Math Modeling Final Project

Ché, Jacob, Sam, Tynan

March 18, 2022

1 Introduction

Radiotherapy is a useful treatment method for curing certain sizes and shapes of cancerous tumors. Approximately 50% of cancer patients receive radiation therapy at some point in their treatment, and radiation therapy ultimately contributes to around 40% of curative treatments [1]. Developing a strategy to determine how much of a tumor is likely to survive after a given radiotherapy treatment configuration is an important first step to predicting patient outcomes and optimizing treatments. In this paper, we develop a mathematical and computational strategy to predict both the amount of tumor that survives and the amount of healthy tissue that is damaged after a multiple beam radiotherapy treatment.

Our strategy starts with several simplifying assumptions that could be loosened in future work. We will be considering a 2D cross-section of tumor and healthy tissue restricted to a 51×51 pixel region. A basic beam that is depositing radiation within a region of water was provided as a starting point. We will be assuming that the beams can only be pointed toward the center of the region. Any necessary translation of the tumor can be accomplished by moving the tumor with respect to the beams. We are assuming a relatively low dose of radiation, so that the cell response to radiation is primarily stochastic.

An outline of our strategy is as follows. First, we generalize the given beam distribution by transforming the beam distribution with a rotation matrix. Second, we add up rotated beams to get a higher radiation deposition in the center region. Third, we assume that the whole region is a tumor to derive a 51×51 element probability matrix T_{ij} with each cell's exposure to the radiation of some given beam configuration. Fourth, we introduce a 'body' matrix B_{ij} that allows for tumor cell and healthy tissue damage to be predicted at the same time. Fifth, we find the proportion of tumor that we would expect to survive a given beam configuration applied to two different tumor shapes.

2 Beam Deposition Data

We are given data as an example for how one laser, centered above the area with the tumor, would disperse its energy. The given data is show in figure 1.

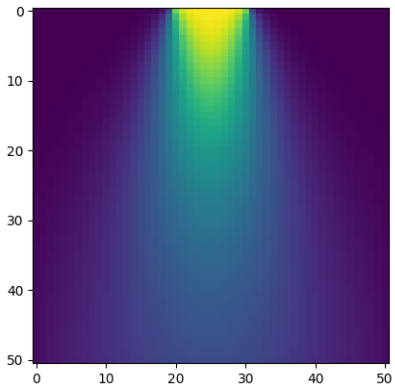


Figure 1: Top-down view of a laser centered above the area with the tumor. Brighter color indicates higher intensity.

We can tweak some parameters of this to make the laser more or less focused, or make the resolution higher. With this, we can think about rotating the picture around for a set of angles and then adding all of the results together to get the total radiation deposited on the area. Next we look at rotating and adding copies of this beam to each other.

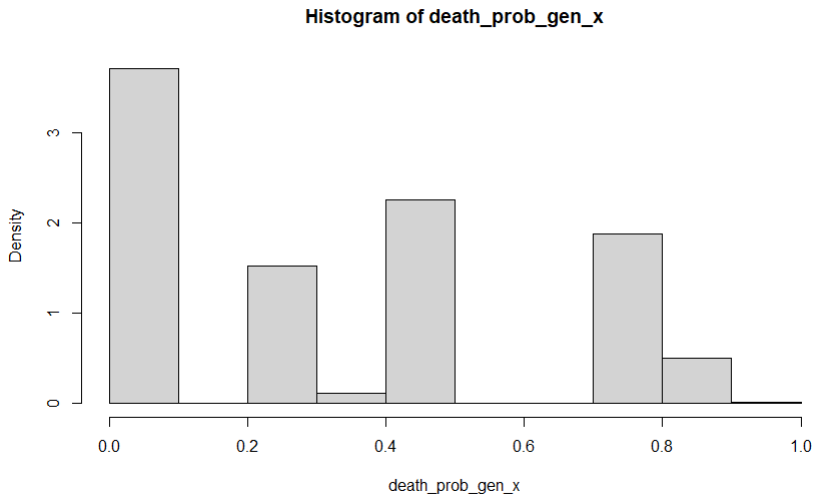


Figure 2: The density of values from figure 1

The histogram in figure 2 shows that the energy dispersed throughout the tumor is fairly spread out, and seems almost uniform, but it is hard to say [1]. "death_prob_gen_x" is the name of the vector used in R to find this histogram and represents the energy dispersed in the tumor.

