College of the Holy Cross

Department of Mathematics and Computer Science

A Topological Analysis of Targeted In-111 Uptake in SPECT Images of Murine Tumors

Author:
Melissa R. McGuirl

Advisor:
David B. Damiano

Reader:
Edward J. Soares

April 15, 2015

Abstract

A method of applying computational topology to quantify heterogenous uptake behavior across a time series of single-photon emission computed tomography (SPECT) images of soft tissue murine tumors has been developed. This behavior cannot be captured by aggregate measures such as injected dose per gram (%ID/g). **Methods:** This method converts a time series of raw data and ROI images into binary tree representations and persistence diagrams. The binary trees and persistence diagrams quantify uptake behavior in neighborhoods of local maxima of intensity throughout a tumor image. This method explores the effects of heterogeneity of uptake on the trends in persistence by separately analyzing persistence points corresponding to high maxima and persistence points corresponding to low maxima. The underlying mathematical construction of this approach is of the zeroth dimension persistent homology. This method has been applied to SPECT tumor images in inviCRO Study 1881-040 In-111, which analyzes differences in aggregate uptake behavior between four mice injected with radioactive tracer In-111 targeted antibody and four mice injected with radioactive tracer In-111 control antibody. **Results:** Statistical analyses indicate that this method effectively distinguishes localized uptake behavior in patients injected with In-111 target antibody from those injected with In-111 control antibody. Separating persistence points into high and low groups shows that there exists heterogenous uptake behavior in patients injected with In-111 control antibody, while uptake behavior in patients injected with In-111 target antibody appears homogenous. **Discussion:** There is statistical evidence to suggest that this topological approach captures heterogeneity in the binding and uptake of control antibody in tumors in mice injected with In-111 control antibody. This behavior is not present in the binding and uptake of targeted antibody in tumors injected with In-111 target antibody.

Acknowledgements

Contents

1 Introduction

Topology is an area of mathematics that studies the structure and connectedness of objects. Properties of topological spaces can be quantified by analyzing the number of connected components of a space in various dimensions. Specifically, persistent homology is a method of detecting and quantifying topological information through a filtration of a domain. In dimension 0, persistent homology analyzes the connected components in a space as one descends through a filtration of its domain.

There are many applications of computational topology in biology, image analysis, and data analysis. For example, in [3] computational topology techniques were successfully applied to classify liver lesions and in [4] discrete topology was used to extract vessel trees from 3-D images. In this thesis, persistent homology is utilized to analyze the effects of tumor targeting antibodies.

Targeted drug therapies are designed to increase amount of treatment concentrated in tumor while minimizing the amount of treatment concentrated in healthy tissue and organs. Antibodies, a vehicle for targeted drug delivery, can be labeled with a radioactive isotope to enable the study of antibody pharmacokinetics through imaging methods. Utilizing computational topology techniques, we have developed a method that is capable of quantifying local Indium-111 (In-111) labeled antibody uptake behavior in single-photon emission computed tomography (SPECT) images of murine tumors.

Due to natural biological processes, such as necrosis and angiogenesis, the concentration of radio labeled antibodies will be unevenly distributed throughout a tumor. Information about local uptake behavior cannot be captured by global measures such as percent injected dose per gram (%ID/g). A new method that is capable of quantifying local, topological information from SPECT images of murine tumors injected with targeted antibodies is presented in this thesis. Applying this method to decay corrected and normalized SPECT tumor images in inviCRO Study 1881-040 In-111 differentiates between the local uptake behavior in patients injected with target antibody and patients injected with the control antibody.

2 Background

2.1 Computational topology and Persistent Homology

Computational topology focuses on the quantification of topological structures in datas. Applying a filtration to a data set allows one to use persistent homology and computational topology to detect and quantify topological information from their data. In dimension zero, persistent homology is concerned with the connected components in the data through a given filtration.

Consider the dimension zero homology of a data structure represented by a 3-D array of values, D. Let $F_0 \subset F_1 \subset ... \subset F_n$ be a filtration of D such that $F_0 = \min(F)$ and $F_n = \max(F)$. At each step in the filtration, F_k consider the set $D_k = \{x \in D | x \leq F_k\}$. In persistent homology, the number of connected components at each step in the filtration is of interest.

Persistence diagrams can be utilized to represent structural information of a data set. To

construct a dimension zero persistence diagram, a birth and death time is assigned to each component. A component C is born at time C_b if $C \in b_b$ and $C \notin D_j$ for all $j < b$. The component C dies at time C_d if (i) there exist points $c \in C$ at time C_d such that at time C_{d-1}, $c \in C'$ and $C \cap C' = \emptyset$ at F_{d-1}, and (ii) component C'_b ¡ C_b. Condition (ii) is known as the Elder Rule. A traditional persistence diagram is a scatter plot of points (C_b, C_d) for each connected component.

In this thesis computational topology is utilized to analyze data of SPECT tumor images. Due the structure of the dataset presented, a descending filtration of the tumor images has been applied to each image and the values greater than a given filtration value at each stage in the filtration of the tumor data are analyzed. In this analysis the Elder Rule is ignored and the death of a component is defined as as the value at which is joins to an existing connected component. A modification of the traditional persistence diagram is utilized to represent the topological information

2.2 Morse theory

A Morse function is a smooth function with only non-degenerate critical points. Given a Morse function, Morse theory allows one to analyze the structure of critical points, level sets, and sub-level sets of the function on a given domain. Specifically, Morse theory extracts topological information from the domain by only looking at the behavior of the function at its critical points.

In the dataset presented in this thesis, each image is represented by a 3-D array of voxel intensities, which are interpreted as values of a Morse function on a 3-D domain. The locally generic nature of smoothed SPECT tumor images makes it reasonable to assume that the data contain no degenerate critical points. Viewing the SPECT tumor image data as a Morse function a 3-D domain, topological information can be extracted from different regions of the tumors. Specifically, the uptake behavior of the tumor images can be quantified by locating local maxima of the SPECT tumor images and analyzing regions around the maxima.

2.3 Tumor biology and drug targeting

Targeted cancer treatments have the potential to increase the amount of drug delivered directly to the tumor while minimizing the amount of treatment concentrated in healthy areas of the body. One method of drug targeting involves attaching treatment to antibodies with specific tumor cell antigen binding sites. If the antibody successfully reaches the targeted tumor cells it will then be internalized by the cell and degraded through chemical processes, however, several other biological processes can either increase or decrease the concentration of antibodies that reach the targeted tumor cells.

Complications in pre-clinical drug targeting studies arise because it is difficult to isolate the binding effects of a targeted antibody. First, antibodies are radiolabeled with an isotope so that the antibody can be tracked through imaging processes. Consequently, the radioactive decay of the chosen isotope will decrease the amount of uptake behavior that can be captured through imaging.

Second, the antibody can only reach the tumor through the blood. If the antibody is cleared from the blood at a slower rate, then it can result in higher tumor uptake compared

to an antibody that clears the blood faster.

Third, the enhanced permeability and retention effect (EPR) can trap antibodies in tumor tissue and therefore increase the antibody concentration. Tumor cells grow by simulating new blood vessels through a process called angiogenesis. The rapid production of new blood vessels results in abnormal vascular structure in tumor tissue which results in poor lymphatic drainage. This is known as the EPR effect. Since antibodies are relatively large molecules they can get trapped and accumulate in the tumor interstitial tissue. The EPR effect can also cause antibodies to accumulate in the tumor at different rates, depending on their location within the tumor. For example, antibodies might be slower to clear from necrotic regions of the tumor due to decreased blood supply.

It is imperative to account for all of the processes that can effect antibody accumulation and clearance in order to fully understand the binding rates of a targeted antibody. Analyzing local regions of the tumor might provide additional information regarding the behavior of the antibodies throughout the tumor. In this thesis various normalization techniques are applied to analyze tumors injected with different antibodies, and the tumors are analyzed on a local scale in an attempt to quantify both the binding rates of the antibodies and heterogeneity within the tumors.

3 Overview of Method

3.1 inviCRO Dataset

inviCRO, a pre-clinical software company that focuses on the design and analysis of imaging trial, provided SPECT image data of murine tumors for analysis. The topological methods of analysis presented in this thesis have been applied to the SPECT data for eight patients in inviCRO Study 1881-040 In-111. Each animal in this study has been injected with a non-xenograft tumor from a mouse cell line. There are two groups in this study, one group was injected with In-111 labeled targeting antibody and the other was injected with In-111 control antibody. No treatment was given to either groups in this study.

The patients in this trial have been images at 3, 24, 48, and 72 hours after injection, but in this thesis the hour 3 images are discarded as insignificant. Group 1 will refer to the four patients injected with In-111 labeled targeting antibody (patients A1101-4) and Group 2 will refer to four patients injected with In-111 control antibody (patients A1105-8). The images have been decay corrected using 67.32 hours as the radioactive half-life of In-111. Uptake measures are in microcuries (μCi).

Two separate normalization techniques have been applied to these images. In the first normalization, all images have been normalized by the mean intensity value of the corresponding hour 24 image when calculating persistence. In the second normalization, each tumor image has been normalized by the mean intensity value of the heart uptake of the corresponding image. Heart normalization reduces the discrepancies in blood clearance rates across the patients in an attempt to isolate the difference in binding rates of the two antibodies.

3.2 Approach

A topologic approach for quantifying the the binding rates of targeted antibodies and heterogeneity within the tumors was developed. The underlying mathematical construction of this method is of the zeroth dimension persistent homology of the image filtered by intensity. This method was implemented in a MATLAB pipeline that takes as its input a time series of raw data and ROI images exported from VivoQuant, inviCRO's image analysis software, and produces both binary tree representations of the tumor images and a time series of 2D scatter plots, called persistence diagrams.

The binary tree representations of the tumor images correspond to the zeroth dimension persistent homology, or connected components, of the tumor image. The leaves of the tree correspond to local maxima, while the internal branches represent two or more connected components that have joined at some stage in a filtration of the tumor image. Converting the binary trees into persistence diagrams provides a way to quantify the local behavior of the tumor images.

The persistence diagrams measure heterogeneity across SPECT images by representing the intensity behavior surrounding local maxima of intensity throughout each image. Combining the persistence diagrams of a series of images taken at subsequent times produces a *time series persistence diagram*.

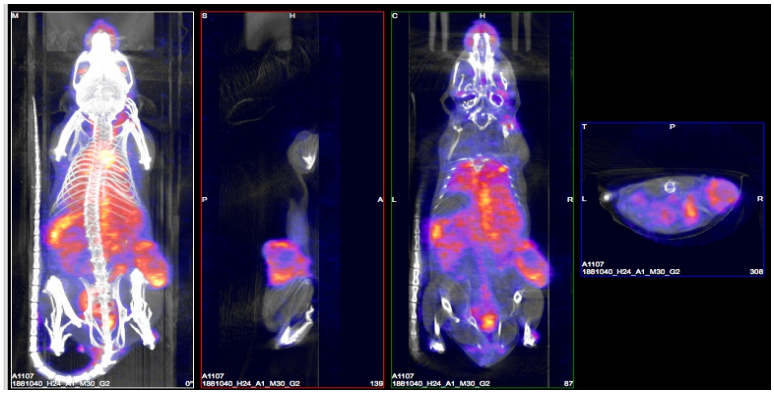
Statistical analyses of the time series persistence diagrams show the significance of the distinctions between the persistence diagrams representing the tumor features of the patients in Group 1 and those representing the tumor features of the patients in Group 2. Various normalization techniques were applied to the persistence diagrams in an attempt to isolate the binding effects of the targeted antibody. The persistence diagrams were normalized by initial time point, hour 24, in order to capture the relative change in antibody accumulation, and separately by the heart's antibody accumulation to capture the antibody binding rate.

