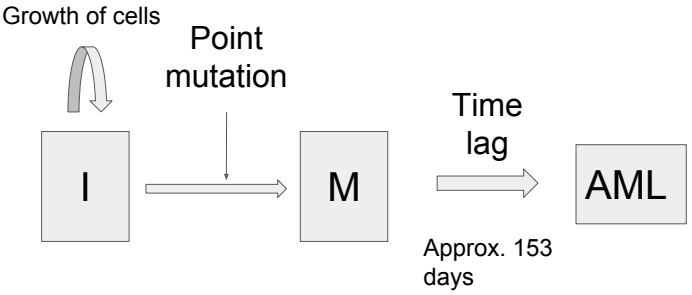


FIGURE 4.10: Cellular states and rates of transitions for post irradiation period, where K_p is the rate of the *Sfpi1* point mutation and the growth rate of Id cells is defined as g_I.

Two rates need to be defined here; the rate of a spontaneous point mutation and the net growth rate of the Id cell population, as discussed in biological aspect 1. Let K_p denote the rate of the *Sfpi1* point mutation occurring per cell. Assuming that the rate of spontaneous point mutations is increased if more damage is induced in cells during radiation exposure [25] [14], and that the number of chromosome aberrations accumulated per cell ($N_{CA}(t)$) during radiation exposure (so for $0 < t \leq T$), can be used as a measure of the amount of damage induced, then K_p can be defined proportional to ($N_{CA}(t)$). After radiation exposure, so for $t > T$, the total number of accumulated chromosome aberrations remains fixed at $N_{CA}(T)$.

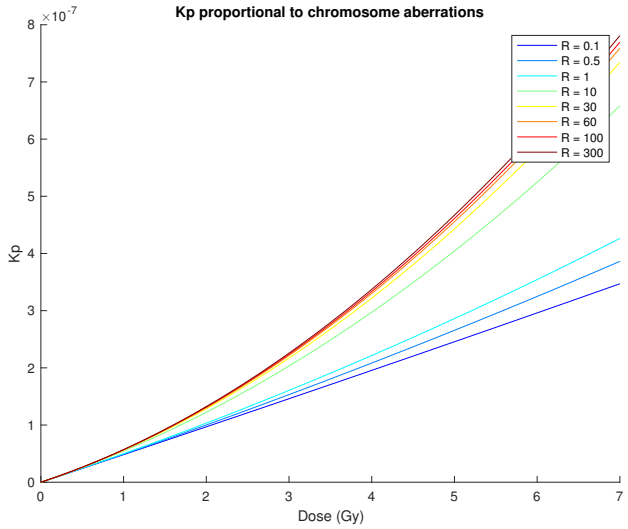


FIGURE 4.11: Rate of point mutation K_p proportional to the number of chromosome aberrations, as defined by equation 4.13

Figure 4.11 shows K_p when taking the form in equation 4.13. It seems biologically reasonable however, that the rate would saturate with increasing number of chromosome aberrations (which increase with increasing dose and dose-rate), since theoretically as $D \to \infty$, $K_p \to \infty$ in equation 4.13. Using the arctan function here provides a way to include this saturation, shown in figure 4.12. Here K_p takes the form in equation 4.14 where $\mu_p, \eta \in \mathbb{R}^+$ are parameters to be estimated using a Simulated Annealing method when fitting the entire model to data [22].

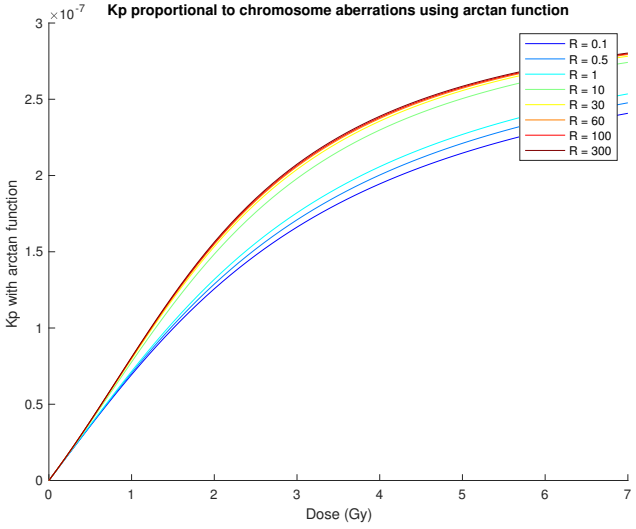


FIGURE 4.12: Rate of point mutation defined by equation 4.14 for different dose-rates.

$$K_p(t) = \mu_p \cdot N_{CA}(t) \tag{4.13}$$

$$K_p(t) = \mu_p \cdot \arctan(\eta \cdot N_{CA}(t)) \tag{4.14}$$

Ban and Kai [25] also suggest that the increase in cell divisions as time post irradiation increases causes the chance of the point mutation to become more likely, but since this would increase K_p the same amount irrespective of dose and dose-rate, it is ignored for simplification of the model.

