University of Massachusetts Boston
ScholarWorks at UMass Boston

Graduate Masters Theses | Doctoral Dissertations and Masters Theses

5-31-2016

Analysis of Gompertzian Growth in Aggregating Multicellular Tumor Nodules

Gwendolyn A. Deger
University of Massachusetts Boston

Follow this and additional works at: http://scholarworks.umb.edu/masters_theses

Part of the Biophysics Commons, and the Physics Commons

Recommended Citation
Deger, Gwendolyn A., "Analysis of Gompertzian Growth in Aggregating Multicellular Tumor Nodules" (2016). *Graduate Masters Theses.* Paper 370.

This Open Access Thesis is brought to you for free and open access by the Doctoral Dissertations and Masters Theses at ScholarWorks at UMass Boston. It has been accepted for inclusion in Graduate Masters Theses by an authorized administrator of ScholarWorks at UMass Boston. For more information, please contact library.uasc@umb.edu.

ANALYSIS OF GOMPERTZIAN GROWTH IN AGGREGATING MULTICELLULAR TUMOR NODULES

A Thesis Presented

by

GWENDOLYN A. DEGER

Submitted to the Office of Graduate Studies,
University of Massachusetts Boston,
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2016

Applied Physics Program

© 2016 by Gwendolyn A. Deger
All rights reserved

ANALYSIS OF GOMPERTZIAN GROWTH IN AGGREGATING MULTICELLULAR TUMOR NODULES

A Thesis Presented

by

GWENDOLYN A. DEGER

Approved as to style and content by:

_____
Jonathan Celli, Assistant Professor
Chairperson of Committee

_____
Chandra Yelleswarapu, Associate Professor
Member

_____
Stephen Arnason, Associate Professor
Member

_____
Stephen Arnason, Program Director
Applied Physics Program

_____
Bala Sundaram, Chairperson
Physics Department

ABSTRACT

ANALYSIS OF GOMPERTZIAN GROWTH IN AGGREGATING MULTICELLULAR TUMOR NODULES

May 2016

Gwendolyn A. Deger, B.A., Grinnell College
M.S., University of Massachusetts Boston

Directed by Professor Jonathan Celli

Past studies have shown that tumor growth generally follows an exponential growth function or, with a limiting growth constraint, the sigmoid Gompertzian function, where a terminal tumor size is reached at late times. The classical Gompertzian description of tumor growth applies in the case of two-dimensional (2D) *in vitro* cell studies due to the effect of physical limitations on possible growth area. This project asked whether Gompertzian form applies to the *in vitro* growth of multifocal 3D tumor nodules, whose size is determined by aggregation events as well as cell proliferation. Previous reports have indicated that these three-dimensional (3D) spheroids appear to reach a terminal size, even though the full available 3D volume is not occupied. In this scenario it is not immediately obvious if individual nodules are growth-constrained by nutrient or oxygen diffusion, or rather if the

ensemble of all nodules exhibits Gompertzian form. 3D *in vitro* ovarian cancer cells were chosen as the population to be studied. The ovarian cancer cells were grown in overlay on a laminin-rich extracellular matrix (ECM). This model system is a common and widely used cell culture platform in cancer cell research. Using this system, division of the ovarian cancer cells into heterogeneous clusters that aggregate into larger clusters, and then reach a steady bimodal distribution of small and large aggregates, was observed. The average volume as well as the total volume of these two cell aggregate groups were measured over time to determine the nodules' growth behavior would plateau without a growth area limitation. Biological processes may limit the size and behavior of cells within sphere-like multicellular nodules differently than a simple layer of cells on a petri dish. The standard deviation of the rapidly growing nodule volume population within a 3D *in vitro* ovarian cancer sample was shown to grow according to a quadratic function, while the population of small nodules stays constant over time. The overall growth behavior of the total volume of the rapidly growing nodules was Gompertzian. The spread between the increasing average size of the large and growing nodule population and the constant average size of the population of small nodules increased exponentially. A particle velocity tracking program was used to search for a relationship between the lateral velocity of the nodules within the field of view and the average size of the rapidly growing nodules. The average lateral velocity of all nodules was shown to weakly decrease over time. This indicates that the behavior of 3D grown ovarian cancer cells follow a dissemination pattern in which small cells or nodules of ovarian cancer cells demonstrate higher dissemination than large nodules. The motion of smaller cell

nodules or single cells may be advantageous in the *in vivo*, as well as the *in vitro* settings. This advantage may produce the bimodal distribution of mobile small aggregates and large slow-moving and growing aggregates, and in turn, this behavior may demonstrate that dissemination of small aggregates of ovarian cancer cells occurs in a 3D environment.

ACKNOWLEDGEMENTS

I am grateful to my advisor, Professor Jonathan Celli, whose expertise and understanding made it possible to work on a topic of great interest to me.

I would like to express my gratitude to my committee: Professor Jonathan Celli, Professor Stephen Arnason, and Professor Chandra Yelleswaparu, for their guidance and support.

I would also like to thank all my professors at the University of Massachusetts Boston. I highly value all that I have learned and been exposed to during my time there.

I also am very grateful to my family and friends for their support.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES x

CHAPTER Page

1. BACKGROUND 1
Motivation and Goals of the Study 1
MATLAB Volume Analysis 5
MATLAB Velocity Analysis 5
Dark Field Microscopy 6
Ovarian Cancer Cell Behavior 7

2. MATERIALS AND METHODS 9
Cell Culture and Microscopy Procedure 9
Matlab Image Volume Analysis Procedure 17
Matlab Image Velocity Analysis Procedure 24

3. GOMPERTZIAN BEHAVIOR OF TUMOR VOLUME 27
Total Volume vs. Time Behavior 27
Average Volume vs. Time Behavior 28
Interpretation of Parameters and Physical Constraints 30

4. TIME EVOLUTION OF VOLUME DISTRIBUTION 31
Aggregation Kinetics 31

5. INSIGHTS FROM TIME LAPSE VIDEO ANALYSIS 34
Volume Changes at Short Times 34
Visualization of Aggregation Events 36

6. DISCUSSION 39
Relationships between measured values 39
Error and Reproducibility 41
Potential Implications and Conclusions 42

APPENDIX

A. MATLAB PROGRAM: “VolDataGather.m” 44

B. MATLAB PROGRAM: “VolDataGatherShortTerm.m” 51

C. MATLAB PROGRAM: “PIVOrganizer.m” 54

REFERENCE LIST 58

BIOGRAPHICAL SKETCH OF AUTHOR 61

LIST OF FIGURES AND TABLES

Figure Page

1. Population Count Definition 7
2. Combination Definition 10
3. Population Increase Definition 11
4. ImageJ Thresholding of Short-term Photograph..... 14
5. ImageJ Thresholding of Long-term Photograph..... 14
6. Darkfield Microscopy Photograph Plate 2 Day 2 16
7. Plate 4 Field 1 Time 10 minutes 17
8. Plate 4 Field 1 Time 20 minutes 17
9. File Organization Convention..... 20
10. Total Volume vs. Time for Plates 1, 2, and 3 27
11. Curve fit of Total Volume vs. Time for Plates 1, 2, and 3..... 28
12. Average Volume vs. Time for Plates 1, 2, and 3 29
13. Curve fit of Average Volume vs. Time for Plates 1, 2, and 3 29
14. Volume Histogram at Multiple Times for Plates 1, 2, and 3 32
15. Average Volume Behavior Over Time for Plates 1, 2, and 3 32
16. Curve fit of Standard Deviation of 2nd Mode vs. Time (1, 2, 3)... 33
17. Analysis of Standard Deviation of 2nd Mode vs. Time (1, 2, 3) ... 33
18. Average Volume vs. Time for Plate 4 34
19. Curve fit of Average Volume vs. Time for Plate 4..... 35
20. Total Volume vs. Time for Plate 4..... 35