3 Rotating the Beam

To rotate the data, we can use the 2D rotation matrix

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

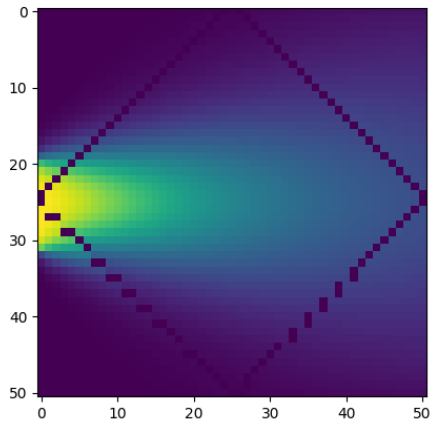
If we center the data so that $(0, 0)$ is at the center of the image (instead of at the top right), then we can rotate the center of each square by any angle θ and then round to the nearest integer. This was accomplished in code. First, we assign each value in the data a coordinate and shift the coordinates so that the origin is at the center of the data. Then, for each coordinate, apply the rotation matrix to get a new coordinate. Finally, shift this new coordinate back to the original data's coordinates (where the origin is at the top right) and round it to the nearest integer. There's some subtlety in the process which has been left out, but the result is *fairly* good, aside from some strange rounding errors causing certain squares to be practically ignored. Additionally, we can see that when rotating a square, the corners get cut off and the sides move inward.

Finally, we can add together different sets of data corresponding to different angles to get a full picture of the dose of radiation. As an example, in figure 4 we add together a set of evenly spaced 30° rotations and see that the area with the most concentrated total dose is in the center.
Now the question is, given a set of rotations, when we have decided approximately how the radiation is spread, how do we determine the probability that a specific tumor cell dies?

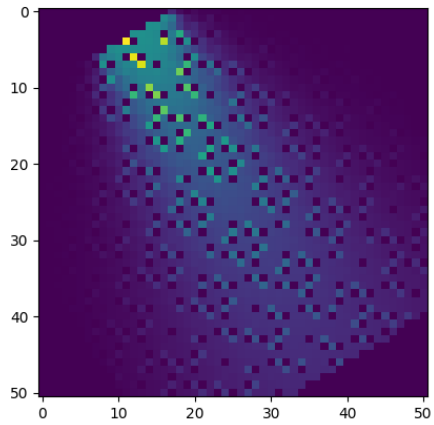
4 Probability of Tumor Damage

We want to consider how likely the tumor cells exposed to the radiation are to be damaged. In order to do this, we need to start by determining at what dose damage is likely to be done. We will use the maximum value from the given data on a single beam (figure 1) and call this maximum value m. This is a good value to use for our simulations because it is something that a radiologist can feasibly control. Also, the actual value does not play a significant role in how the calculations happen since we

[1] Looking at a Cullen-Frey graph suggests this histogram does indeed most closely follow a uniform distribution.



(a) A 90 degree rotation applied to the original data in figure 1



(b) A 30 degree rotation applied to the original data in figure 1

Figure 3

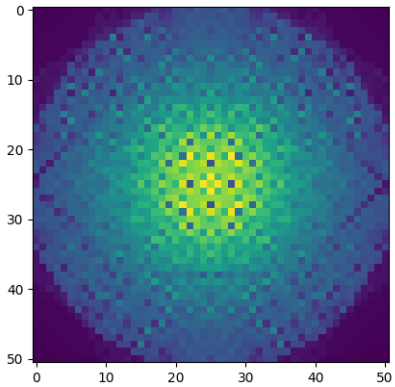


Figure 4: Twelve 30° rotations added together.

will be normalizing the data.