Net growth rate of Id cells

Let the net growth rate of Id cells be defined as g_I. As discussed in biological aspect 1, it is not clear exactly the effect on this growth rate. It seems certain that the growth rate must be larger than the rate of a spontaneous point mutation, and that since the remaining H cells are ignored after radiation exposure, a logistic growth of Id cells is assumed to ensure the number of Id cells does not grow to a population higher than 10000. Under the assumption that the net growth rate increases with a higher amount of radiation-induced damage, g_I is also taken to be proportional to the number of chromosome aberrations. Hence g_I is defined as in equation 4.15.

$$g_I = \epsilon_I \cdot N_{CA}(t) \cdot (1 - Id(t)/10^4) \tag{4.15}$$

4.2.6 Model Summary

To summarise, the model rates are defined as in 4.16, 4.17 and 4.18, where K_d is defined as in 4.12 and K_p is defined as in 4.14. To clarify, for $t > T$, $\nu(t) = 0$, $K_d = 0$ and $N_{CA}(t) = N_{CA}(T)$.

$$\dot{H} = K_d \cdot H - \nu(t) \cdot H \tag{4.16}$$

$$\dot{Id} = K_d \cdot H - \nu(t) \cdot Id + g_I \cdot Id - K_p \cdot Id \tag{4.17}$$

$$\dot{M} = K_p \cdot Id \tag{4.18}$$

The same value of 153 days is used as in the dose-dependent model [8], for the lag time between the existence of the first malignant cell and the diagnosis of AML. The end time (t_{end}) of the simulation is set to 477 days as described in assumption 5. To simulate the time of AML diagnosis, the *Tau-Leaping* algorithm as described in chapter 3 is used to track the time of existence of the first M cell, defined as $t_{M1} \in \mathbb{R}^+$. If $0 < t_{M1} < t_{end}$, then it is assumed that AML will develop at time $t_{M1} + 153$. In order to estimate the incidence of AML for a given dose and dose-rate, the simulation is run a high number of times to count the number of cases where AML is diagnosed.

A Simulated Annealing method is used to find the optimum values for parameters $\phi, \mu_d, \delta, \gamma, \mu_p, \eta$ and ϵ_I, using data from Major 1979 [22]. See chapter 3 for a description of this method.

Chapter 5

Results

This chapter presents the results from the parameter estimations for the Linear-Quadratic model, the model predicted incidence of AML and recommendations for DDREFs for AML.

5.1 Parameter Estimation of the Linear-Quadratic Model using Bayesian Inference

Using data on surviving fractions in response to dose for different dose-rates [26], the parameter values α, B and C were estimated using a Bayesian Inference method. The package *RStan* was used, a package for R that implements the *No-U-Turn Sampler* algorithm, as described in Chapter 3. Prior distributions of the parameters were set as:

$$\alpha \sim \text{unif}(0, 1) \tag{5.1}$$

$$B \sim \text{unif}(0, 1) \tag{5.2}$$

$$C \sim \text{unif}(0, 10) \tag{5.3}$$

The algorithm was run using a warm up of 1000 iterations and another 20000 iterations for the full sampling. The posterior distributions of the parameters are displayed in figures 5.1, 5.2 and 5.3, and show Guassian shaped distributions, therefore the median values from these distributions are chosen for the parameter values of the model.

Figure 5.4 shows the trace plots of the sampling of the parameter spaces, where theta[1], theta[2] and theta[3] refer to α, B and C respectively. The grey shaded area shows the warm up period.