Figure Page

21. Curve fit of Total Volume vs. Time for Plate 4..... 36

22. Velocity Average vs. Time for Plate 4..... 37

23. Velocity Histogram at Multiple Times for Plate 4..... 38

24. Analysis of Velocity Histogram at Multiple Times for Plate 4 38

Table Page

1. Distribution of Measured Volumes for Plates 1, 2, 3 vs. Time .. 31

CHAPTER 1

BACKGROUND

Motivation and Goals of the Study

The primary question addressed by this project was: how does the classical Gompertzian description of tumor growth apply to the case of multifocal tumor nodules? The primary goal of this project was to use MATLAB to analyze the changes in volume and lateral speed across the field of view of cell clusters in a specific set of ovarian cancer cell photographs, explores the growth behavior of both the collective and the individual masses within a 3D *in vitro* system. A secondary goal is to examine the behavior of the volume frequency distribution over a long period of time (in the form of a volume histogram). Another secondary goal is to examine the changes in the sum of all volumes measureable at each time over time. Another secondary goal is to examine the changes in the average of all volumes measureable at each time over time.

Past studies have shown that cancer tumor growth generally follows an exponential growth function or, with a limiting growth factor, the sigmoid Gompertzian function (Edinger, Sweeney, et al. 1999; Greenspan 1972; Johnson, Edwards, et al. 2007; Ward 1997; Tomlinson & Bodmer 1995; Swan 1990; Zelen 1966; Norton 1988; Marušić, Vuk-Pavlovic, et al. 1994; Kozusko & Bajzer 2003).

The Gompertzian function is characterized by initial exponential growth that plateaus and is constant from that time point onwards. A form of the Gompertzian that has been used to describe *in vivo* metastatic tumor growth in various host organisms in the past is:

$$f(t) = e^{a(1-e^{-bt})}$$

f(t) indicates the number of cells or their weight after time *t*, and *a* and *b* are experimental coefficients determining slope of the curve (Laird, Tyler, et al. 1965; Laird 1965; Bajzer, Vuk-Pavlović, et al. 1997).

In this study, a form of the Gompertzian equation was used to describe the relationship between the initial volume of a nodule (V_o) and the final volume of a nodule (V) after time t.

$$V = V_o e^{(\frac{A}{\alpha}(1-\mathrm{e}^{-(\alpha)}))}$$

In this case A and α are experimental coefficients determining slope of the curve and volume at which the nodule growth plateaus.

Cell motility is also a function of the growth environment (Chicoine and Silbergeld 1995; Johnson, Leight, et al. 2007). The classical Gompertzian description of tumor growth applies in the case of 2D *in vitro* cell studies due to the effect of physical limitations on possible growth area. This project explored the possibility that a classical Gompertzian also describes the growth of multifocal tumor nodules grown in an environment without limitations on possible cluster shape behavior and growth area. 3D *in vitro* ovarian cancer cells were chosen as the population to be studied. The ovarian cancer cells were grown within an extracellular matrix. This model system, cells overlaid

on a bed of extracellular matrix, is a common and widely used cell culture platform in cancer cell research (Nelson & Bissell 2006).

Biological processes, such as the limitation of the size of a hypoxic center or limitations on the vasculature able to infiltrate and deliver nutrients to the central cells of a nodule, may limit the size and behavior of cells within sphere-like multicellular nodules differently than a simple layer of cells on a petri dish (Castro, Klamt, et al. 2003).

Multicellular nodules have been shown to aggregate while also maintaining a bimodal distribution of small and large aggregates (Celli, Rizvi, et al. 2010). The first mode is constant at all times and the second increases over time. Describing the pattern in the noticeable increase over time of the standard deviation of the second mode of the volume frequency distribution is a goal of this project. As the standard deviation increases, the range of different nodule sizes increases (Celli, Rizvi, et al. 2010). Nodule size has been found to be dependent on its surroundings (Tchafa, Shaw, et al. 2012), so it is useful to understand how multiple nodules interacting within a 3D *in vitro* environment (similar to an *in vivo* environment) may affect one another and possibly determine the behavior of their neighbors.

A motivation for looking at the behavior of the sum of volumes at each time over time is that it reflects the population of cancer cells within the *in vitro* environment in which they were grown. The ovarian cancer cells studied in this project were grown in a 3D *in vitro* method. The growth behavior of cells grown in a 2D configuration is known and well-studied, but the behavior of cells grown in a 3D configuration is not fully understood (Lü, W., Zhang 2014). The growth behavior of cells grown in a 2D configuration grow exponentially until they fill their 2D container and then they stop

growing. The Gompertzian curve has been studied as a representation of cancer cell population behavior, creating a motivation to try a fit of a Gompertzian curve on the resulting total sum of volumes versus time (Laird 1964).

The motivation for studying the behavior of the average of all volumes measured at each time over time is that this behavior reflects the clumping behavior of individual cells into nodules or conversely, the disseminating of individual cells from a larger tumor nodule separately. Ovarian cancer cell behavior is not yet fully understood. Because understanding the behavior of ovarian cancer cells can be useful to treatment planning, it is useful to study the behavior of ovarian cancer cells *in vitro* in an *in vivo*-like 3D configuration for the sake of defining that behavior.

Another motivation for creating useful MATLAB image analysis scripts and programs is that they can be modified to run on a large amount of data at once, with data in the form of photographs of objects. The MATLAB scripts and programs used to analyze the ovarian cancer cell photographs may be useful for analyzing other simple photographs of other objects, or for analyzing many ovarian cancer cell photographs at once in an effort to reproduce this project's results on a larger scale.

In summary, the goal for this project is to measure the behavior of the sum of the volume of cancer cells within their changing biological 3D environment and to measure the behavior of the lateral speed across the field of view of cancer cells within the changing biological 3D environment.

MATLAB Volume Analysis

MATLAB was chosen as the image analysis tool because it allows the user to write scripts and programs that can identify simple shapes against a simple background. The images of ovarian cancer cells used were grayscale, and the tumor nodules were faint but distinct against a mostly black or gray background. The MATLAB program was able to threshold the grayscale photos and cut out the background, highlighting the tumor nodules, using an algorithm known as Otsu's method.

Otsu's method works by assuming that each value in a grayscale picture is assigned a number so that a histogram of color (described by a number for each shade) vs. pixel count describes the frequency of each shade. Otsu's method assumes that the histogram is bimodal (meaning that it has two peaks). It also assumes that one mode represents the background and the other mode represents the foreground. By selecting for only background shades, the background can be isolated within the image, or if the shades of the foreground (and in this project's case: cancer cell shapes) are selected, the shape of the foreground images can be isolated (Otsu 1979). Once the foreground images are isolated, we can use MATLAB to count the number of pixels within each individual continuous nodule (object) and use this to describe the cross sectional area of each spherical cancer nodule.

MATLAB Velocity Analysis

PIVlab is a MATLAB program created by William Thielicke and available online from the MathWorks File Exchange website

(http://www.mathworks.com/matlabcentral/fileexchange/27659-pivlab-time-resolved-particle-image-velocimetry--piv--tool). PIVlab tracks isolated images within a set of photographs taken over time.