In an attempt to analyze uptake behavior in different regions of the tumors, the persistence points were separated into high and low classes based on the value of their corresponding maxima. Analyses of the high and low time series persistence diagrams have been compared to analyses of normalized percent injected dose per gram (%ID/g) and tumor-to-heart ratio calculations on high, low, and all voxels to demonstrate that this method successfully measures tumor heterogeneity which cannot be captured by standard, global measures.

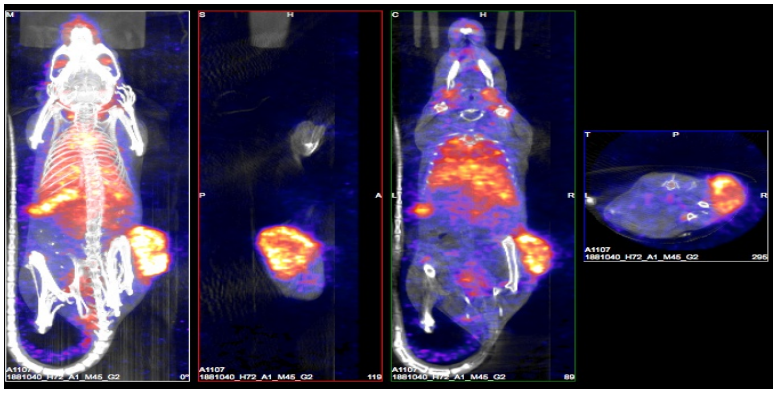
4 Algorithm

4.1 Image Extraction

SPECT image data was exported from VivoQuant using the cropping and 3-D ROI (region of interest) tool. Tumor ROI raw data for each SPECT image was imported into MATLAB, along with cropped regions of non-tumor tissue and heart tissue for comparison. Images were smoothed using a Gaussian smoothing algorithm with box-width 3 and decay corrected using the 67.32 hour radioactive half-life of In-111. Images were separately normalized by the mean intensity value of the corresponding hour 24 image and the mean intensity value of the corresponding heart ROI when calculating persistence.



(a)



(b)

Figure 1: (a) A1107 H24 CT image and sagittal, coronal, and transverse SPECT images. (b) A1107 H72 CT image and sagittal, coronal, and transverse SPECT images.

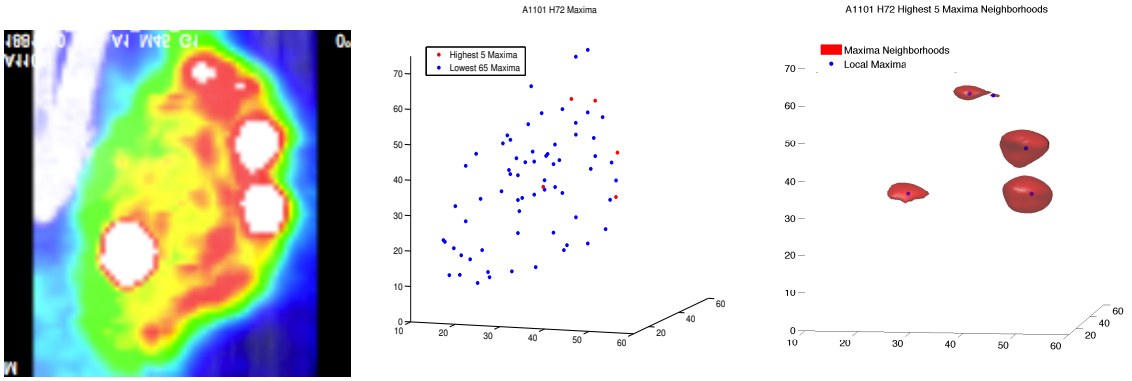
4.2 MATLAB Pipeline

Each image in the dataset was fed through a MATLAB pipeline for analysis.

4.2.1 Identification of Maxima

For each image, the pipeline identifies a collection of local maxima. A voxel is designated as a local maximum if the intensity value M_i for that voxel is strictly larger than the values of the remaining voxels in a $3 \times 3 \times 3$ box of voxels centered at the voxel of interest. The location and uptake value of each maximum are recorded, and maxima whose uptake values are lower than the 95th percentile of non-tumor tissue uptake are discarded.

For each stored maximum, the algorithm creates a connected neighborhood of voxels. It maximizes the number of voxels in each neighborhood so that each neighborhood contains exactly one maximum, no two maximum neighborhoods intersect, and for each neighborhood the minimum uptake intensity value m_i of the neighborhood occurs on the boundary.



(a) (b) (c)

Figure 2: (a) VivoQuant cropped A1101 H72 tumor image (The axes have been flipped to match the MATLAB images). (b) Maxima points of A1101 H72 tumor image. (c) Maxima neighborhoods of highest 5 maxima of A1101 H72 tumor image.

4.3 Construction of optimal maxima neighborhoods

To find the optimal collection of neighborhoods around each maximum, the algorithm first picks an intensity value, called an *isovalue*, between the first and second maxima value, where the maxima values are ranked in descending order. Then, it takes all of the voxels in the 3-D array whose value is greater than or equal to the isovalue and identifies connected components, called *isosurfaces*. This process is repeated as isovalues are iteratively chosen between critical points until all of the maxima have appeared in the level set. Once all of the maxima appear in the level set, isovalues below the lowest maxima value are selected until all of the components in the level set have joined to one connected component.

The surface information for each isovalue is store in a MATLAB array. The columns of these arrays correspond to critical values and the rows correspond to isovalues. The isosurfaces are identified in the following way:

1. Get dimension of smoothed image and build a mesh grid.

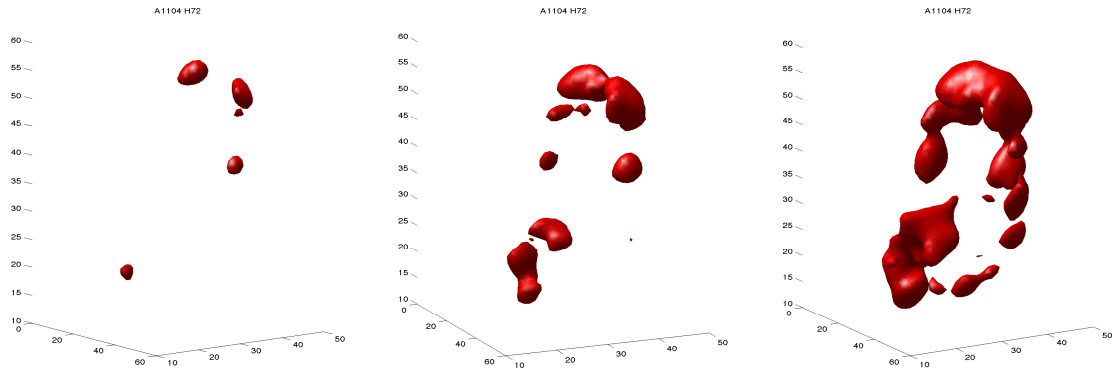
2. Use MATLAB's *isosurface* function to get the faces and vertices list of the voxels in the isosurface. Use the face list to create an edge list, and then use the vertices list to create an incidence matrix.

3. Use *graphconcomp* from MATLAB's bioinformatics package to calculate the number of connected components in the isosurface and match each vertex with a connected component.

4. Covert each component's vertex list into a Delaunay triangulation through MATLAB's *delaunayTriangulation* function and use MATLAB's *nearestNeighbor* function to find which vertices and surfaces are closest to the critical points in the level set.

5. Assign an integer value to each surface and store the surface information in the array containing the image's surface information.

Note: Here it is assumed that the components in the isosurface are oriented surfaces that partition the set of maxima into those that lie in the interior of the surface (voxel intensities greater than isovalue) and those that lie in the exterior (voxel intensities less than isovalue).



(a) (b) (c)

Figure 3: (a) Isosurfaces corresponding to isovalue = 3.38 $\times 10^{-04}$ of patient A1104 H72. (b) Isosurfaces corresponding to isovalue = 3.0 $\times 10^{-04}$ of patient A1104 H72. (c) Isosurfaces corresponding to isovalue = 2.615 $\times$ 10^{-04} of patient A1104 H72.

For each isovalue, the algorithm checks for bad joins and bad splits. A bad join refers to more than two connected components joining to one connected component. A bad join occurs when two consecutive isovalues are too far apart in value, and so the algorithm checks and fixes bad joins in the following way:

1. Compare rows in the array containing the isosurface information for each image that correspond to consecutive isovalues.

2. Compare the column entries of these two rows to see if columns containing more than 2 distinct integer entries in one row have the same integer entry in the next row.

3. Fix joins by inserting more isovalues in-between the values where the bad join occurred until no more bad joins exist.

Similarly, a bad split occurs when critical points that were previously identified as being on the same component have become misidentified as splitting onto two different components. This happens when an isosurface surrounds one or more minima as the isovalues decrease, which causes a void to form that encloses the minim(um)a. The algorithm checks for splits in the same way it checks for joins, except it moves in the opposite direction. Specifically, the algorithm checks for bad splits in the following way:

1. Identity the maxima on each component in the previous isosurface.

2. Record which component(s) the maxima on the previous component are identified with in the current isosurface.

3. If the maxima on the previous component are not on the same surface in the current isosurface then a split has been found. In this case, identify the split surfaces and find the size of each surface by counting the number of vertices in its vertex list.

4. Identify the larger component of the split surfaces as the maxima component and the smaller component as the minima component.

5. For each maxima in the isosurface that has been identified as being on a minima surface, replace the minima surface label with the correct maxima surface label.

Once all of the bad splits and bad joins have been correctly identified and fixed, the algorithm produces a MATLAB array that contains the surface information for each isovalue in the filtration of the tumor image. Surface labels 1-(the number of critical points) correspond to surfaces around a single critical point, and surface labels greater than the number of critical points correspond to surfaces containing more than one critical point. A surface is born at the isovalue corresponding to the row in which the surface label first appears, and a surface dies at the isovalue corresponding to the first row in which the surface no longer appears. Since the algorithm descends through the filtration of the image, the death times of a surface are always less than their corresponding birth times.

4.3.1 Construction of Binary Trees

Binary trees are used to represent the structure of the Morse functions corresponding to the 3-D SPECT image data. The leaves of the tree represent critical points of the Morse function, and branches of the tree represent critical points joining. The length of leaves andbranches correspond to the length of the range of isovalues for which the component appears in the level set.