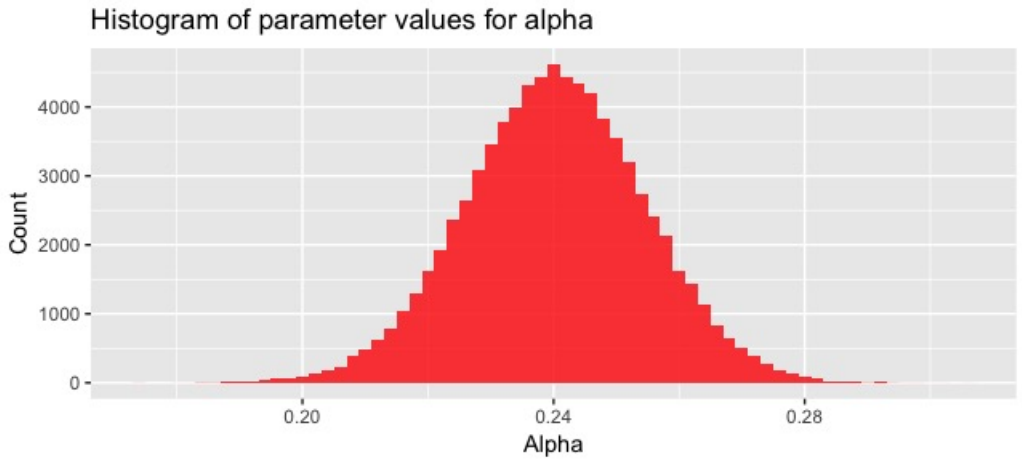


FIGURE 5.1: Posterior Distribution of α

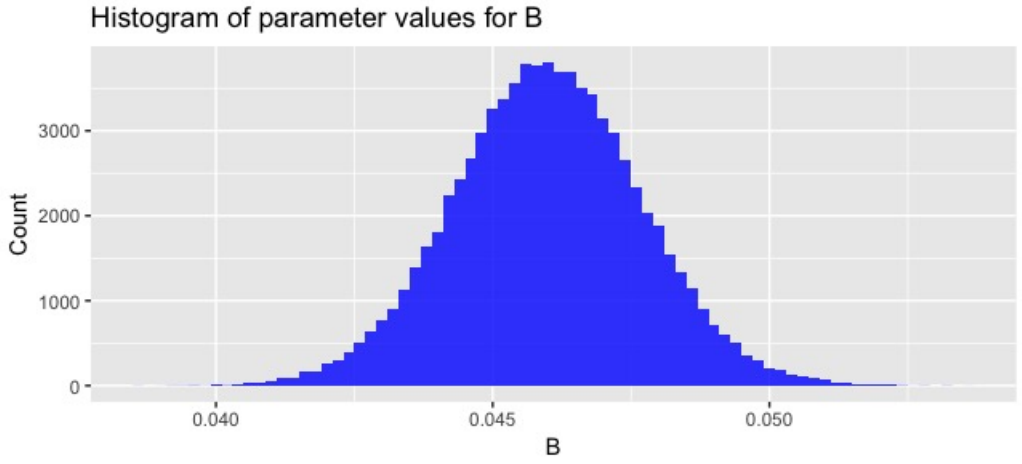


FIGURE 5.2: Posterior Distribution of B

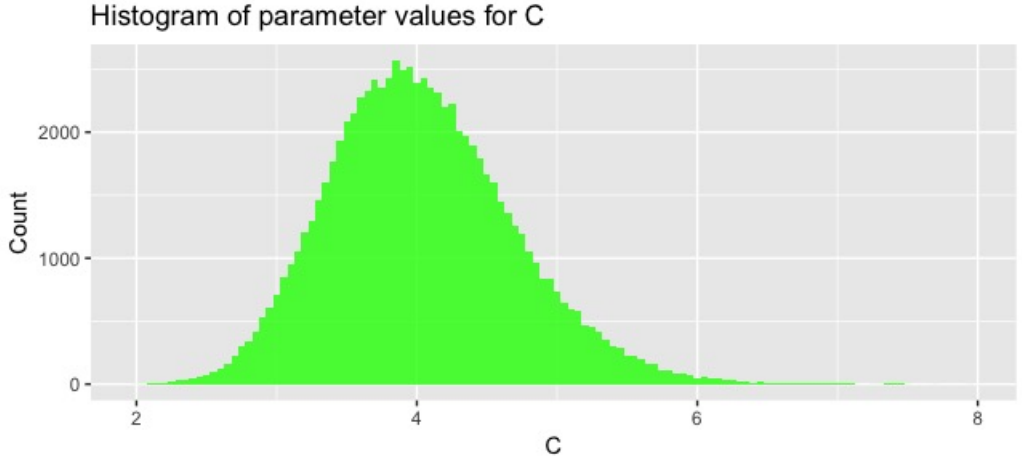


FIGURE 5.3: Posterior Distribution of C

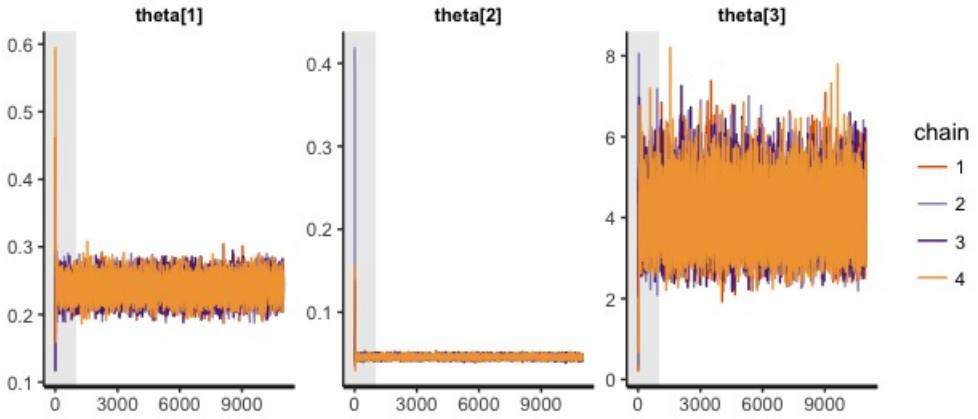
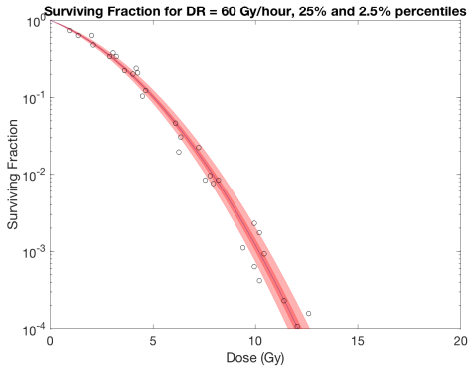
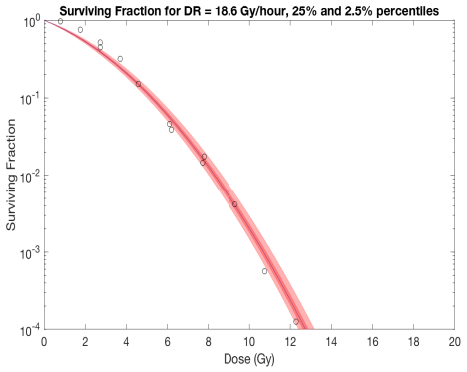


FIGURE 5.4: Trace Plot of Bayesian Inference

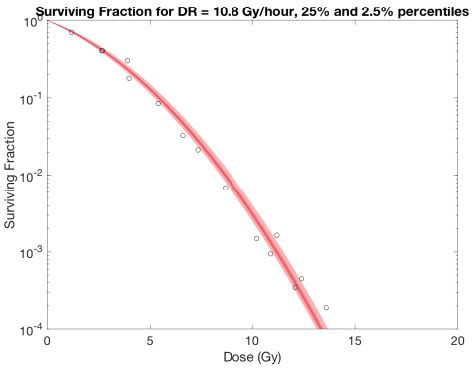
The surviving fraction model predictions are shown in figure 5.5 for different dose-rates, along