PIVlab works by using particle image velocimetry (PIV), "tracking" and "identifying" assemblies of foreground objects and measuring their displacement between two images. The process of particle image velocimetry takes place when individual objects cannot be tracked successfully, as opposed to particle tracking velocimetry (PTV), which takes place when individual objects can be tracked. The general method of a PIV program is: the objects are identified as masses (individual or nodules). Each individual particle can be matched with many others from photo to photo, but the incorrect matches only add up to noise while the true match that describes the overall displacement dominates this sum and stands out as a vector (lateral velocity vector).[1]

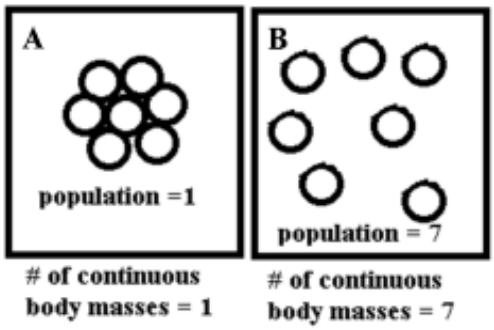
Dark Field Microscopy

The photos of ovarian cancer cells were taken with a dark field microscope, which produces photographs with dark backgrounds and light colored objects. This happens because the light source of a dark field microscope is angled in a way that the objective lens of the microscope does not directly collect that light. The samples themselves show up as bright shapes on the dark background because only the light diffracted by these objects reaches the eye piece. Because the background is dark and the objects are brighter

and nearer to white, dark field microscopy photographs can easily be thresholded to reverse the shades and leave the background white and the objects black and grayscale (Davidson & Abramowitz 2016).

Ovarian Cancer Cell Behavior

Let *organism* be defined as a continuous body mass made up of one or more bodies. Let *population* be defined as the total number of organisms. Examples of population counts are shown in Figure 1.



(Figure 1: The population count for A is 1 because all seven visible individual bodies are combined within a continuous body mass. The population count for B is 7 because all seven visible individual bodies are separated by space.)

For the special case of the cancer cell population studied, acini are defined as specific types of continuous bodies composed of greater than one ovarian cancer cell. Nodule and acini will be used interchangeably in this project. In this case, the ovarian cancer cell is the specific case unit of the continuous body. The type of ovarian cancer being studied in this exercise is epithelial ovarian cancer.

Ovarian cancer is the foremost cause of death from gynecological cancer in the developed world. In the United States, 21,290 new diagnoses of ovarian cancer and

14,180 deaths were expected in 2015. The overall 5-year survival rate for women with ovarian cancer is 45.6% (Howlader, Noone, et al. 2016). About 80% of patients with ovarian cancer present with metastatic disease. The epithelial cells of the ovary constitute 1% of the total ovarian mass but constitute 90% of the ovarian neoplasms. Epithelial ovarian cancer (EOC) spreads initially by direct extension into adjacent organs, especially the fallopian tubes and uterus, with occasional spread into the rectum, bladder and pelvic side wall (Farghaly 2013). Ovarian cancer is known to spread throughout the peritoneal cavity by means of "seeding" (Delong and Burkhart 2008).

Seeding is the exfoliation of cells into the peritoneal cavity. Therefore, it seems that ovarian cancer cells can, in certain circumstances, shed individual cells that move off from a larger tumor if the proper circumstances occur. We can infer information about the behavior of average volume over time and the average lateral velocity over time using "seeding" as a base model of behavior. Circumstances in which the tumors stop growing and acquiring new individual cancer cells from their surroundings as a method of growth must happen in a way that although ovarian cancer cells divide and grow exponentially, their average latera velocity may stay constant. It may be that as most of the nodules get larger and slower, they still release smaller 'seed' cells that still may move relatively fast, increasing the average velocity of cancer cells within the peritoneal cavity.

CHAPTER 2

MATERIALS AND METHODS

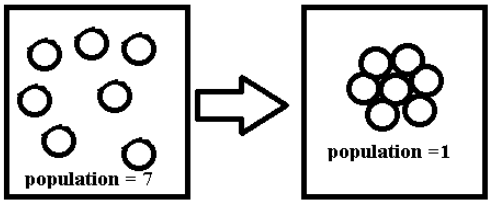
Cell Culture and Microscopy Procedure

Photographs of cells used in this exercise were collected within a previous study. The information on cell culture origin for this specific exercise in image analysis is taken directly from: Celli, J. P., Rizvi, I., Evans, C. L., Abu-Yousif, A. O., & Hasan, T. (2010). Quantitative imaging reveals heterogeneous growth dynamics and treatment-dependent residual tumor distributions in a three-dimensional ovarian cancer model. Journal of Biomedical Optics, 15(5), 051603-051610. doi:10.1117/1.3483903.

> ‘Epithelial ovarian cancer NIH:OVCAR-5 cells were obtained from Thomas Hamilton (Fox Chase Cancer Institute, Philadelphia, Pennsylvania). The cells were grown in RPMI-1640 (Roswell Park Memorial Institute) medium (Mediatech Inc., Herndon, Virginia, USA) supplemented with 10% heat-inactivated fetal calf serum (GIBCO Life Technologies, Grand Island, New York, USA), 100 U⁄mL penicillin, and 100 μg⁄mL streptomycin. For 3-D cell culture, growth factor reduced (GFR) Matrigel (BD Biosciences, San Jose, California) was used as a basement membrane, which has appropriate gel structure and established

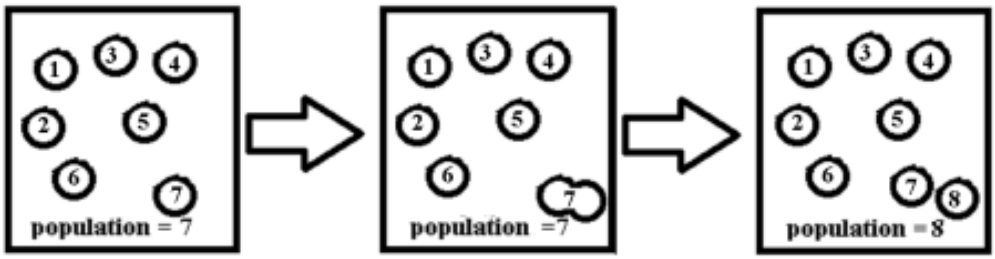
biological activity to promote growth and differentiation of a variety of cells. (Kleinman, McGarvery, et al. 1986) To prepare 3-D cultures 500-μL volumes of NIH:OVCAR-5 cells in single cell suspension of 15,000 cells⁄mL were grown on beds of GFR Matrigel (BD Biosciences) on the glass slide inset of 35-mm MatTek culture dishes (MatTek Corporation, Ashland, Massachusetts, USA). GFR Matrigel beds were initially prepared by ejecting 150 μL of GFR Matrigel solution at ∼4 °C on the chilled MatTek plates and carefully rotating the dish at a slight angle to evenly distribute it over the 10-mm-radius glass portion of the MatTek dishes that were then incubated at 37 °C for 30 min prior to plating cells to allow gelation to occur. Following initial plating of cells, they were allowed to adhere to the Matrigel bed before addition of complete growth medium with 2% GFR Matrigel. All cultures were maintained in an incubator at 37°C in an atmosphere of 5% CO_2.'

Photographs must be collected of the changing population of organisms over time. Population changes were only caused by two specific allowable causes of change in population. The first is the combination of at least one organism in the population with at least one other organism to form a larger acini as shown in Figure 2.



(Figure 2: A reduction in population by combination.)

The second allowable cause of change in population is the separation of at least one organism in the population into multiple whole organisms as shown in Figure 3. In the case of cells, this process is called mitosis and creates two whole daughter cells identical to the parent cell.



(Figure 3: A gradual increase in population by mitosis.)