The algorithm uses MATLAB's *phytree* function to create the binary tree representation of each image. The inputs for *phytree* are a two-column array corresponding to the leaves and branches of the tree and a vector containing the lengths of the leaves and branches of the tree. The rows in the two-column array input correspond to the tree's branches and the entries in each row are the children of that branch. The number of rows in this array will be one less than the number of critical points in the image.

The algorithm identifies branches in the MATLAB array containing each image's surface information as those whose surface label is greater than the number of critical points in the image. For each branch, starting with the branch whose label is (number of critical points + 1), the algorithm finds the children of that branch by identifying the surface labels in the row above the row where the branch first appears. The surface labels of the children for each branch are inserted into a two-column array, which will be the first input for *phytree*. The second input for *phytree* is a distance vector corresponding to the length of the range of isovalues for which each component persists, which can be extracted directly from the MATLAB array containing each image's surface information.

The binary tree representations contain surface information about the dimension 0 persistent homology of each image. Therefore, these structures can be utilized to identify topological information about the antibody accumulation throughout local neighborhoods of the SPECT tumor images.

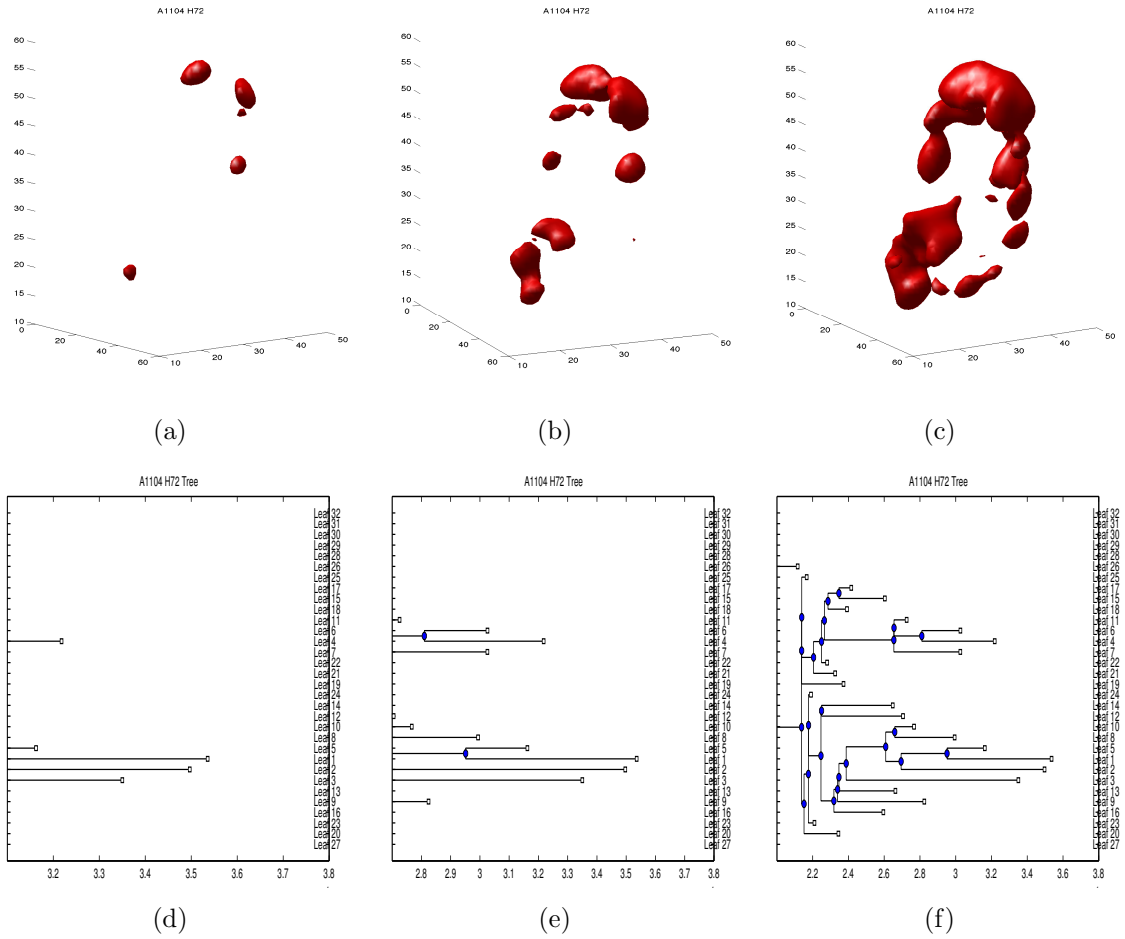


Figure 4: (a) Isosurfaces corresponding to isovalue = 3.38 $\times 10^{-04}$ of patient A1104 H72. (b) Isosurfaces corresponding to isovalue = 3.0 $\times 10^{-04}$ of patient A1104 H72. (c) Isosurfaces corresponding to isovalue = 2.615 $\times$ 10^{-04} of patient A1104 H72. (d) Binary tree corresponding to isovalue = 3.38 $\times 10^{-04}$ of patient A1104 H72. (e) Binary tree corresponding to isovalue = 3.0 $\times 10^{-04}$ of patient A1104 H72. (f) Binary tree corresponding to isovalue = 2.615 $\times$ 10^{-04} of patient A1104 H72.

4.3.2 Construction of Persistence Diagrams

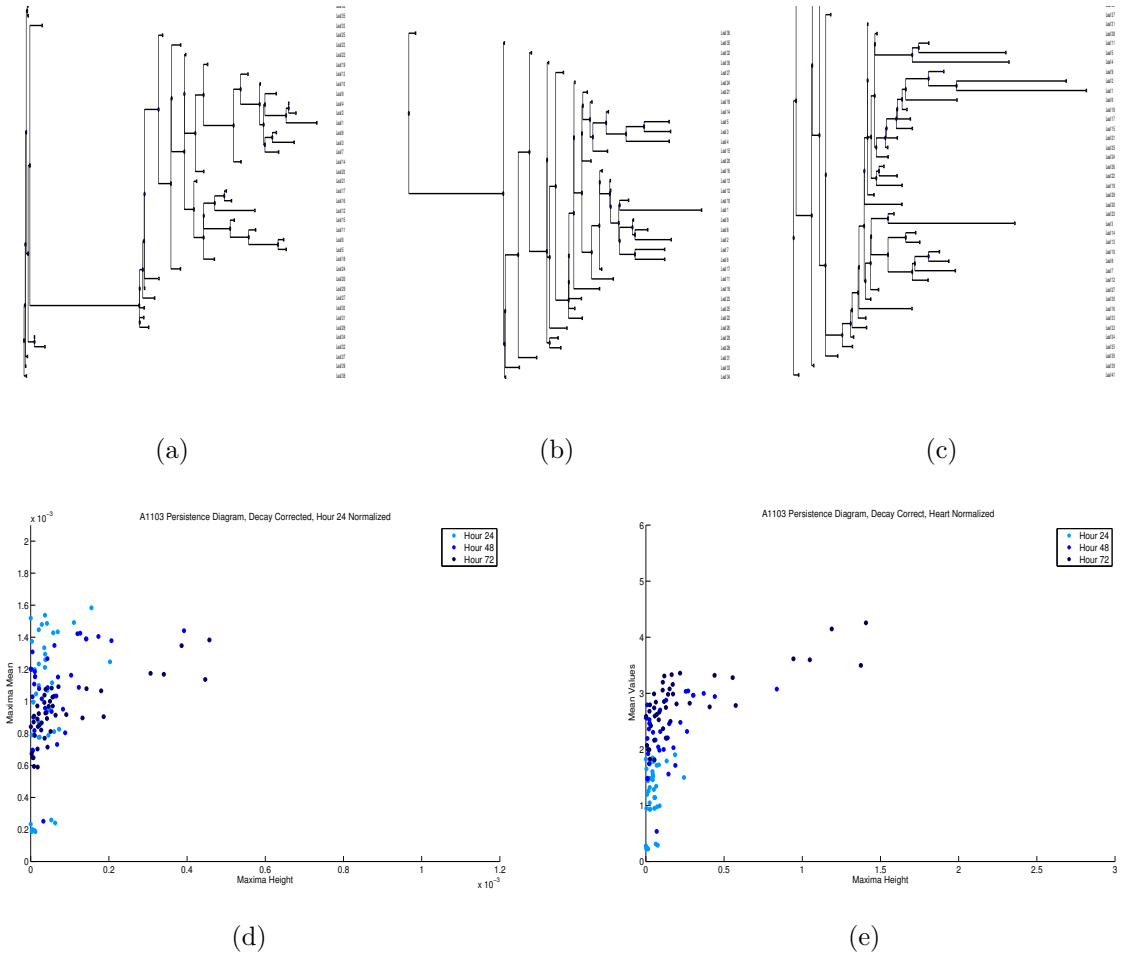


Figure 5: (a) Binary tree representation of image A1103 H24. (b) Binary tree representation of image A1103 H48. (c) Binary tree representation of image A1103 H72.(d) Decay corrected and hour 24 normalized time series persistence diagram of patient A1103. (e) Decay corrected and heart normalized time series persistence diagram of patient A1103.

To better quantify the topological information of the tumor images, the binary tree representations are converted into 2-D scatter plots called persistence diagrams. The persistence diagram for an image is a scatter plot of the points

$$(x_i, y_i) = (M_i - m_i, \frac{M_i + m_i}{2}),$$

where x_i is the length of the i^{th} leaf in the binary tree and y_i is the midpoint value of the i^{th} leaf in the tree. Similarly, x_i can be thought of as the range of intensity in the neighborhood of the maximum, and y_i approximates the mean intensity in the neighborhood of the maximum. Thus, the persistence plot contains measures of the neighborhoods of maxima throughout the tumor. For each patient, the pipeline combines the H24, H48, and H72 persistence diagrams to produce a time series persistence diagram.

In order to explore the influence of regions of high uptake on the trends in persistence, the algorithm separates persistence points into two groups based on the value of the local maxima to which they correspond. An *intensity threshold* is set at the 95th percentile of maximum values in an image. Persistence points corresponding to local maxima with values below the intensity threshold are classified as low maxima and persistence points corresponding to the local maxima with values above the intensity threshold are classified as high maxima. Persistence diagrams are constructed for each patient for both the low and high maxima.