with the 2.5 % and 25 % percentiles of the parameter posterior distributions.



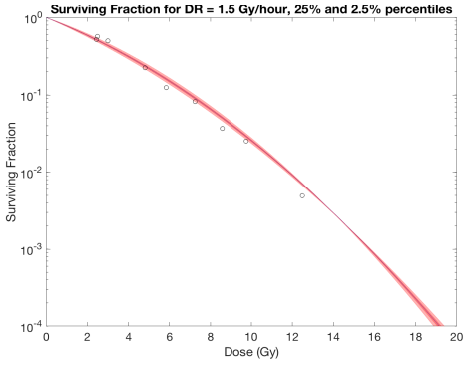
((A)) dose-rate = 60 Gy/hour



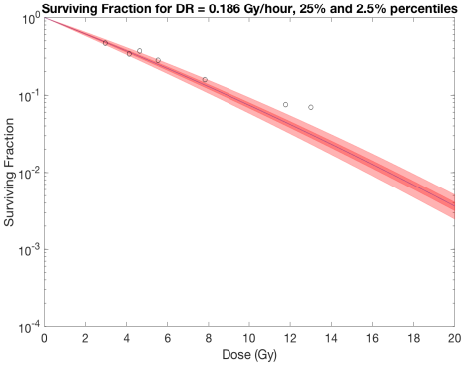
((B)) dose-rate = 18.6 Gy/hour



((C)) dose-rate = 10.8 Gy/hour



((D)) dose-rate = 1.5 Gy/hour



((E)) dose-rate = 0.186 Gy/hour

FIGURE 5.5: Model predictions showing 2.5% and 25% percentiles of parameter posterior distributions for various dose-rates.

Figure 5.6 shows the combined predicted surviving fractions and the 2.5 % and 25 % percentiles.