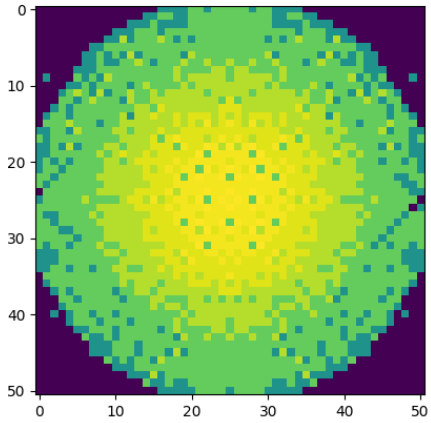
Now that we have our benchmark value for when a tumor cell is likely to be damaged (m), we will use it to normalize the data for a beam being rotated around the center of the tumor as shown in figure 4. Each data point can be divided by our benchmark value to bring the scale of the data to the scale of counting numbers, which is beneficial for determining how many times a cell received the benchmark dosage. So for each position T_{ij}, we normalize the beam intensity value at that position to be $\frac{T_{ij}}{m}$. For example, a value of 6 indicates that the cell received 6 times the benchmark dose. We will assume that each time a cell receives the benchmark dose, it has a probability, p, of being damaged. This is a linear response to increases in dosage, which is a reasonable assumption for relatively low doses as shown in Maqbool's Introduction to Medical Physics [2]. This leads us to consider the survival of a tumor cell as following a geometric distribution.

Define X to be the number of times a cell has received a damaging dose of radiation. Since X follows a geometric distribution, we need to adjust our data so that it is discrete. We can easily do this by rounding each data point to the nearest integer. We have:

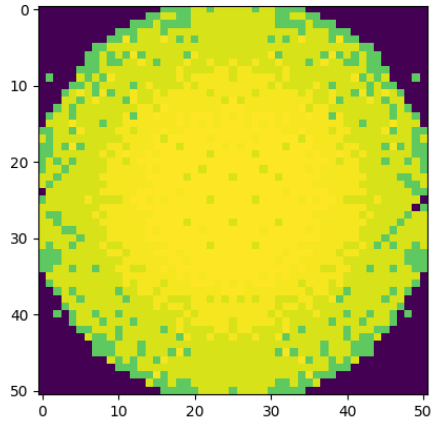
$$P(X = x) = (1 - p)^{x-1} p,$$
$$P(X \leq x) = 1 - (1 - p)^{x},$$

where p is the probability that a tumor cell is damaged from a benchmark dose of radiation, and x is how many times a cell

received that benchmark dosage. In order for this distribution to hold, we assume that once a tumor cell is damaged, it cannot be further damaged. In other words, we are concerned with whether a tumor will or won't be damaged, not the level of damage.



(a) $p = 0.5$



(b) $p = 0.75$

Figure 5: Probability of tumor cells being damaged according to the rotation defined in figure 4

It is worth noting that the strange discontinuities arise from limitations in rotating the single beam about the tumor in a limited resolution as well as the rounding the data to allow for a discrete probability distribution. Also, the value of p chosen would need to be refined with more domain knowledge or having access to data. With more data, we could also consider better options for the probability distribution.