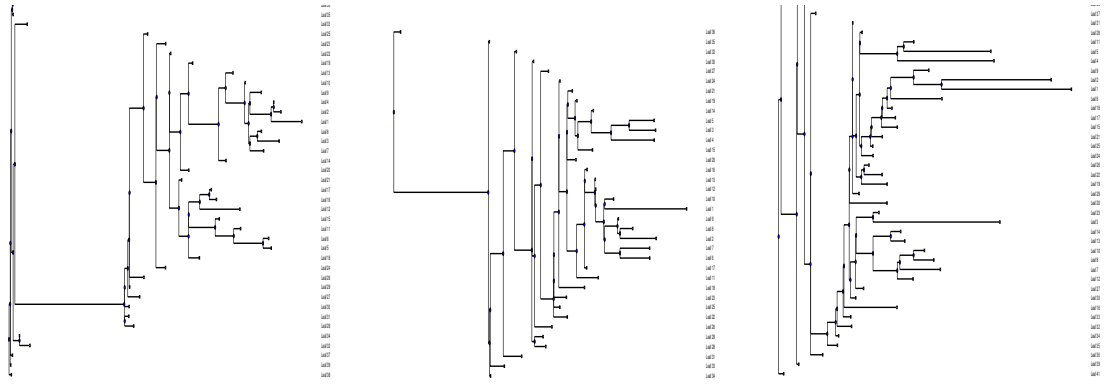
Persistence diagrams corresponding to patients A1101-A1104 were combined to create an aggregate Group 1 time series persistence diagram representing the topological information of the targeted antibody, and persistence diagrams corresponding to patients A1105-A1108 were combined to create an aggregate Group 2 time series persistence diagram representing the topological information of the control antibody.

Two collection of persistence diagrams are created for each patient. The first collection of persistence diagrams contains decay corrected persistence points normalized by the initial time point, hour 24, in order to capture the relative change in antibody accumulation. The second collection of persistence diagrams contains decay corrected persistence points normalized by the heart's antibody accumulation in an attempt to capture the antibody binding rate.

5 Results

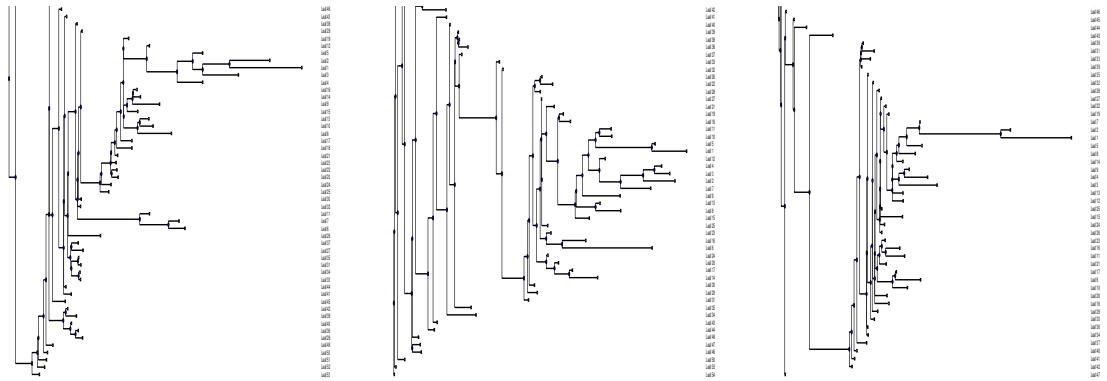
5.1 Binary tree representations

The following figures are binary tree representations of tumor images from inviCRO Study 1881-040 In-111. The first time series of binary trees comes from a patient in Group 1, representing the tumor images of mice inject with the targeted antibody. The second time series of binary trees comes from a patient in Group 2, representing the tumor images of mice inject with the control antibody.



(a) (b) (c)

Figure 6: (a) Binary tree representation of image A1103 H24. (b) Binary tree representation of image A1103 H48. (c) Binary tree representation of image A1103 H72.



(a) (b) (c)

Figure 7: (a) Binary tree representation of image A1107 H24. (b) Binary tree representation of image A1107 H48. (c) Binary tree representation of image A1107 H72.

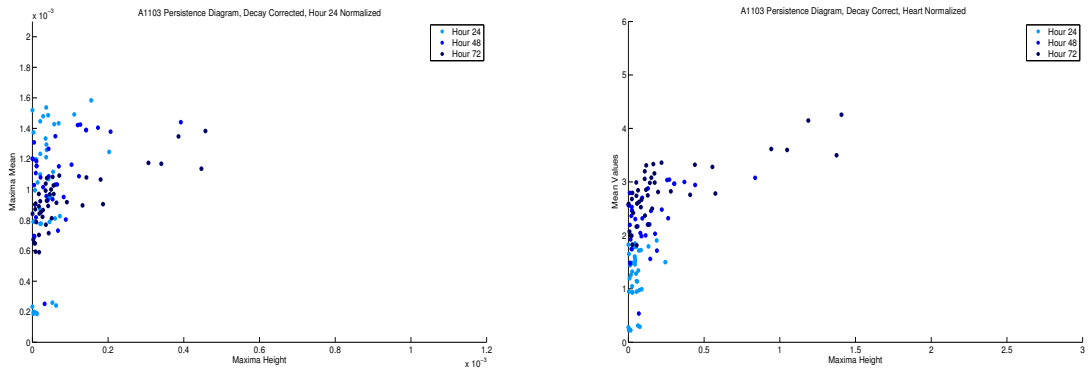
It appears as through the leaves in the binary tree representation of patient A1103 move closer to the root of the tree over time, while the leaves of the binary tree representation of patient A1107 remain approximately constant with respect to their location in the tree over time. The length of the leaves also appear to be increasing over time in patient A1103, but they remain approximately constant in patient A1107. To better quantify these difference, each leaf is assigned coordinates

$$(x_i, y_i) = (M_i - m_i, \frac{M_i + m_i}{2}),$$

where x_i is the length of the i^{th} leaf in the binary tree and y_i is the midpoint value of the i^{th} leaf in the tree. The points for all of these points in the time series of binary trees are combined to create time series persistence diagrams for each patient in the study.

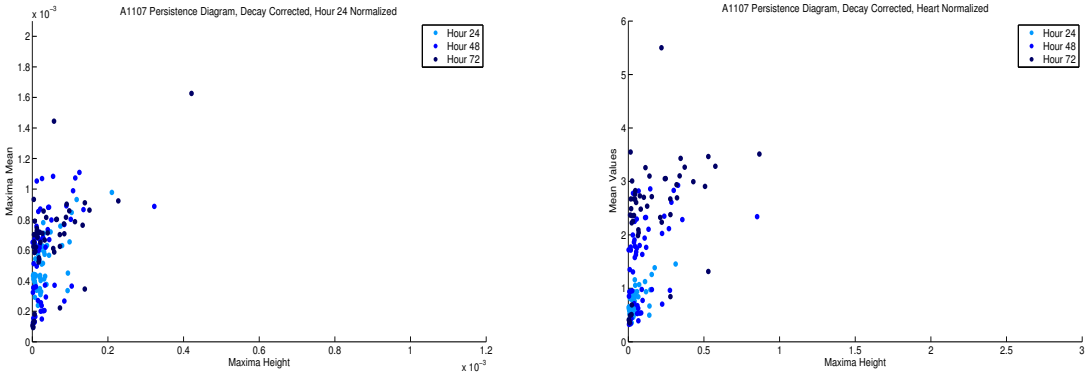
5.2 Persistence diagrams

The following figures are pairs of time series persistence diagrams of tumor images from inviCRO Study 1881-040 In-111. The first pair of time series persistence diagrams comes from a patient in Group 1, representing the tumor images of mice inject with the targeted antibody. The second pair of time series persistence diagrams comes from a patient in Group 2, representing the tumor images of mice inject with the control antibody. Each pair of persistence diagrams contains one diagram where the persistence points are normalized by the initial time point, hour 24 and one diagram normalized by the heart's antibody accumulation.



(a) (b)

Figure 8: (a) Decay corrected and hour 24 normalized time series persistence diagram of patient A1103. (b) Decay corrected and heart normalized time series persistence diagram of patient A1103.

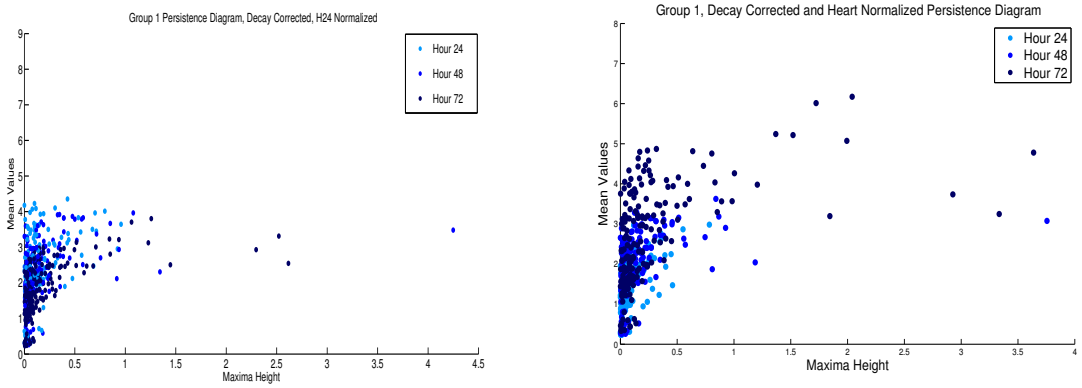


(a) (b)

Figure 9: (a) Decay corrected and hour 24 normalized time series persistence diagram of patient A1107. (b) Decay corrected and heart normalized time series persistence diagram of patient A1107.

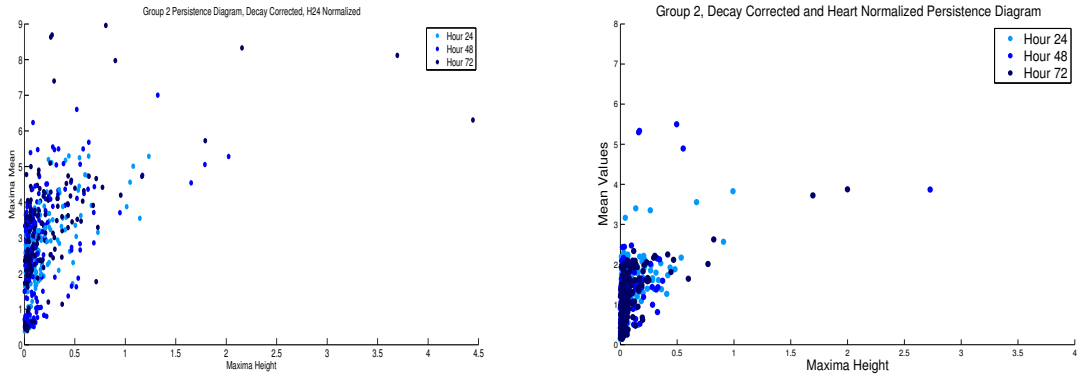
In the decay corrected and hour 24 normalized time series persistence diagrams, the y-coordinates tend to decrease over time for patient A1103, but there is no apparent trend for the y-coordinates corresponding to patient A1107. This behavior corresponds to the location of the leaves in the binary tree representations of patients A1103 and A1107. Contrastingly, in the decay corrected and heart normalized time series persistence diagrams, the y-coordinates tend to increase over time for patient A1103, but there is no apparent trend for the y-coordinates corresponding to patient A1107.

The persistence points for all of the patients in Group 1 were combined to create an aggregate Group 1 time series persistence diagram. Similarly, the persistence points for all of the patients in Group 2 were combined to create an aggregate Group 2 time series persistence diagram. The group persistence diagrams are presented below.