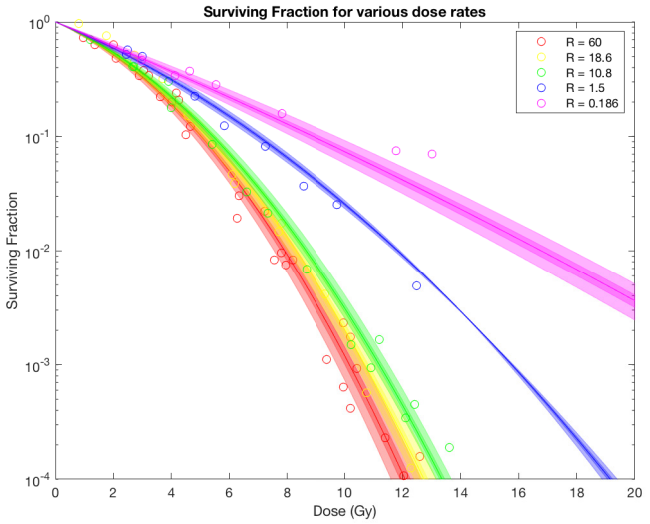


FIGURE 5.6: Surviving fractions estimated by Bayesian Inference for dose-rates from Matsuya data [26], showing the data points as circles, the 25% and 2.5% percentiles by the shaded regions and the model estimates using median parameter values by the dark lines.

5.2 Parameter estimation for model using Simulated Annealing

The results from Simulated Annealing yield the following parameter values for the dose-rate dependent model.

$$\phi = 1.5, \tag{5.4}$$

$$\delta = 0.001, \tag{5.5}$$

$$\gamma = 5, \tag{5.6}$$

$$\mu_d = 5.9, \tag{5.7}$$

$$\mu_p = 1.5e^{-8}, \tag{5.8}$$

$$\epsilon_I = 3.395e^{-5}, \tag{5.9}$$

$$\eta = 2 \tag{5.10}$$

These results make sense biologically; the parameter for the rate of the point mutation occurring (μ_p) in an I_d cell is much lower than the parameter for the rate of net growth of I_d cells (ϵ_I).

5.3 AML Incidence

In this section the predicted incidence of AML from the model is presented. Firstly, some observations from available data are discussed [28][22][29]. For high dose-rates such as 5 Gy/minute and single body exposure, the incidence is seen to peak at a dose of between 2 and 3 Gy with incidence rates between 10% and 30%. The incidence is seen to drop for dose-rates of 0.57 Gy/minute but a peak is still observed between 2 and 3 Gy. At lower dose-rates such as 0.04 mGy/hour, the incidence is seen to be more uniform for different doses, with a lower incidence rate of around 5%.

Figure 5.7 shows the model predictions of AML incidence for dose-rates used in Major 1979, and the data points of Major 1979.

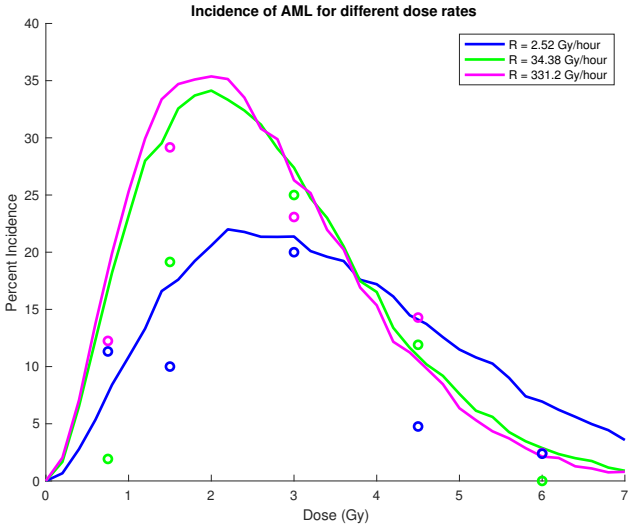


FIGURE 5.7: Model predictions of the incidence of AML, where dose-rates are in Gy/hour, with Major 1979 data plotted as circles. [22].

To see the effect of even lower dose-rates, figure 5.8 shows the predicted incidence of AML from the model for dose-rates down to 0.01 Gy/hour.

The trends from the data are also visible in the model predictions, where the incidence becomes lower for lower dose-rates.

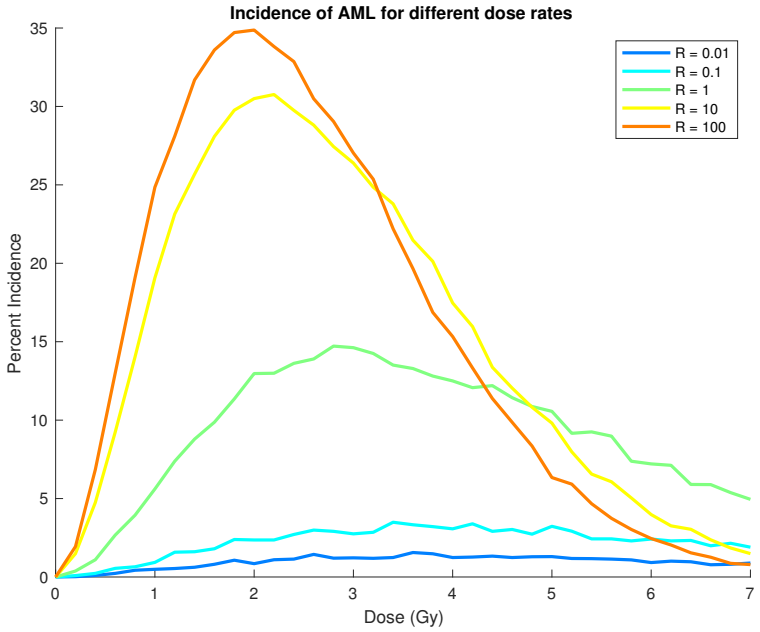


FIGURE 5.8: Model incidence of AML for different dose-rates, where dose-rate is Gy/hour.