In summary, the variables are limited to population of organisms versus time and size of organisms versus time, with the sampling area of the photographs remaining constant.

The photographs must be saved as .tif or .tiff files. The following information on cell preparation and imaging for this specific exercise in image analysis is taken directly from the source: Celli, J. P., Rizvi, I., Evans, C. L., Abu-Yousif, A. O., & Hasan, T. (2010). Quantitative imaging reveals heterogeneous growth dynamics and treatment-dependent residual tumor distributions in a three-dimensional ovarian cancer model. Journal of Biomedical Optics, 15(5), 051603-051610. doi:10.1117/1.3483903.

> ‘To characterize development of size distributions over time, 3-D cultures were routinely imaged by dark field microscopy at 5× using a Zeiss Axiovert inverted microscope (Carl Zeiss MicroImaging, Inc.,

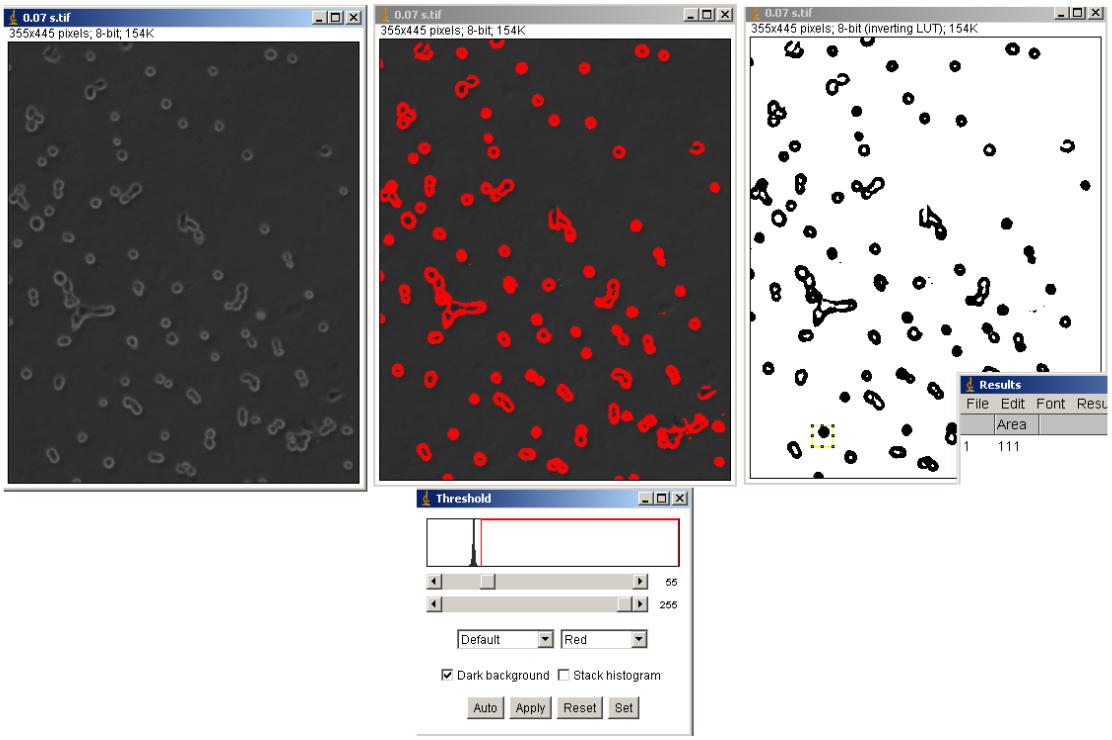
Thornwood, NY, USA) fitted with a QuantiFIRE™ cooled 12-bit monochrome CCD camera with 2048 ×2048-pixel chip (Optronics, Goleta, California). Images were acquired in a focal plane just above the surface of the gel on which 3-D structures are formed. For each plate, five dark-field images of each culture dish were acquired at each time point in the tagged image file format (TIFF) and saved for off-line processing. Time-lapse microscopy sequences were obtained using a Nikon TE2000-S inverted microscope with a 10× phase contrast objective in an enclosed weather station (Nikon Instruments, Inc, Melville, New York, USA).'

Note that, for this specific exercise, the acini formed 3D shapes, but the photographs only collected 2D population data. 3D information was extrapolated by assuming that the 3D shape formed by the acini was spherical. The total volume of each acini was calculated using the radius of the 2D acini shown in the collected photographs.

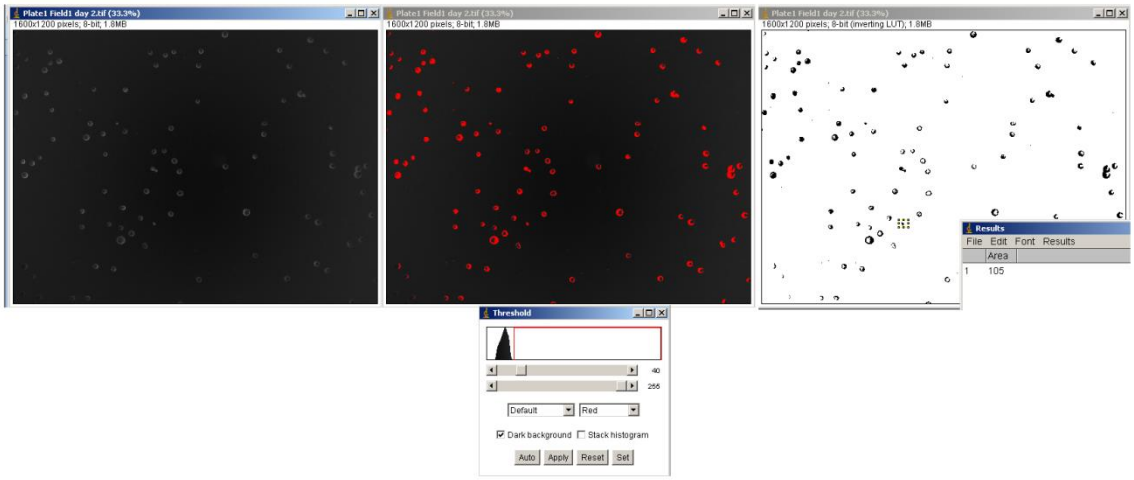
The minimum characteristic size of the specific organism being studied must be determined. The expected size in area units of μm^2 is equal to the variable called "minsize" used within the MATLAB .m files used to analyze images. Minsize must be set manually when the code is run. The minsize of a single ovarian cancer cell of the type and size used for this exercise was 100 μm^2. This value was approximated from manual cell cross sectional area measurements using the program ImageJ. This process is shown in Figure 4 and Figure 5. Photographs taken at early times were chosen as the cells had not yet combined into larger nodules, and the individual cells were visible and distinguishable from nodules.

The process used for calculating the cross sectional area measurements of individual cells was as follows:

1) In ImageJ, open the image file.
2) Through the standard menu bar, choose Image>Adjust>Threshold
3) Adjust the lower slider bar all the way to the right and the upper slider bar just far enough left that the cells or objects, but not the background, are highlighted in red, and choose Apply.
4) Using the arrow, highlight around a single cell or object in your photograph, choose from the menu bar Analyze>Analyze Particles. Within the Analyze Particles options window, using the standard options of size = 0-Infinity and with the 'Display Results' box checked, select OK. The resulting Results window will display the cross sectional area shown in the photograph of the highlighted cell, as measured by ImageJ.