(a) (b)

Figure 10: (a) Decay corrected and hour 24 normalized time series persistence diagram of all patients in Group 1. (b) Decay corrected and heart normalized time series persistence diagram of all patients in Group 1.



(a) (b)

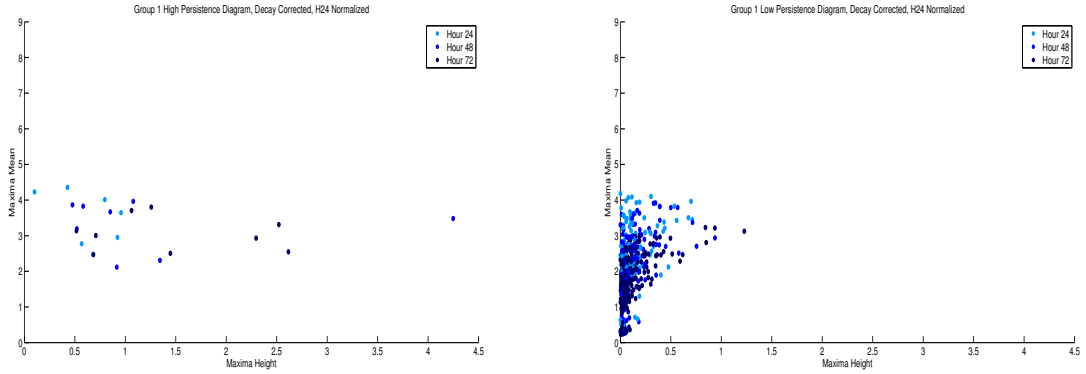
Figure 11: (a) Decay corrected and hour 24 normalized time series persistence diagram of all patients in Group 2. (b) Decay corrected and heart normalized time series persistence diagram of all patients in Group 2.

Similar to the behavior in the patient persistence diagrams, in the decay corrected and hour 24 normalized time series persistence diagrams the y-coordinates tend to decrease over time for Group 1, but there is no apparent trend for the y-coordinates corresponding to Group 2. In the decay corrected and heart normalized time series persistence diagrams, the y-coordinates tend to increase over time for Group 1, while there is no clear trend for the y-coordinates corresponding to Group 2.

The persistence points were separated into two groups based on the value of the local maxima to which they correspond, in an attempt to analyze the influence of regions of high uptake on the trends in persistence. An *intensity threshold* is set at the 95th percentile of maximum values in an image. Persistence points corresponding to local maxima with values below the intensity threshold are classified as low maxima and persistence points

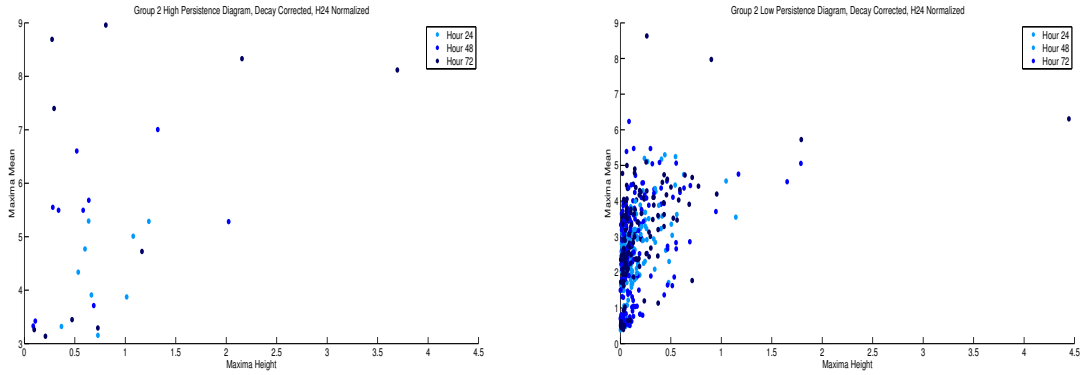
corresponding to the local maxima with values above the intensity threshold are classified as high maxima. Persistence diagrams are constructed for each patient for both the low and high maxima.

The high and low normalized persistence plots corresponding to Group 1 and Group 2 are presented below.



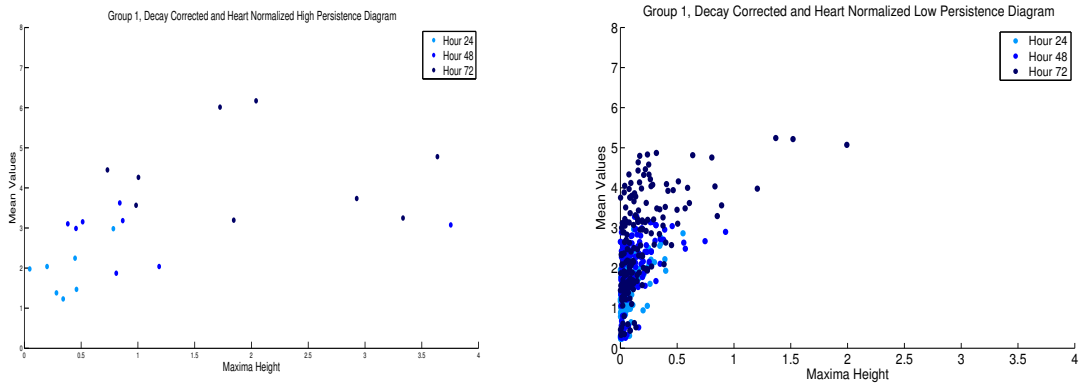
(a) (b)

Figure 12: (a) Decay corrected and hour 24 normalized time series persistence diagram of the high maxima corresponding to all patients in Group 1. (b) Decay corrected and hour 24 normalized time series persistence diagram of the low maxima corresponding to all patients in Group 1.



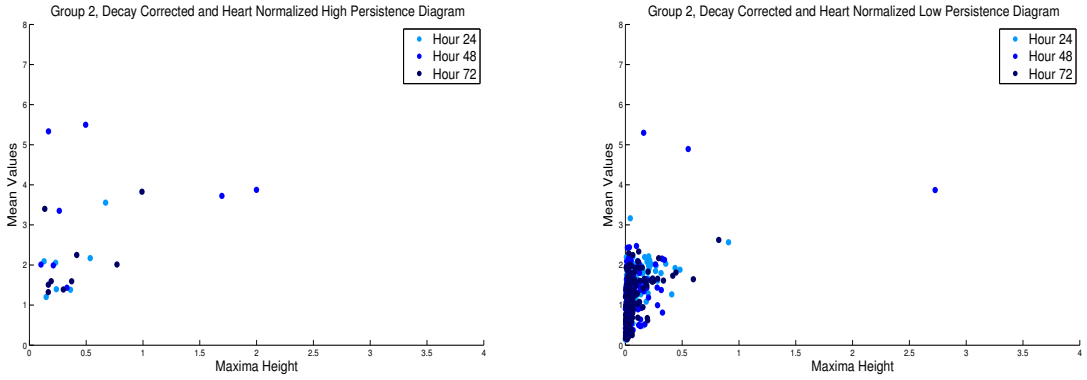
(a) (b)

Figure 13: (a) Decay corrected and hour 24 normalized time series persistence diagram of the high maxima corresponding to all patients in Group 2. (b) Decay corrected and hour 24 normalized time series persistence diagram of the low maxima corresponding to all patients in Group 2.



(a) (b)

Figure 14: (a) Decay corrected and heart normalized time series persistence diagram of the high maxima corresponding to all patients in Group 1. (b) Decay corrected and heart normalized time series persistence diagram of the low maxima corresponding to all patients in Group 1.



(a) (b)

Figure 15: (a) Decay corrected and heart normalized time series persistence diagram of the high maxima corresponding to all patients in Group 2. (b) Decay corrected and heart normalized time series persistence diagram of the low maxima corresponding to all patients in Group 2.

Statistical analyses were applied to the collection of group persistence diagram presented above in order test if persistence diagrams can distinguish patients injected with the targeted antibody from those injected with the control antibody.

5.3 Statistical analyses of Persistence diagrams

Kruskal-Wallis tests and random effects modeling were applied to the collection of time series persistence diagrams, as well as to the separate high and low persistence points. These tests were applied to aggregate time series persistence diagrams comparing all of the persistence points in Group 1 against all of the persistence points in Group 2. The results were

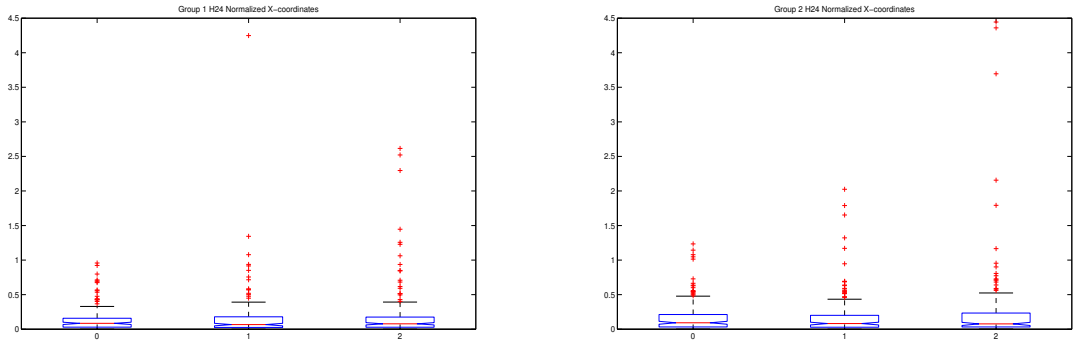
further compared to analyses of the H24, H48, and H72 normalized %ID/g values for each patient.

5.3.1 Kruskal-Wallis tests

Kruskal-Wallis tests indicate that the medians of the decay corrected, hour 24 normalized mean uptake values per maxima neighborhoods (y-coordinates of the persistence diagrams) are significantly different across the time series in both Group 1 and Group 2 with p-values 1.16×10^{-10} and 5.07×10^{-04}, respectively. In contrast, the medians of the decay corrected, hour 24 normalized ranges of intensity per maxima neighborhood (x-coordinates of the persistence diagrams) are not significantly different across the time series in both Group 1 and Group 2, with p-values 0.746 and 0.767, respectively. Thus, the mean uptake values of the maxima neighborhoods will detect changes in uptake while the range of uptake in the maximum neighborhoods will not with respect to the decay corrected and hour 24 normalized data.