5.4 Dose and Dose-Rate Effectiveness Factors

This section provides recommendations on dose and dose-rate effectiveness factors for the incidence of AML. Figure 5.9 shows the dose-rate effectiveness factors for AML, with a dose-rate of 100 Gy/hour as a reference. Focusing on the low dose region of less than 1 Gy, the results show that as the dose-rate is lowered, the effectiveness factor becomes higher. This means that the risk of AML from a dose-rate of 100 Gy/hour should be divided by a higher value as the dose-rate lowers. For a dose-rate of 0.05 Gy/hour, the model suggests that the DDREF could be between 30 and 50 for doses less than 1 Gy, much higher than the current DDREF of 2. However, as the total absorbed dose increases above 1 Gy, the model suggests that the dose-rate has less of an effect on the risk of AML and the dose-rate effectiveness factor lowers significantly. This suggests that much more attention should be given to estimating the risk of low dose radiation.

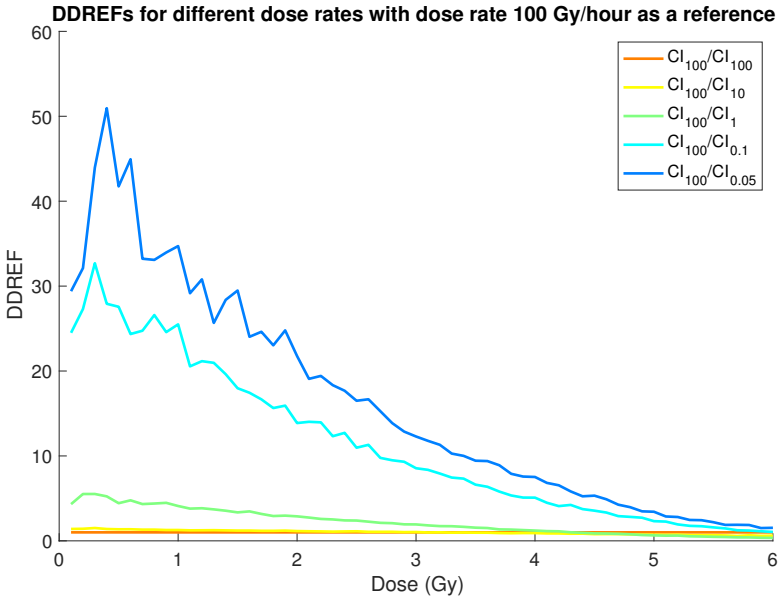


FIGURE 5.9: Dose-rate effectiveness factors with the reference dose-rate as 100 Gy/hour, i.e. CI_{100}/CI_1 denotes the cancer incidence from a dose-rate of 100 Gy/hour divided by the cancer incidence from a dose-rate of 1 Gy/hour.

Chapter 6

Conclusion and Discussion

This chapter will summarise the results found in this thesis, and discuss recommendations on further research.

6.1 Conclusion

In this thesis a dose and dose-rate dependent model for murine Acute Myeloid Leukaemia (AML) has been developed. From this model, risk estimates of radiation-induced AML from low dose and low dose-rate radiation can be deduced. As stated in the introductory chapter, current risk estimates for low dose and low dose-rate radiation use a factor (DDREF) of 2, whereby the risk from high doses at a high dose-rate is divided by this factor. This DDREF is used very generally for all types of radiation damage, and is therefore thought to lead to over or under estimations [7] [5].

The model developed in this thesis shows a non-linear response to dose and non-linear dose-rate effects, where for example given a dose of 3 Gy, the estimated incidence rate at low dose-rates is much less than half of the incidence rate for high dose-rates. This suggests that the DDREF of 2 should be replaced with a dose and dose-rate effectiveness function so that risk from any given dose and dose-rate can be better estimated. The model provides an estimate of the incidence from a given dose and dose-rate, but as a development of this a function could be derived to calculate the dose-rate effectiveness factor for a given dose.

6.2 Recommendations for further research

This section will describe some possible improvements to the model and some other biological observations that could be included which could effect the incidence of AML.