Since we know the probability for each cell being damaged, we want to figure out how much of the tumor we expect to survive. To accomplish this, we will find the expected number of cells in the region that will not be damaged. Define Y_{ij} to be an indicator random variable for the survival of a cell located at T_{ij} and $Y = \sum_{i=1}^{N} \sum_{j=1}^{N} Y_{ij}$. Then we have that

$$Y_{ij} = \begin{cases} 1 & \text{if cell survives} \\ 0 & \text{cell is damaged} \end{cases}.$$

Since Y_{ij} is an indicator random variable, its expected value is the probability of that outcome. So we have the expected amount of the tumor to survive the radiation is

$$\begin{aligned} E[Y] &= E\left[\sum_{i=1}^{N}\sum_{j=1}^{N} Y_{ij}\right] \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} P(Y_{ij} = 1) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} \left(1 - T_{ij}\right). \end{aligned}$$

We have the expected number of tumor cells that survive the radiation, however we would like to express this as a proportion of the entire tumor, so we take $E[Y]$ and divide by the total area of the tumor. From our examples in figure 5, when $p = 0.5$ we expect 34.15% of the tumor is expected to survive and when $p = 0.75$ we expect 21.47% to survive.

Now, we are going to provide our own justification for the use of the Geometric Distribution. First, there are three assumptions that need to be met in order for the Geometric Distribution to be used appropriately: each event is independent, it is either a success or failure, and the probability, p, is the same for every trial. Clearly, all of these are met, so we now will show mathematical justification.

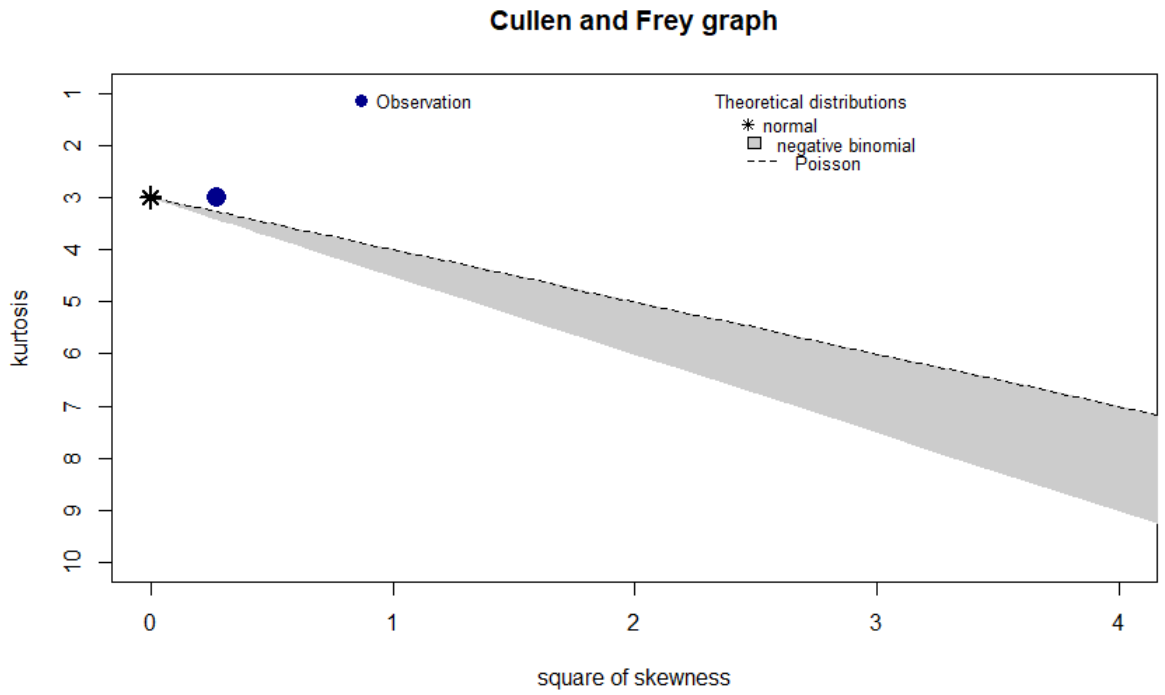
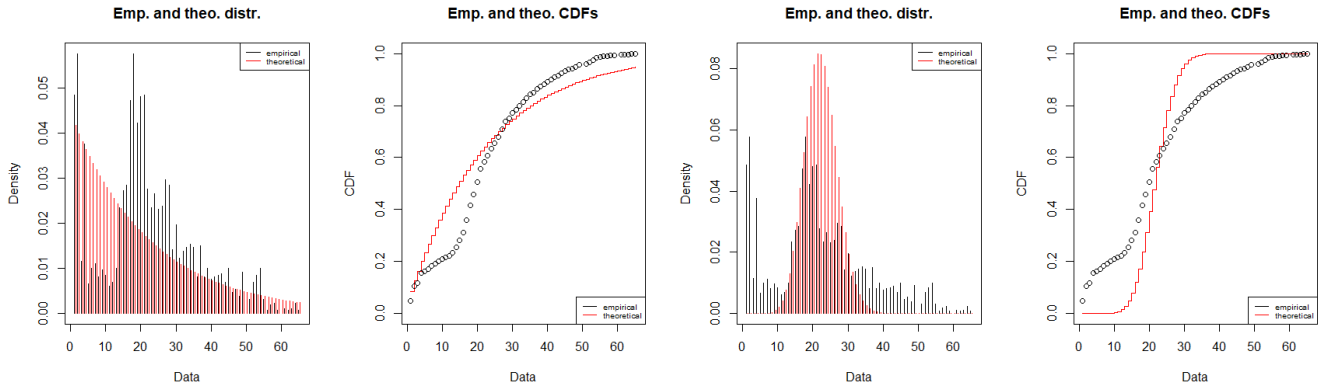