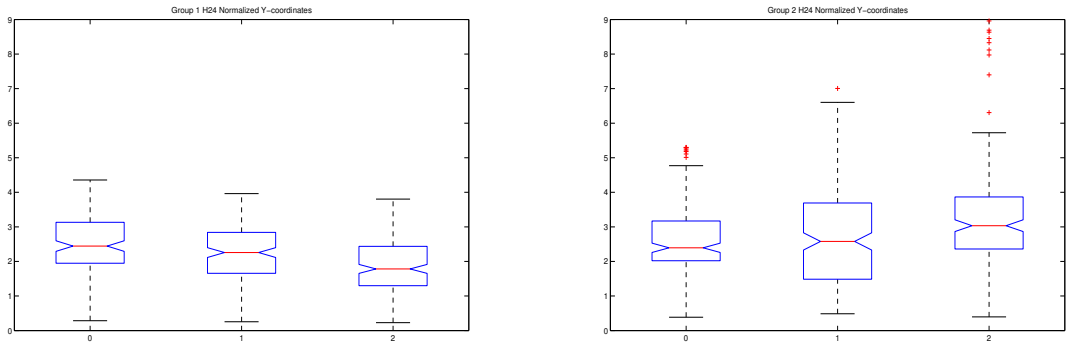
Applying Kruskal-Wallis tests to the decay corrected, heart normalized persistence points indicate that both the mean uptake values of the maxima neighborhoods and the range of uptake in the maximum neighborhoods will detect changes in uptake with respect to the decay corrected and hour 24 normalized data. The medians of the decay corrected, heart normalized mean uptake values per maxima neighborhoods (y-coordinates of the persistence diagrams) are significantly different across the time series in both Group 1 and Group 2 with p-values 2.06×10^{-25} and 2.39×10^{-34}, respectively. Similarly, the medians of the decay corrected, hour 24 normalized ranges of intensity per maxima neighborhood (x-coordinates of the persistence diagrams) are significantly different across the time series in both Group 1 and Group 2, with p-values 2.06×10^{-25} and 0.0015, respectively.

The following box plots demonstrate the trends of the persistence points as a function of time. In these plots, 0 refers to hour 24, 1 refers to hour 48, and 2 refers to hour 72.



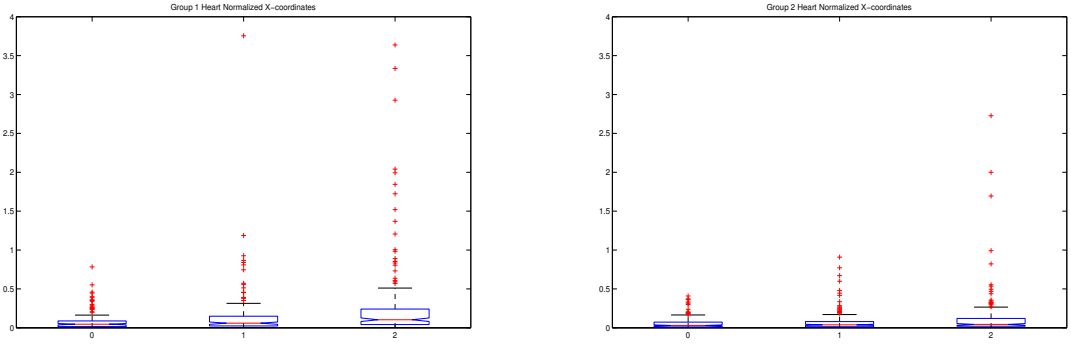
(a) (b)

Figure 16: (a) Notched box plots of Group 1's hour 24 normalized ranges of intensity per maxima neighborhood (x-coordinates). (b) Notched box plots of Group 2's hour 24 normalized ranges of intensity per maxima neighborhood (x-coordinates).



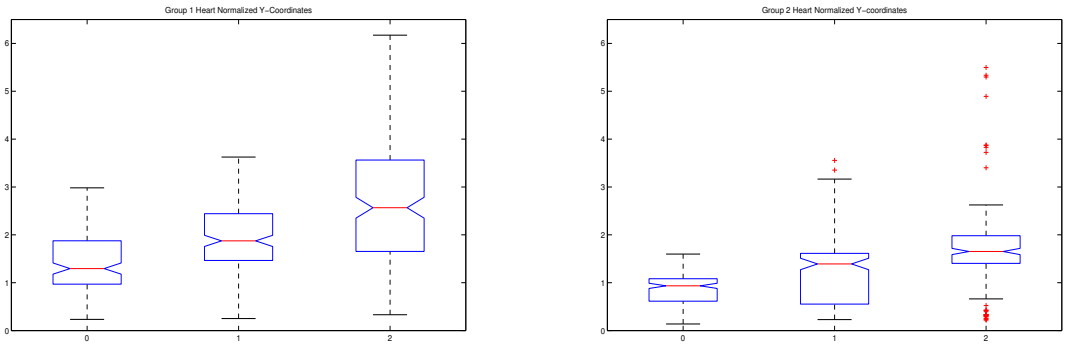
(a) (b)

Figure 17: (a) Notched box plots of Group 1's hour 24 normalized mean values per maxima neighborhood (y-coordinates). (b) Notched box plots of Group 2's hour 24 normalized mean values per maxima neighborhood (y-coordinates).



(a) (b)

Figure 18: (a) Notched box plots of Group 1's heart normalized ranges of intensity per maxima neighborhood (x-coordinates). (b) Notched box plots of Group 2's heart normalized ranges of intensity per maxima neighborhood (x-coordinates).



(a) (b)

Figure 19: (a) Notched box plots of Group 1's heart normalized mean values per maxima neighborhood (y-coordinates). (b) Notched box plots of Group 2's heart normalized mean values per maxima neighborhood (y-coordinates).

The box plots above show that the hour 24 normalized x-coordinates remain approximately constant over time in both Group 1 and Group 2. Contrastingly, the hour 24 normalized y-coordinates of the persistence points decrease over time in Group 1 while they increase over time in Group 2. In regards to the heart normalized persistence points, both the the x-coordinates and y-coordinates increase over time in both Group 1 and Group 2. Random effects modeling was applied to the coordinates of the persistence points as a function of time to test for differences of the rate of change of the persistence points in the two groups in the study.

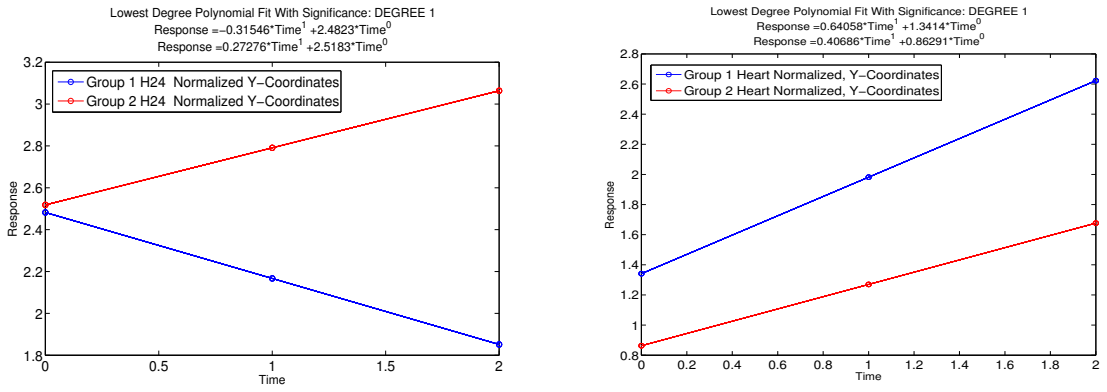
5.3.2 Random effects modeling on the X-coordinates of the persistence points

5.3.3 Random effects modeling on the Y-coordinates of the persistence points

Random effects modeling was applied to the coordinates of the persistence points as a function of time. Trends were assumed linear.

Linear trend comparisons of the group 1 and group 2 mean values per maxima neighborhood (y-coordinates) as a function of time demonstrate that the differences between the Group 1 and Group 2 slopes of the mean values per maxima neighborhood are statistically different with respect to both the hour 24 and heart normalizations. The Group 1 and Group 2 slopes of the hour 24 normalized mean values per maxima neighborhood are statistically different with p-value 5.19×10^{-11}, while the Group 1 and Group 2 slopes of the heart normalized mean values per maxima neighborhood are statistically different with p-value 0.0001.

The figures below show the trend fits of the Group 1 and Group 2 mean values per maxima neighborhood (y-coordinates) as a function of time with respect to both the hour 24 (a) and heart (b) normalizations.



(a) (b)

Figure 20: (a) Linear trend comparisons of the Group 1 and Group 2 hour 24 normalized mean values per maxima neighborhood (y-coordinates). (b) Linear trend comparisons of the Group 1 and Group 2 heart normalized mean values per maxima neighborhood (y-coordinates).

In order to analyze the influence of regions of high uptake on the trends in persistence and in an attempt to capture heterogenous tumor uptake behavior, the persistence points were separated into two groups based on the value of the local maxima to which they correspond. An *intensity threshold* is set at the 95th percentile of maximum values in an image. Persistence points corresponding to local maxima with values below the intensity threshold are classified as low maxima and persistence points corresponding to the local maxima with values above the intensity threshold are classified as high maxima. Within group and between group linear trend comparisons were applied to the high and low persistence points.

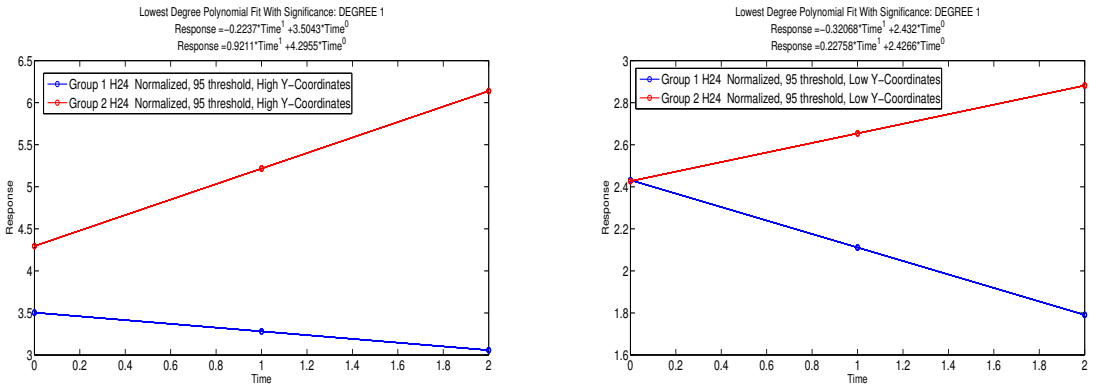
The Group 1 and Group 2 slopes of the high hour 24 normalized mean values per maxima neighborhood are statistically different with p-value 0.016 and the Group 1 and Group 2 slopes of the low hour 24 normalized mean values per maxima neighborhood are statistically different with p-value 5.062×10^{-11}. Contrastingly, the Group 1 and Group 2 slopes of the high heart normalized mean values per maxima neighborhood are not statistically different with p-value 0.273, while the Group 1 and Group 2 slopes of the low heart normalized mean values per maxima neighborhood are statistically different with p-value 4.072×10^{-5}.

Since the Group 1 vs. Group 2 comparisons of the high heart normalized mean values per maxima neighborhood are not statistically different, while the Group 1 vs. Group 2 comparisons of the low heart normalized mean values per maxima neighborhood are statistically different there is evidence to suggest that, with respect to the heart normalized data, the distinguishing features between the group of tumors injected with the targeted antibody and the group of tumors injected with the control antibody come from areas of lower uptake.