(Figure 4: Example of ImageJ thresholded actual minimum cell sizes at start of short term Plate 4 data)



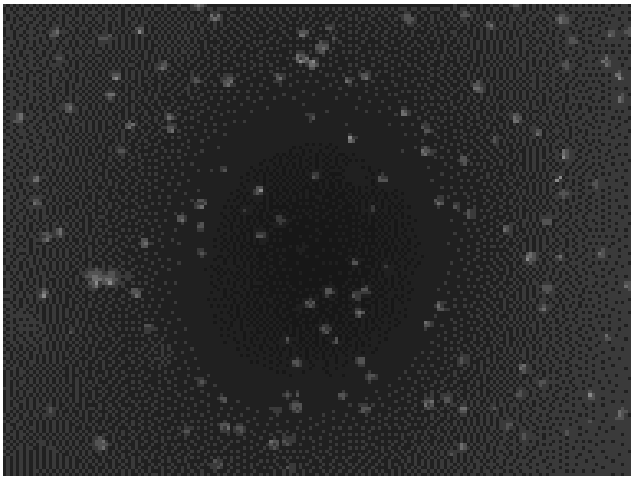
(Figure 5: Example of ImageJ thresholded actual minimum cell sizes on Day 2)

Each .tif file must be labeled as PlateA_FieldB_DayC.tif, where C is the day t the picture was taken, and A and B give location information for the picture within the population's larger ecosystem.

Using the units of time set C, the MATLAB variable “time_scale” is equal to 1 if the time is recorded in the photograph file names in units of days (following the convention described above) and the desired unit of time in the results .mat file is days. Or, ‘time_scale’ is equal to 86400 if the time is recorded in the photograph file names in units of days (following the convention described above) and the desired units of time in the results .mat file is seconds. Or, ‘time_scale’ is equal to 1 if the time is recorded in the photograph file names in units of seconds (not following the convention described above) and the desired units of time in the results .mat file is seconds.

For a general case, the Plate Number identifies the total ecosystem, while Field Number identifies the limited field of view within that total ecosystem. For this specific exercise done on ovarian cancer cells, PlateA, where A is equal to 1, 2, or 3, indicates the different plates of cells used in this specific exercise. A plate is defined as a plastic container in which cells are grown.

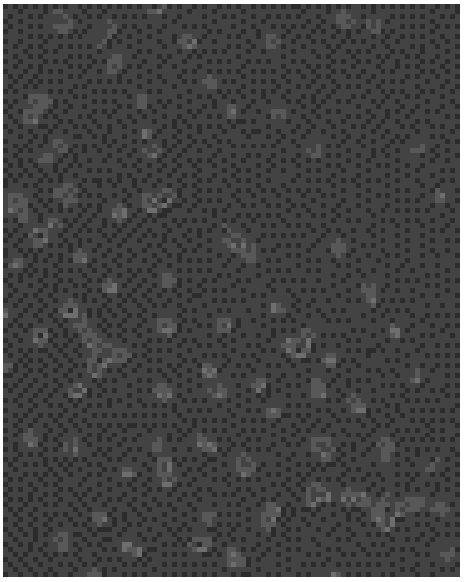
Field is defined as a set portion of the plate and is shorthand for the designated field of view. The field of view must be held constant over time. In this exercise, the field of view is rectangular, as shown in the file Plate2_Field1_Day2.tif as seen in Figure 6.



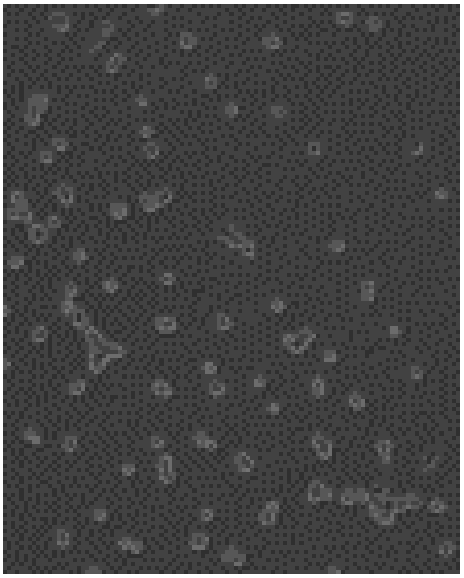
(Figure 6: Darkfield photograph of Field 1within Plate 2 on Day2.)

The variable defined as calfactor is determined by the scale of the photographs taken in units of 1.41 μm/pixel.

For this specific set of data, photographs were taken on Days: 2, 3, 4, 7, 8, 9, 10, 14, 16, 17, 18, and 21 at Plate 1 Field 1, Plate 1 Field 2, Plate 1 Field 3, Plate 1 Field 4, Plate 1 Field 5, Plate 2 Field 1, Plate 2 Field 2, Plate 2 Field 3, Plate 2 Field 4, Plate 2 Field 5, Plate 3 Field 1, Plate 3 Field 2, Plate 3 Field 3, Plate 3 Field 4, and Plate 3 Field 5. 452 photographs were taken of Plate 4 Field 1 at a shorter time interval of every 10 minutes. Examples of Plate 4 Field 1 are shown in Figure 7 and Figure 8.



(Figure 7: Plate 4 Field 1 at t = 10 mins)



(Figure 8: Plate 4 Field 1 at t = 20 mins.)

MATLAB Image Volume Analysis Procedure

A MATLAB program VolDataGather.m (an edited version of the original file: get_acini_info.m by author: Jonathan Celli, PhD, Professor at University of Massachusetts Boston, 2010) was written to analyze the collected cell photographs once sorted into folders labeled by the day the photographs were taken. The photograph .tif file naming convention that works for VolDataGather.m is “Plate# Field# day##.tif”. This convention allows for the .mat file that results from VolDataGather.m to be named as “sizeinfoday##.mat” and is saved in the same folder as the photographs.

The dependent variables for VolDataGather.m are calfactor, minsize, time_scale, and bins. ‘calfactor’ is the calibration sizing factor difference between picture and microscope. ‘minsize’ is the minimum allowable cell cross-sectional area, in this particular case, 100 μm^2 was chosen. ‘time_scale’ is the time conversion factor from days to the desired time scale. ‘time_scale’ is equal to 1 if the time is recorded in the photograph file names in units of days (following the convention described above) and the desired units of time in the results .mat file is days. Or, ‘time_scale’ is equal to 86400 if the time is recorded in the photograph file names in units of days (following the convention described above) and the desired units of time in the results .mat file is days. Or, ‘time_scale’ is equal to 1 if the time is recorded in the photograph file names in units of seconds (not following the convention described above) and the desired units of time in the results .mat file is seconds. ‘bins’ is initially a guess of how many bins will be appropriate in the histogram, and then once the data is collected for all groups of data at each time t, a script is used to calculate the exact bin number necessary, and then the program is re-run on all the data using the number of bins calculated earlier.

How to calculate the value = bins:

1) Run VolDataGather(calfactor,minsize,time_scale,bins) on all groups of .tif files at all times t, with bins = a guess.
2) Load an individual sizeinfo<day#>.mat file and run MaxSize(Day#,1)=max(Vlist(:,1)); on each before loading the next .mat file.
3) Repeat step 2 on all sizeinfo<day#>.mat files.

4) Use the script [a,b]=find(max(MaxSize)); DayofMax=a; binMax=max(MaxSize); on the resulting MaxSize list.
5) Solve the following equation for i: binMax = minVol + baseStep * 2^(i-1)
6) Round the bin number up to the nearest whole number. This rounded up whole number is equal to the accurate value of 'bins'.
7) Run VolDataGather(calfactor,minsize,time_scale,bins) again on all groups of .tif files at all times t, with bins equal to the accurate calculated number.