Figure 6: The Cullen-Frey Graph using data used in figure 4

This is a graph known by the Cullen and Frey Graph or the Pearson graph. It takes the data that is provided and compares two qualities of the data to that of different distributions. Kurtosis is a measure of how heavy-tailed the data is, or how frequency has changed as opposed to the frequency at the mean. Skewness is a measure of how asymmetric the data is. The blue dot represents the data that was given [1]. We see that it is fairly close to the line that indicates a Poisson distribution. The geometric distribution is not shown here.

[1] In this case, the same data used in figure 4



(a) Geometric

(b) Poisson

Figure 7

In figure 7, we see theoretical and empirical distributions and cumulative density functions. The goal is to have both plots, theoretical and empirical, match. It seems that the general shape is better using the Poisson distribution, but the CDF matches better with the geometric distribution. However, using a Poisson distribution would not be appropriate in this scenario because Poisson distributions typically describe the time elapsed between events, which does not make sense in this context. Thus, we assume the data fits a Geometric Distribution.

5 Application

Thus far, we have been treating the model as if only the tumor exists. This is not typically the case. Sometimes tumors appear next to or can grow around a vital organ. In these situations, it's important that we limit the organ's radiation exposure. This will be addressed by limiting the rotation of the beam. We want ensure that our model allows us to limit exposure to healthy cells while still ensuring a high probability of damaging tumor cells. There are two main cases that we want to consider: a tumor (or group of tumors) exists around an organ, and a tumor exists adjacent to an organ.

To introduce both tumor and healthy cells, we first need to express the output of the tumor damage probability as a matrix, T_{ij}, with matrix elements i and j. We next introduce a "body" matrix comprised of the sum of a tumor matrix and a healthy tissue matrix, $B_{ij} = \mathcal{T}_{ij} + \mathcal{H}_{ij}$, with the same matrix elements. Assuming that tumor cells and healthy tissue diffuse radiation equally, we only need to consider that healthy cells are less damaged than tumor cells. So, the body matrix will have 1's at the location of the tumor and some value, r, that corresponds to the healthy cell's increased 'resistance' to radiation damage. For locations without tissue, the body matrix will have 0's, however we do assume that there is still water or some non-reactive tissue surrounding the organ to preserve the original beam diffusion. A small example body matrix with $r = .5$ and the tumor located in the middle of an organ is given below:

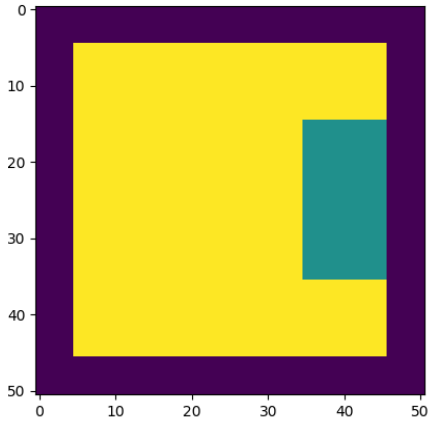
$$B_{ij} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & .5 & .5 & .5 & 0 \\ 0 & .5 & 0 & .5 & 0 \\ 0 & .5 & .5 & .5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \mathcal{T}_{ij} + \mathcal{H}_{ij}$$

If this mathematical model were applied in medical practice, this matrix could be found by CT-scan, X-ray, or another imaging process. The different tissues in that image would need to be identified manually and assigned respective r values.

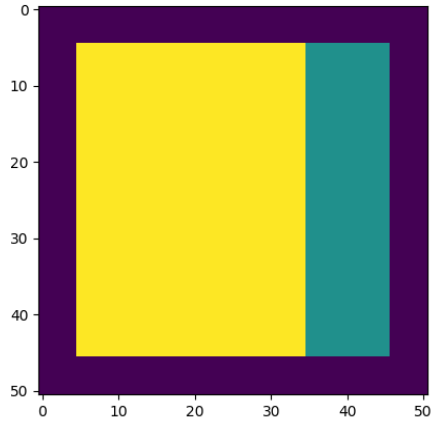
Then, to get a new matrix that measures the probability of damage to both tumor and healthy cells we do an element-wise Hadamard product between T_{ij} and B_{ij} which is defined as follows where i and j span to the maximum row and column indices:

$$T_{ij} \circ B_{ij} = \begin{bmatrix} T_{11}B_{11} & \dots & T_{1j}B_{1j} \\ \vdots & \ddots & \vdots \\ T_{i1}B_{i1} & \dots & T_{ij}B_{ij} \end{bmatrix}$$

We will consider two cases. First, a case in which a tumor is partially surrounding an organ. Second, where a tumor and an organ are adjacent to each other. The body matrices for these cases are shown in figure 8a and figure 8b. We chose to include an outside boundary of non-reactive tissue to minimize the error introduced in the corners of the rotated beam matrix. An additional justification to this boundary is that organs and tumors exist within a body and would not be exposed to the full power of the beam at the edges.



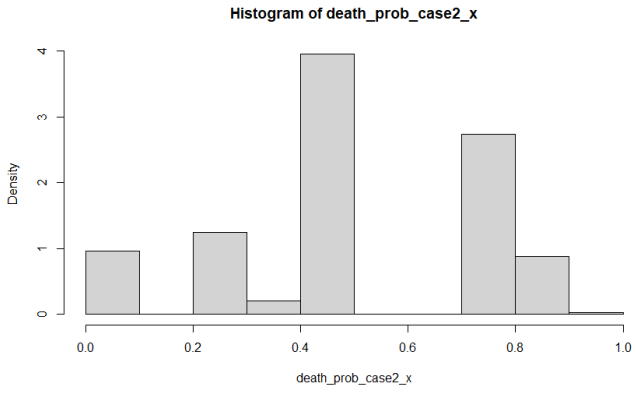
(a) Case 1: Tumor surrounding an organ.