It is important to note that the sample size of the high persistence points is significantly smaller than the sample size of the low persistence points. Therefore, the p-values for the linear trend comparisons of the high persistence points with be artificially higher than the p-values for the linear trend comparisons of the low persistence points.

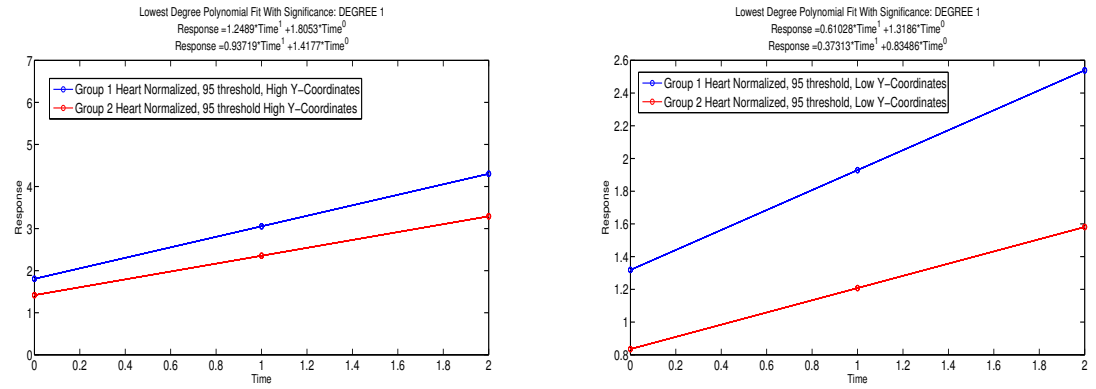
The figures below show the trend fits of the high and low Group 1 and Group 2 mean values per maxima neighborhood (y-coordinates) as a function of time with respect to both

the hour 24 and heart normalizations.



(a) (b)

Figure 21: (a) Linear trend comparisons of the Group 1 and Group 2 high hour 24 normalized mean values per maxima neighborhood (y-coordinates). (b) Linear trend comparisons of the Group 1 and Group 2 low hour 24 normalized mean values per maxima neighborhood (y-coordinates).



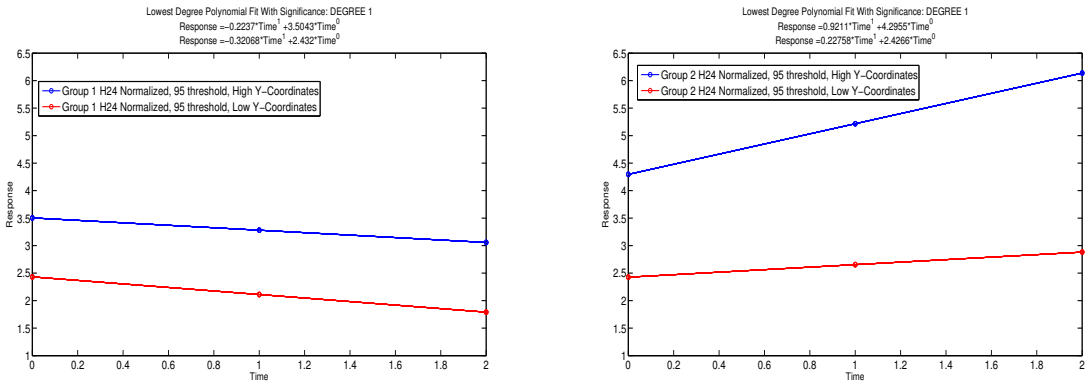
(a) (b)

Figure 22: (a) Linear trend comparisons of the Group 1 and Group 2 high heart normalized mean values per maxima neighborhood (y-coordinates). (b) Linear trend comparisons of the Group 1 and Group 2 low heart normalized mean values per maxima neighborhood (y-coordinates).

Within group linear trend comparisons were applied to the high and low persistence points to test for heterogeneous tumor uptake behavior. The slopes of the Group 1 high and low hour 24 normalized mean values per maxima neighborhood are not statistically different with p-value 0.650, while the slopes of the Group 2 high and low hour 24 normalized mean values per maxima neighborhood are statistically different with p-value 0.0180. This analysis suggests that there exists homogenous uptake behavior in Group 1 that is not present in Group 2.

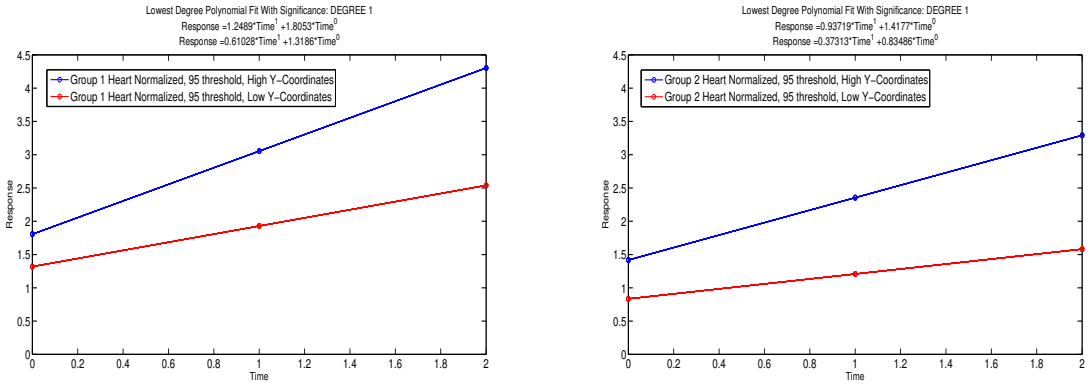
Homogenous uptake behavior in Group 1, however, is not captured with the heart normalized data. The slopes of the Group 1 high and low heart mean values per maxima neighborhood are statistically different with p-value 0.006, and similarly the slopes of the Group 2 high and low heart normalized mean values per maxima neighborhood are statistically different with p-value 2.476×10^{-5}.

The figures below show the trend fits of the within group, high vs. low mean values per maxima neighborhood (y-coordinates) as a function of time with respect to both the hour 24 and heart normalizations.



(a) (b)

Figure 23: (a) Linear trend comparisons of the Group 1 high vs. low hour 24 normalized mean values per maxima neighborhood (y-coordinates). (b) Linear trend comparisons of the Group 2 high vs. low hour 24 normalized mean values per maxima neighborhood (y-coordinates).



(a) (b)

Figure 24: (a) Linear trend comparisons of the Group 1 high vs. low heart normalized mean values per maxima neighborhood (y-coordinates). (b) Linear trend comparisons of the Group 2 high vs. low heart normalized mean values per maxima neighborhood (y-coordinates).

6 %ID/g and tumor-to-heart ratio

In order to test if the method of analyzing tumor uptake with persistence provides additional information than standard analyses, the results of our analysis were compared to standard metrics. The standard metrics used for comparison are percent injected dose per gram (%ID/g) and the tumor-to-heart ratio (T:H). %ID/g is a common measure of uptake which can be calculated by

$$\%\text{ID/g} = \frac{\text{uptake}}{\text{injected dose} \times 100} \times \frac{\text{volume}}{1000},$$

where uptake is the total SPECT sum over entire tumor region in μCi, injected dose is the decay corrected injection dose of antibody, and volume is the total volume of tumor region in mm^3. The %ID/g calculations were normalized by the hour 24 %ID/g value for each patient.

T:H is a measure used to normalize the blood clearance rate across patients in a trial. T:H is defined as the quotient of the tumor %ID/g and the heart %ID/g. The hour 24 normalized persistence diagrams are compared to the tumor %ID/g calculations and the heart normalized persistence diagrams are compared to the T:H calculations. The calculations were repeated for separate high and low regions of the tumor, where high regions correspond to voxels whose value is in the 95th percentile of the tumor voxel values, and low regions correspond to voxels whose value is less than the 95th percentile of tumor voxel values.

The following figures show the tumor %ID/g, heart %ID/g, and T:H calculations for all voxels, high voxels, and low voxels, where the calculations corresponding to Group 1 patients are in the top row and calculations corresponding to Group 2 patients are in the bottom row.

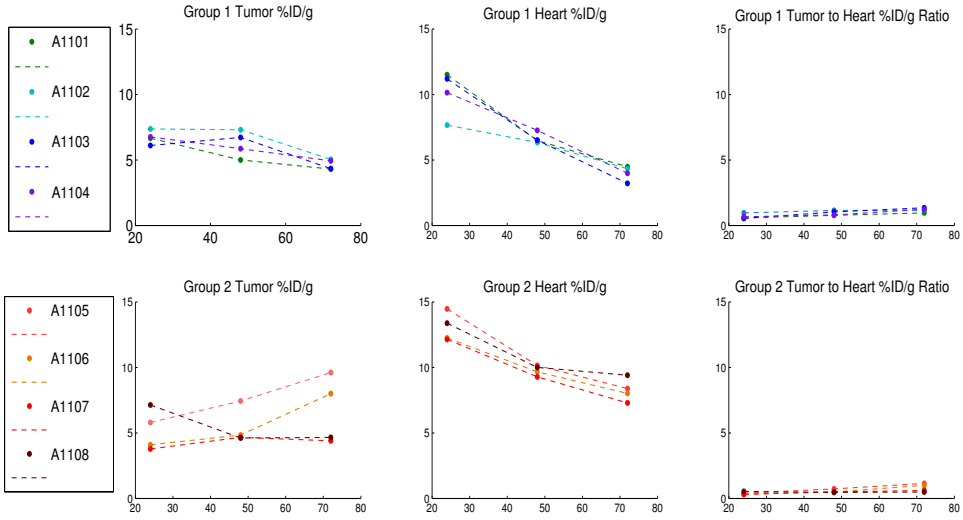


Figure 25: Tumor %ID/g, heart %ID/g, and T:H calculations of all voxels from right to left with patients in group 1 represented in the top row and patients in group 2 represented in the bottom row

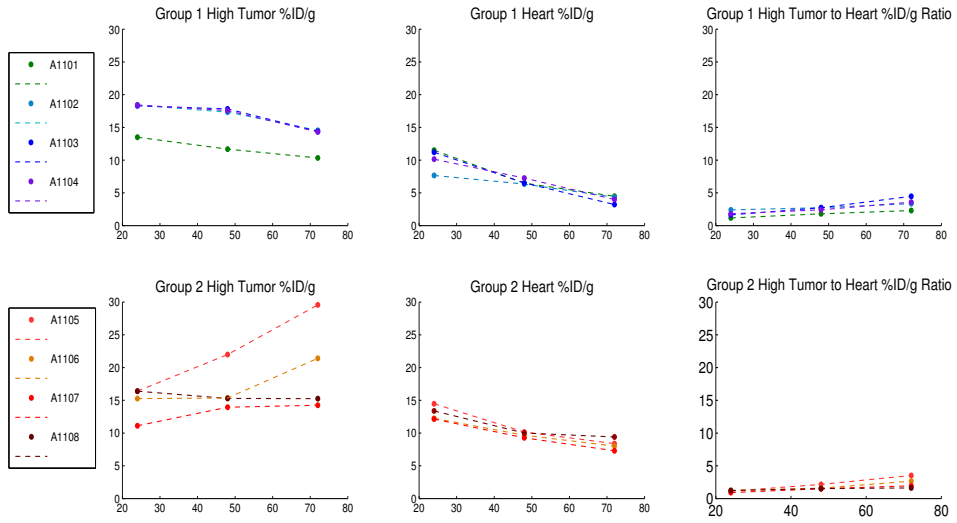


Figure 26: Tumor %ID/g, heart %ID/g, and T:H calculations of high voxels from right to left with patients in group 1 represented in the top row and patients in group 2 represented in the bottom row