Firstly, adding to the model the chance of an H cell obtaining an *Sfpi1* point mutation first, and then the *Sfpi1* deletion, could have an effect on the dose or dose-rate response to murine AML. This was ignored for simplification of the model but this route could be possible for low dose-rates. Another detail that could be included in the model is the effect of cell cycle sensitivity [9]. Since cells show variations in sensitivity when in different phases of the cell cycle, the phase that the cell is currently in when hit by radiation could affect the chances of the *Sfpi1* deletion or a point mutation occurring.

The Linear-Quadratic model for cell survival predicts that at low doses, the dose response decreases in a linear fashion and the dose-rate effects are negligible. Other literature [30],[31],[32] suggests that there is a hyper radio-sensitivity at doses of less than 1 Gy, as shown in figure 6.1.

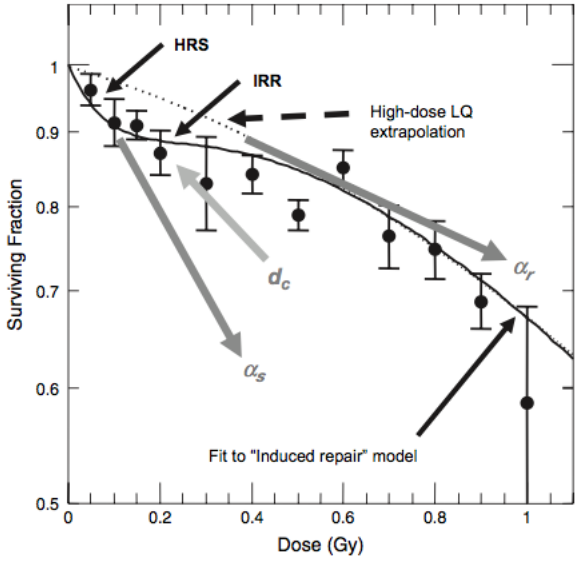


FIGURE 6.1: Hyper radio-sensitivity observed in cell survival for low doses [32] shown by the solid black line, and the dotted line is extrapolation from the Linear-Quadratic model

This observation is also a contradiction to the linear no-threshold model, in which separate exposures to radiation and their risks are added up in order to give a risk estimate for the cumulative dose received. In fact the hyper radio-sensitivity observed here suggests that the risk of two separate exposures adding up to a particular dose could be higher than the risk of receiving the total dose instantaneously. For example, the cell

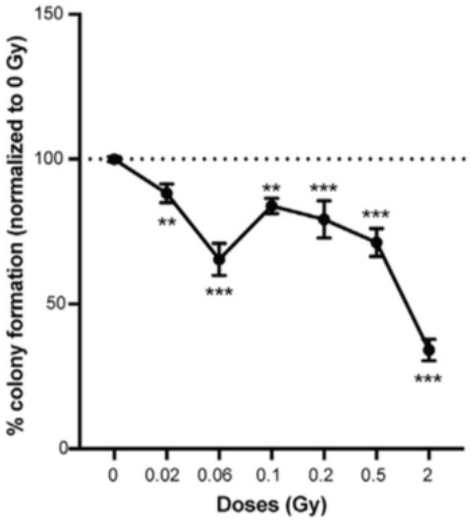


FIGURE 6.2: Hyper radio-sensitivity observed in cell survival for very low doses [30].

damage upon receiving a dose of 0.2 Gy is very similar to the cell damage upon receiving a dose of 0.4 Gy, suggesting that more damage could be done by receiving a dose of 0.4 Gy split over two sessions, than receiving 0.4 Gy all at once.

Figure 6.2 also shows an observed hyper radio-sensitivity at very low doses of less than 0.1 Gy, which is a particularly interesting observation applicable to medical imaging. For example, this observation suggests that receiving one CT scan containing 0.1 Gy would cause much less damage than receiving 2 CT scans containing 0.05 Gy each. Therefore, to improve further the risk estimates of low dose radiation, this hyper radio-sensitivity could be used to better describe the cell survival at low doses.

The data set used [22] contains a small number of data points and there is no indication of the size of possible errors in the experiments, so this is a limitation to the reliability of the parameter estimation of the model. With a larger data set the model parameters could be better estimated. Once more experiments are carried out that monitor the induction of the *Sfpi1* deletion in healthy cells after absorption of radiation at different dose-rates, then there can be some clarification on the dynamics of I_d cell formation. Also experiments tracking the numbers of I_d cells for a period after radiation exposure at different doses and dose-rates would potentially give more insight into the growth rate of these cells, and whether the dose affects the growth rate.

As a development of this model, it would be interesting to look at the effect of radiation delivered at a fractionated dose-rate in order to give recommendations to radiotherapy procedures, as it could be that differences are observed when the same dose is delivered continuously or fractionated, due to possible adaptation of cell repair mechanisms. It

would also be particularly useful and interesting in the future to be able to improve the applicability of this model to the incidence of human AML.