This value was at time = 18 days and was equal to 160098749.340446 μm^3. Once a sizeinfo<time>.mat file was loaded, the script used to determine the largest nodule volume was MaxDay#=max(Vlist(:,1)). Each sizeinfo <day#>.mat file was loaded separately and MaxDay2, MaxDay3, MaxDay4, MaxDay7, MaxDay8, MaxDay9, MaxDay10, MaxDay14, MaxDay16, MaxDay17, MaxDay18, MaxDay21 were determined and compared.

The number of bins was determined by:

The maximum volume data value was at time = 18 days in the data set this program was originally written for, and was equal to 160098749.340446 $(\mu m)^3$. In this case:

$$Volume_{Max} = BinMin_{\,i} + baseStep_{from\,\min of\ initial\ bin\ to\,\min of\ second\ bin}\ 2^{(i-1)}$$

$$Volume_{Max\ Value} = 160098749.340446\ \mu m^3$$

$$Initial\ Bin\ Start\ Value = \frac{4}{3}\pi(\frac{A}{\pi})^{3/2} = \frac{4}{3}\pi(\frac{100\mu m^2}{\pi})^{3/2}. = 752.2528\mu m^3$$

$$160098749.340446\ \mu m^3 = 752.2528\mu m^3 + 100 * 2\hat{}(i-1)$$

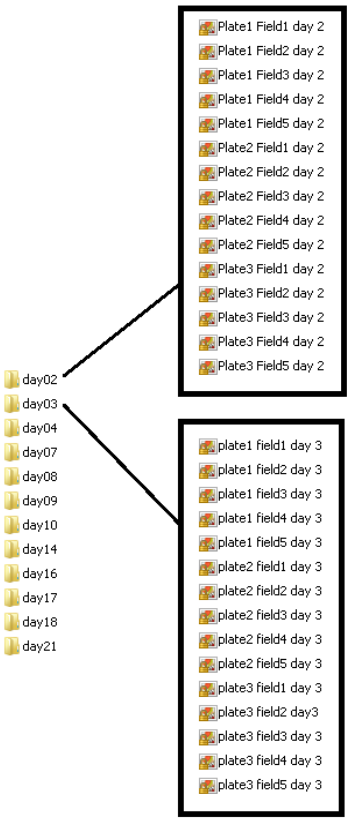
$$i = 21.6105$$

As bin number must be a whole number, 21.6105 was rounded up to 22 bins.

The final input and values used for this particular data were:

VolDataGather(1.41,100,1,22).

The photographs were sorted by day, not field or plate. There were a total of 12 folders, labeled day 2, day 3, day 4, day 7, day 8, day 9, day 10, day 14, day 16, day 17, day 18, and day 21. Each folder had 15 photographs taken at different field and plate spaces, but on a specified day after plate preparation.



(Figure 9: Example File Organization Convention.)

This information on MATLAB analysis for this specific exercise in image analysis, is taken directly from the source: Celli, J. P., Rizvi, I., Evans, C. L., Abu-Yousif, A. O., & Hasan, T. (2010). Quantitative imaging reveals heterogeneous growth dynamics and treatment-dependent residual tumor distributions in a three-dimensional ovarian cancer model. Journal of Biomedical Optics, 15(5), 051603-051610. doi:10.1117/1.3483903.

> 'Image data was processed using custom scripts developed using the Image Processing Toolbox in the MATLAB software package (MathWorks, Natick, Massachusetts, USA). To calculate size distributions from image data we developed a batch analysis routine in which sets of high-contrast dark-field images were thresholded (calling on a built-in automated routine based on Otsu's method), made binary and segmented to identify *in vitro* nodules. The routine would count the number of pixels in each region while rejecting partial features at the edge of each field of view and calibrate sizes to square micrometers.'

After the collected photographs were thresholded and processed using the above method in MATLAB, the MATLAB program VolDataGather.m measured organism sizes in units of square micrometers as 2-D cross-sectional area *A*, which is directly reported from calibration of pixel counts organisms detected. 2D nodule area measurements of size less than 100 μm^2 were rejected from the list of nodule areas, as cells smaller than this were not measured. However, nodules are

approximately spherical and it is a reasonable approximation to calculate equivalent Volume *V* from area by $V = \frac{4}{3}\pi(\frac{A}{\pi})^{3/2}$.

The volume of each acini nodule was added to a list of all nodule volumes and saved in a .mat file that includes data tables of 'Vlist','histogram','Vavg','Vtot'. The output file is named according to the day: sizeinfo <time>.mat. For all individual times at which photographs were taken, there is a separate .mat file. There are a total of 464 individual times at which all the photographs were taken, and consequently 464 separate sizeinfo<time>.mat files. 452 of these times represent the .tif files taken at short intervals of Plate 4 Field 1. Twelve of these times represent the .tif files taken at longer intervals once a day of all fields on Plates 1, 2, and 3.

'Vlist' is a two-column table, where the first column is the measured size of individual nodules and the second column records the time at which the picture of the nodule was taken.

The 'histogram' function produces a size distribution histogram from the list of nodule sizes at each time point. The single histogram produced from data taken at each time is a 2 by 22 matrix where the first column is a list of the minimum edge of each bin and the second column lists the number of organisms whose volume (in units of μm^3) is greater than or equal to the minimum edge of each bin, but less than the minimum edge of the next biggest bin minimum edge. The number of bins was determined manually by determining the largest nodule volume at all times and is entered as a dependent variable. Detailed instructions for determining the dependent variable ‘bins’ value is described above in this section.

‘Vavg’ is a single value for each time. It is the average volume of all volumes of organisms measured for each measurement time. It is calculated after every image at that specific time has been analyzed by using the built-in MATLAB function mean. $Vavg = mean(Vlist)$ calculated the average volume of all organisms detected by the MATLAB image analysis tools at that specific time.

‘Vtot’ is a single value for each time. It is the total volume of all volumes of organisms measured for each day. It is calculated after every image at that specific time has been analyzed by using the built-in MATLAB function *sum*. $Vtot = sum(Vlist)$ calculated the average volume of all organisms detected by the MATLAB image analysis tools at that specific time.

Once the program VolDataGather.m was created, the MATLAB input *VolDataGather(1.41,100,1,22)* (with inputs of calfactor, minsize, time_scale, and bins) analyzed the photographs and created the 12 separate *sizeinfo<time>.mat* files.

VolDataGather.m was copied and pasted into the 12 folders with all the .tif files from times t = [Day 2, Day 3, Day 4, Day 7, Day 8, Day 9, Day 10, Day 14, Day 16, Day 17, Day 18, Day 21] which contained the photographs from Plates 1, 2, and 3.

VolDataGatherShortTerm.m is an edited version of VolDataGather.m and was edited to gather ‘Vavg’ and ‘Vtot’ for each of the 452 times *t* of photographs taken of Plate 4 Field 1. The difference between VolDataGatherShortTerm.m and VolDataGather.m are only in how they organize, list, and save the measured data. The result of VolDataGatherShortTerm.m was a 2x452 matrix where the first column was time and the second column was Vavg (the average organism volume) at that time, and

2x452 matrix where the first column was time and the second column was Vtot (the total sum of organism volume) at that time.

VolDataGather.m can be found in Appendix 1 and VolDataGatherShortTerm.m can be found in Appendix 2.

MATLAB Image Velocity Analysis Procedure

A prepared PIV (particle image velocimetry) MATLAB program was found and downloaded from http://www.mathworks.com/matlabcentral/fileexchange/27659-pivlab-time-resolved-particle-image-velocimetry--piv--tool. The program was written by William Thielicke. An excerpt from the above website states:

> "PIVlab is a time-resolved particle image velocimetry (PIV) software that does not only calculate the velocity distribution within particle image pairs, but can also be used to derive, display and export multiple parameters of the flow pattern. A user-friendly graphical user interface (GUI) makes PIV analyses and data post-processing fast and efficient."