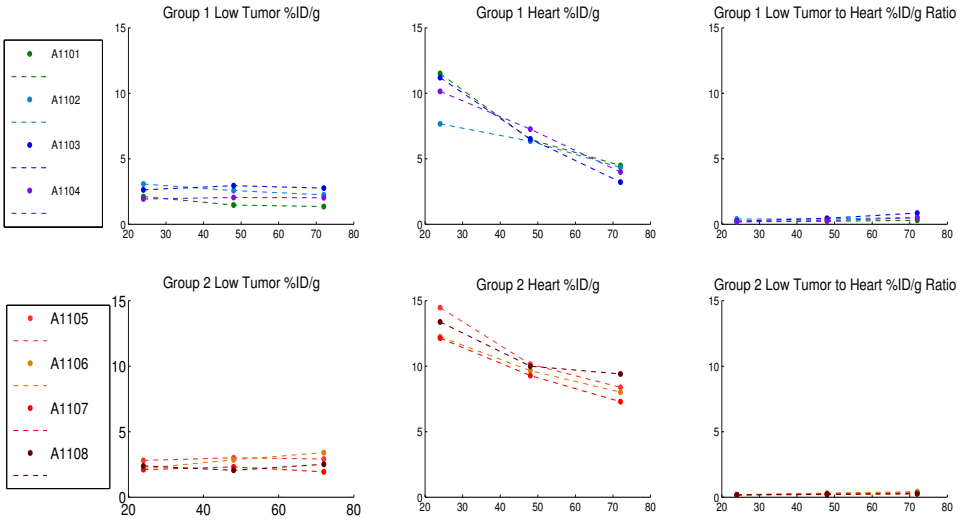
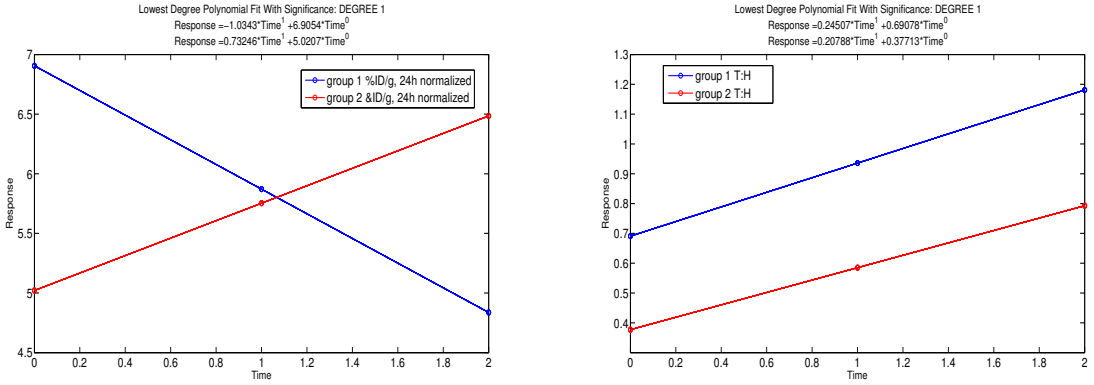


Figure 27: Tumor %ID/g, heart %ID/g, and T:H calculations of low voxels from right to left with patients in group 1 represented in the top row and patients in group 2 represented in the bottom row

Linear fit comparisons of the Group 1 and Group 2 hour 24 normalized %ID/g values

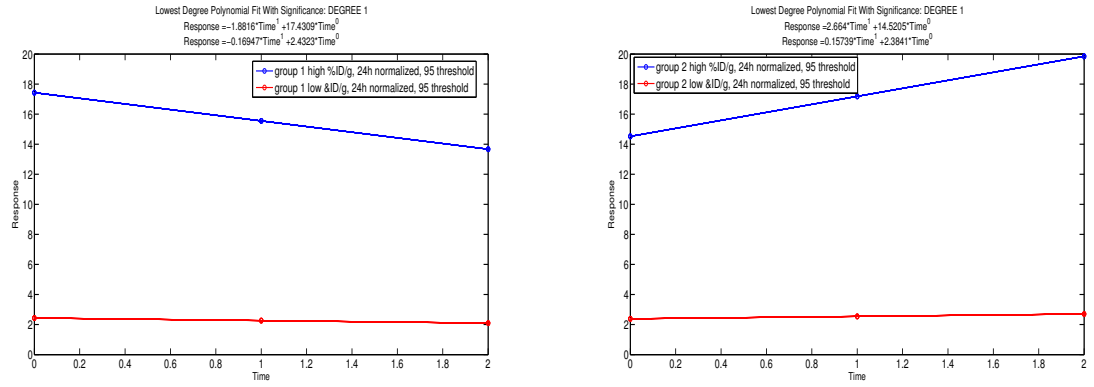
as a function of time yield results similar to the linear fit comparisons of the Group 1 and Group 2 persistence diagrams. As was the case with the hour 24 normalized Y-coordinate of the persistence points, the %ID/g values in Group 1 decrease over time and the %ID/g values in Group 2 increase over time. The distinction between %ID/g values and time series persistence diagrams becomes evident when the %ID/g values are separated into two groups based on voxel values. In contrast to the mean values, the differences between %ID/g of high voxels and %ID/g of low voxels are not statistically different in both Group 1 and Group 2 with p-values 0.0624 and 0.139, respectively.

Linear fit comparisons of the T:H calculations fail to distinguish group 1 from group 2 with a p-value of 0.686. Similarly, the T:H calculations on the high and low voxels failed to distinguish group 1 from group 2 with p-values 0.529 and 0.201, respectively. In contrast to %ID/g, the differences between the T:H calculations of high voxels and the T:H calculations of low voxels are statistically different in both Group 1 and Group 2 with p-values 0.006 and 0.003, respectively.



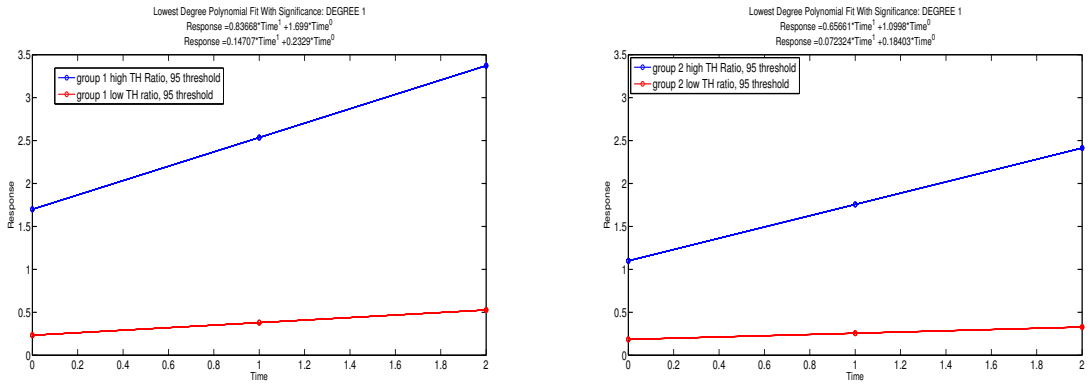
(a) (b)

Figure 28: (a) Linear trend comparisons of the Group 1 and Group 2 %ID/g calculations. (b) Linear trend comparisons of the Group 1 and Group 2 T:H calculations.



(a) (b)

Figure 29: (a) Linear trend comparisons of the Group 1 high and low %ID/g calculations. (b) Linear trend comparisons of the Group 2 high and low %ID/g calculations.



(a) (b)

Figure 30: (a) Linear trend comparisons of the Group 1 high and low T:H calculations. (b) Linear trend comparisons of the Group 2 high and low T:H calculations.

7 Interpretation of Results

8 Conclusion

9 References

[1] http://www.mathworks.com/products/matlab
[2] http://www.vivoquant.com
[3] Aaron Adcock, Daniel Rubin, and Gunnar Carlsson. Classification of Hepatic Lesions using the Matching Metric. *Computer Vision and Image Understanding*, Vol. 121 (2012).
[4] Herbert Edelsbrunner and John L. Harer. Computational Topology. American Mathematical Society, Providence (2009).
[5] Khaled Greish. Enhanced Permeability and Retention (EPR) Effect for Anticancer Nanomedicine Drug Targeting. *Cancer Nanotechnology, Methods in Molecular Biology*, 624. Humana Press (2010).
[6] Carola Heneweer, Jason P. Holland, Vadim Divilov, Sean Carlin, and Jason S. Lewis. Magnitude of Enhanced Permeability and Retention Effect in Tumors with Different Phenotypes: 89Zr-Albumin as a Model System. *The Journal of Nuclear Medicine:* Vol. 52 No. 4(2011).
[7] Jack Hoppin, Kelly Davis Orcutt, Jacob Y. Hesterman, Matthew D. Silva, Dengfeng Cheng, Christian Lackas, and Mary Rusckowski. Assessing Antibody Pharmacokinetics in Mice with In Vivo Imaging. *The Journal of Pharmacology and Experimental Therapeutics* (2010).
[8] Wenhui Ma, Zhe Wang, Weidong Yang, Xiaowei Ma, Fei Kang, and Jing Wang. Biodistribution and SPECT Imaging Study of Tc-99m Labeling NGR Peptide in Nude Mice Bearing Human HepG2 Hepatoma. *Biomed Research International* (2014).
[9] Robert MacPherson and Benjamin Schweinhart, Measuring Shape with Topology. *Journal of Mathematical Physics*, 53, 073516 (2012).

[10] John Milnor. Morse Theory. Princeton, NJ: Princeton: University Press (1963).
[11] John Neter, William Wasserman, and Michael H. Kutner. Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs. Homewood, IL: Irwin (1990).
[12] Gary W. Oehlert. A first course in design and analysis of experiments. New York: W.H. Freeman (2000).
[13] Ganga Srinivasan and Sagar Mane. EPR effect: Promising approach for tumor targeted drug delivery. *World Journal of Pharmaceutical Sciences* (2014).
[14] Andrzej Szymczak, Arthur Stillman, Allen Tannenbaum, and Konstantin Mischaikow. Coronary vessel trees from 3D imagery: A topological approach. *Medical Image Analysis*, 10(4), 548-559 (2006).
[15] Afra Zomorodian. Topology for Computing. Cambridge University Press, Cambridge (2009).