As a last remark and self-reflection, this thesis has taught me some very useful statistical techniques, improved my programming skills greatly and allowed me to learn a lot of interesting cell biology. My interest in the biology meant that it took me a while to narrow down a specific topic of research, and upon doing the process again, I would define a research problem earlier to enable further development of the model.

Bibliography

[1] UNSCEAR 2008 REPORT Vol. I. Sources and effects of ionizing radiation.

[2] Mettler et al (2008). Effective doses in radiology and diagnostic nuclear medicine: A catalog.

[3] M Beyzadeoglu et al (2010). Basic radiation oncology. .

[4] Larry Bodgi et al. Mathematical models of radiation action on living cells: From the target theory to the modern approaches. a historical and critical review. .

[5] Dose and dose-rate effectiveness factor (ddref): Recommendation by german commission on radiological protection (2014).

[6] Maurice Tubiana et al. The linear no-threshold relationship is inconsistent with radiation biologic and experimental data (2009). .

[7] K H Chadwick (2017). *Towards a new dose and dose-rate effectiveness factor (DDREF): Some comments.*

[8] Dekkers et al (2010). *A two-mutation model of radiation-induced acute myeloid leukemia using historical mouse data*. .

[9] E. J. Hall. Radiobiology for radiobiologists. 2012.

[10] Stephen J. McMahon et al (2016). Mechanistic modelling of dna repair and cellular survival following radiation-induced dna damage.

[11] URL http://www.ebbiology.com/cell-cycle.

[12] URL https://dynamic.libretexts.org/print/url=https://chem.libretexts.org/Courses/Eastern_Mennonite_University/EMU_3A_Chemistry_for_the_Life_Sciences_(Cessna)/19_3A_Nucleic_Acids/19.5_3A_Mutations_and_Genetic_Diseases.pdf.

[13] T. Verbiest et el (2018). *Tracking preleukemic cells in vivo to reveal the sequence of molecular events in radiation leukemogenesis*.

[14] J. B. Little. Radiation-induced genomic instability. *International Journal of Radiation Biology*, 1998.

[15] URL https://www.cancer.gov/.

[16] URL https://www.nhs.uk/conditions/acute-myeloid-leukaemia/symptoms/.

[17] Daniel A Pryma. Nuclear medicine: Practical physics, artifacts, and pitfalls.

[18] C. H. Olme (2013). *Live cell detection of chromosome 2 deletion and Sfpi1/PU1 loss in radiation-induced mouse acute myeloid leukaemia*. .

[19] T. Verbiest and C. Badie et al (2015). *PU1 downregulation in murine radiation-induced acute myeloid leukaemia (AML): from molecular mechanism to human AML*.

[20] Hao Yuan Kueh et al (2013). *Positive feedback between PU1 and the cell cycle controls myeloid differentiation*.

[21] H C Turner et al (2015). Effect of dose rate on residual γ-h2ax levels and frequency of micronuclei in x-irradiated mouse lymphocytes.

[22] I. Major (1979). *Induction of Myeloid Leukaemia by Whole-Body Single Exposure of CBA Male Mice To X-rays*.

[23] D Bryder E Manesso, J Teles and C Peterson. Dynamical modelling of haematopoiesis: an integrated view over the system in homeostasis and under perturbation.

[24] Yuanlin Peng et el (2009). *Radiation Leukemogenesis in Mice: Loss of PU1 on Chromosome 2 in CBA and C57BL/6 Mice after Irradiation with 1 GeV/nucleon 56Fe Ions, X Rays or γ Rays. Part I. Experimental Observations*. *BioOne*. URL http://www.bioone.org/doi/full/10.1667/RR1547.1.

[25] N Ban and M Kai (2009). Implication of replicative stress related stem cell ageing.

[26] et al (2018) Yusuke Matsuya, Stephen McMahon. Investigation of dose-rate effects and cell-cycle distribution under protracted exposure to ionizing radiation for various dose-rates.

[27] Yixin Yao and Wei Dai (2014). Genomic instability and cancer.

[28] Amleto Castellani (2013). Epidemiology and quantitation of environmental risk in humans from radiation and other agents. Volume 96 of Nato ASI Subseries A, .

[29] C. D. R. Dunn (2000) C. J. Pallister. *Progress in Haematology*. 2.

[30] Rodrigues-Moreira et al. Low-dose irradiation promotes persistent oxidative stress and decreases self-renewal in hematopoietic stem cells (2017). .

[31] Oluwole Olobatuyi et al. Effects of g2-checkpoint dynamics on low-dose hyper-radiosensitivity (2018). .

[32] Brian Marples and Spencer J Collis. Low-dose hyper-radiosensitivity: Past, present, and future (2007).