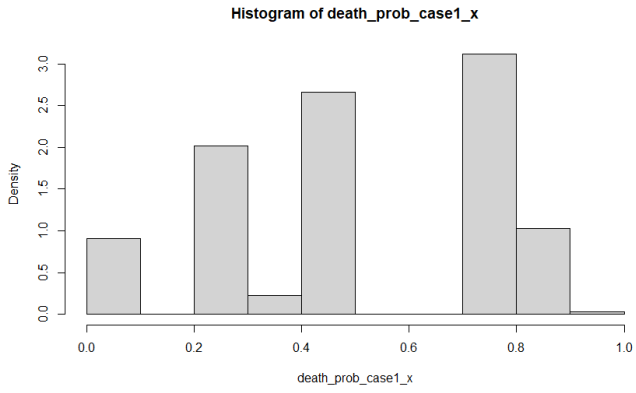
(b) Case 2: Tumor adjacent to an organ.

Figure 8: Full body matrices represented as colored plots for two cases of organ growth. The tumor is yellow, healthy tissue is green, and non-reactive tissue is purple.

For these cases, we consider rotating the beam counter-clockwise from the top of the region in 30° intervals around 180° using the procedure in section 3. Then, we use the procedure in section 4 to generate a T_{ij} matrix to then use the Hadamard product on the $\mathcal{T}$ and $\mathcal{H}$ pieces of the body matrices. The rotated beam and new T_{ij} matrix are shown in figure 10a and figure 10b. To create our T_{ij} matrix, we used $p = 0.5$ as in figure 5a.

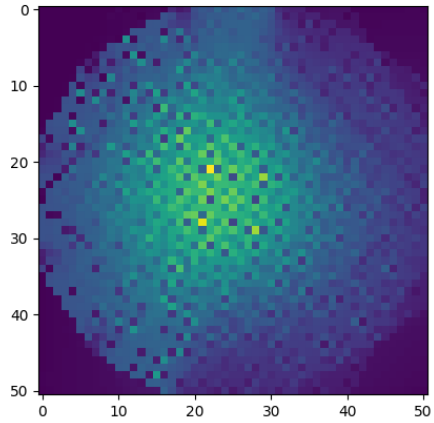


(a) Histogram of energy dispersed in a tumor surrounding an organ.

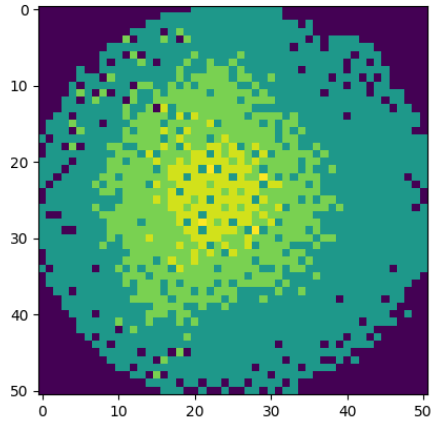


(b) Histogram of energy dispersed in a tumor adjacent to an organ.

These are histograms that describe the density of values for the energy dispersed throughout the tumor for their respective colored plots above. These histograms are interesting because we see a consistent gap in energy dispersed in the values around 0.6.



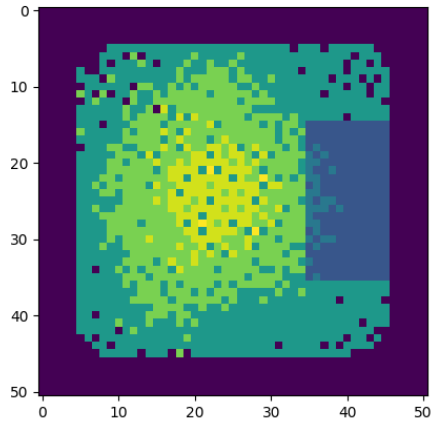
(a) Beam rotation in 30° intervals around 180°



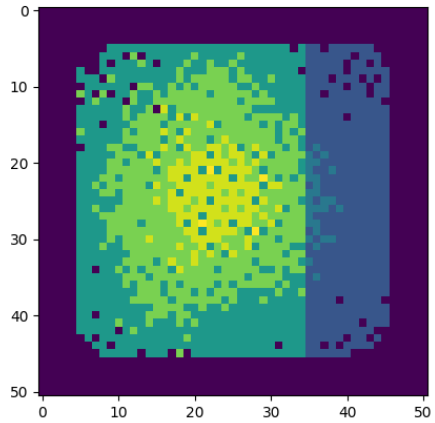
(b) Corresponding T_{ij} matrix

Figure 10: Beam matrix and corresponding T_{ij} matrix that will be applied to the elements of the body matrix.