PIVlab was used to process the images taken at Plate 4 Field 1 only. Velocity image analysis was not done on the other fields because the photographs were not taken at short enough intervals to give meaningful data. The photographs taken at Plate 4 Field

1 were taken every 10 minutes and the motion of the organisms during this time was continuous enough to allow velocity analysis to be useful.

The PIVlab package was designed to output a text file of velocity vectors v (y direction) and u (x direction) from some origin x,y on a grid predetermined by the program. These vectors described the motion of organisms between each frame. Since there were 452 frames, there were a total of 451 text files that represented the 451 transitions between frames.

A custom script called PIVOrganizer.m was written to analyze and organize this data. The magnitude of each velocity vector was calculated from its x and y components. The outcome of PIVOrganizer.m was a .mat file that included ‘mhistogram’, ‘mavg’, ‘mmax’.

‘mavg’ is a 451x1 matrix where the first value is the average magnitude of the velocity of all moving nodules between frame 1 and 2, the second value is the average magnitude of the velocity of all moving nodules between frame 2 and 3, and so on. ‘mhistogram’ is a histogram of the magnitudes of the cells’ velocities in units of pixels/transition between frames for each 451 text files for a total of 451 histograms. A single value average velocity magnitude was calculated at each time t. ‘mmax’ is a 451x1 matrix made up of single value maximum velocity magnitudes calculated at each time t, where the first value is the maximum magnitude velocity of the all moving nodules between frame 1 and 2, the second value is maximum magnitude velocity of the all moving nodules between frame 2 and 3, and so on.

The input for PIVOrganizer is ‘bins’. ‘bins’ is guessed at initially, and then adjusted to allow for the average values to be visible when the resulting data is graphed.

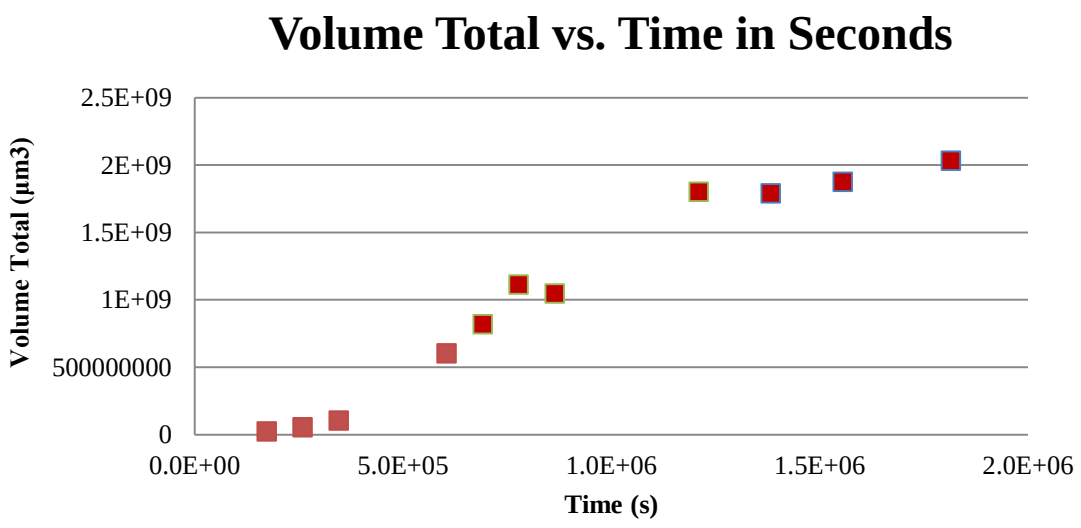
PIVOrganizer.m can be found in Appendix 3. In this case, bins = 20 was used, and with PIVOrganizer.m in the same folder as the .mat files produced by the downloaded PIVlab package, the resulting input into MATLAB was “PIVOrganizer(20)”.

CHAPTER 3

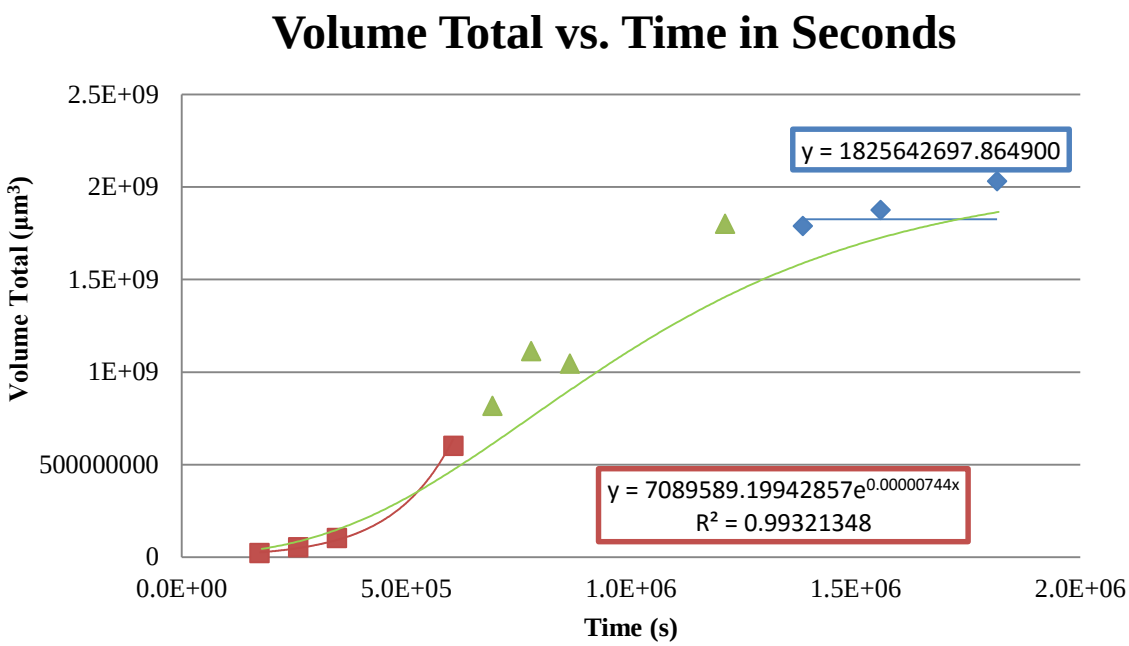
GOMPERTZIAN BEHAVIOR OF TUMOR VOLUME

Total Volume vs. Time Behavior

From Figure 11, it is apparent the total volume measured by the MATLAB program on Plates 1, 2, and 3 seemed to increase and begin to level off. The Total Volume (sum of all volumes recorded at each time) for Plates 1, 2, and 3 increases faster near lower times t and slower near higher times t. Because a Gompertzian function follows this same pattern, through trial and error, and starting by approximating the function as exponential at early times t and constant at later times t, a Gompertzian function was fit to all of the data at all times.