With this probability matrix, we can apply the beam configuration to any body configuration that we desire. Application on the body matrices of Case 1 and Case 2 are shown in figure 11a and figure 11b. Qualitatively, the tumor cells in these matrices appear to experience more of the radiation deposition compared to the healthy tissue because there is some deposition or probability of damage bias toward the leftward side. The next step in our process is to quantify how much of the tumor we would expect to survive in this configuration using the method described at the end of section 4. Because we have the capability in our process, we will also quantify how much healthy tissue damage occurs.



(a) Case 1: Probability of damage to body tissue.



(b) Case 2: Probability of damage to body tissue.

Figure 11: Beam matrix and corresponding T_{ij} matrix that will be applied to the elements of the body matrix.

Going through the process outlined in section 4 to find the expected number of tumor cells damaged by the radiation, we get that for Case 1, where the tumor partially surrounds organ tissue, we expect 40.17% of the tumor to survive and 73.25% of the healthy organ tissue to survive. In Case 2, where the tumor is adjacent to an organ, we expect 37.24% of the tumor to survive and 75.47% of the healthy organ tissue to survive.

6 Limitations

First, we acknowledge the rudimentary rotation method of the given data. As seen in figure 3, this method loses a fair amount of detail, especially for rotation angles which are not multiples of 90 degrees. This compounds when we do multiple rotations and add them together, and so the picture we have been working with is not as accurate as we may hope. This could be fixed in a couple ways: first, instead of rounding to the nearest integer after rotating, we could spread out the original cell's dose of radiation to all of the cells nearby to where it lands after the rotation. This would make the picture much smoother and may be more accurate to how rotating the beam would actually spread the radiation. Second, we could somehow implement a way to simulate a change of angle in the original beam code which was provided. This is infeasible for us, since the main calculation function in the original beam code is treated as a black box.

Also, in the probability section, we assume that we can use a geometric distribution to determine the probability of tumor cells being damaged. With more domain knowledge or access a data set, we would be able to consider other probability distributions rather than making a somewhat arbitrary choice. It could be that real data shows that we can use a continuous distribution rather than a discrete distribution which would remove more rounding issues in our model. In addition, without access to data, we have no way of knowing what the distribution parameters (in our case, p) should be.

7 Conclusion

In this paper, we set out to create a model that would predict how a beam of radiation would interact with a tumor in several scenarios, like when the entire region is a tumor, or when it is adjacent to an organ or surrounding an organ. After investigating how the beam acted in water, we created simulations of the beam being rotated. Then we went into the statistics and probability involved in deciding whether a tumor cell will die depending on how exposed it was to the beam's radiation, while finding and justifying that this follows a geometric distribution. Using all of this data and our statistical process, we determined how likely certain cells were to survive given a set of beam rotation angles.

References

[1] Rajamanickam Baskar, Kuo Ann Lee, Richard Yeo, and Kheng-Wei Yeoh. Cancer and radiation therapy: Current advances and future directions. *Int J Med Sci*, 9:193–199, 2012.

[2] Wazir Muhammad, Amjad Hussain, and Muhammad Maqbool. *Basic Concepts in Radiation Dosimetry*, pages 9–41. 11 2017.

Python code for finding the expected tumor survival (written by Jacob).
Python code for rotating a beam (written by Tynan, with some contributions from Jacob).
Python code for diffusion of a single beam in water (provided by Dr. Barnard).