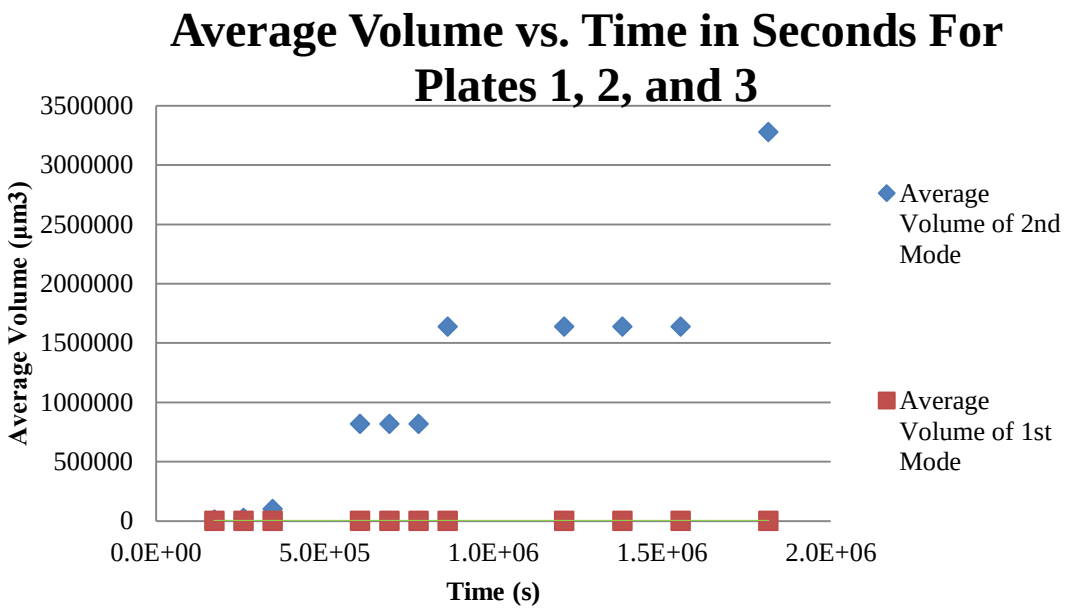
(Figure 10: Total Volume vs. time for Plates 1, 2, and 3.)



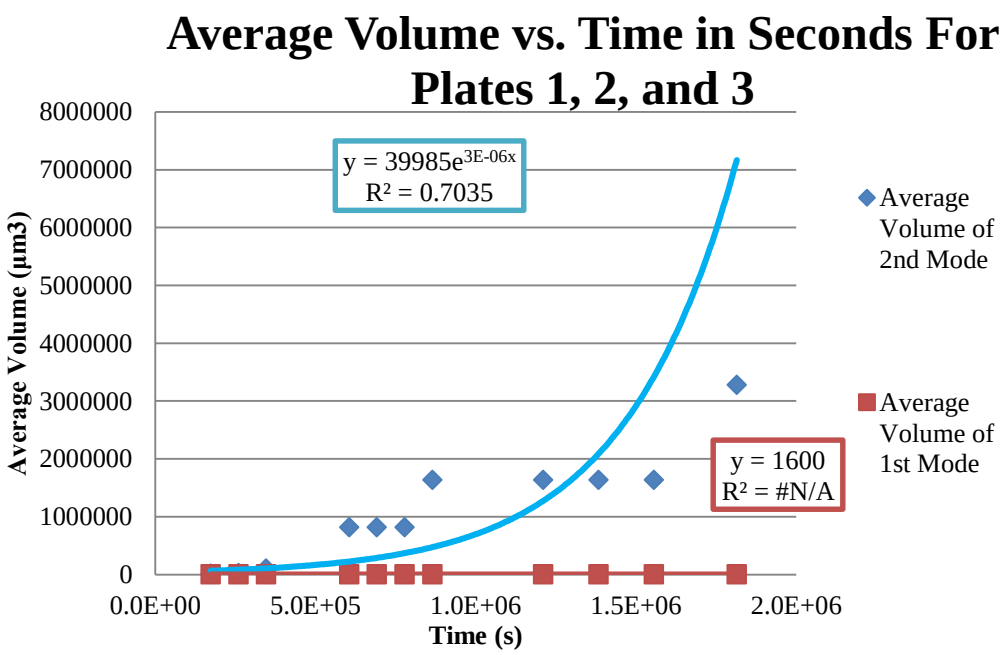
(Figure 11: Curve fit of Total Volume vs. time for Plates 1, 2, and 3. The Gompertzian equation was fit to the Volume Total vs. time manually in the form of: $V = V_o e^{(\frac{A}{\alpha}(1-e^{-(\alpha)}))}$, where $V_o = 7089589.19942857\mu m^3$, $\alpha = .0000024$, and $A = .0000127$.)

Average Volume vs. Time Behavior

From Figure 13, it is apparent that the first mode of the volume histograms that described the long-term behavior of the nodules and cells on Plates 1, 2, and 3 was constant over time, the second mode of the volume histograms increased over time, and the average volume increased over time.



(Figure 12: This is the resulting data from MATLAB analysis that measured average volume vs. time, and recorded a list of volumes at each time for Plates 1, 2, and 3. The Average Volume of the 1st and 2nd modes was found by analyzing and organizing the histograms within Excel.)



(Figure 13: This is the resulting curve fit of the Average Volume for Mode 1 and 2 at each time nodule size was measured at Plate 1, 2, and 3.)

Interpretation of Parameters and Physical Constraints

An exponential growth curve fit was used on the Average Volume of the 2nd Mode data to compare the behavior of the growing population of nodules to exponential growth. The Average Volume of the 1st mode was clearly constant. The first mode average volume was fitted at 1600 μm^3, while the second mode average volume increased over time. At lower times, there are more cells with a small average volume. At higher times, there are less cells with a larger average volume.

Possible physical constraints acting on this system are the size of the plate in the x and y directions and the size of the scaffolding used to grow the cells in the z direction.

CHAPTER 4

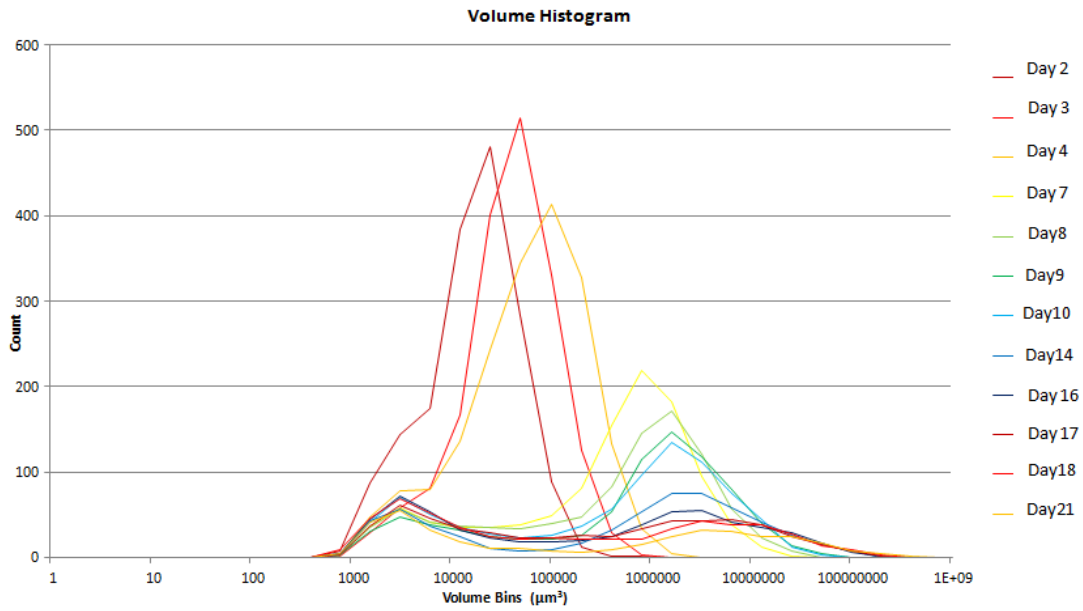
TIME EVOLUTION OF VOLUME DISTRIBUTION

Aggregation Kinetics

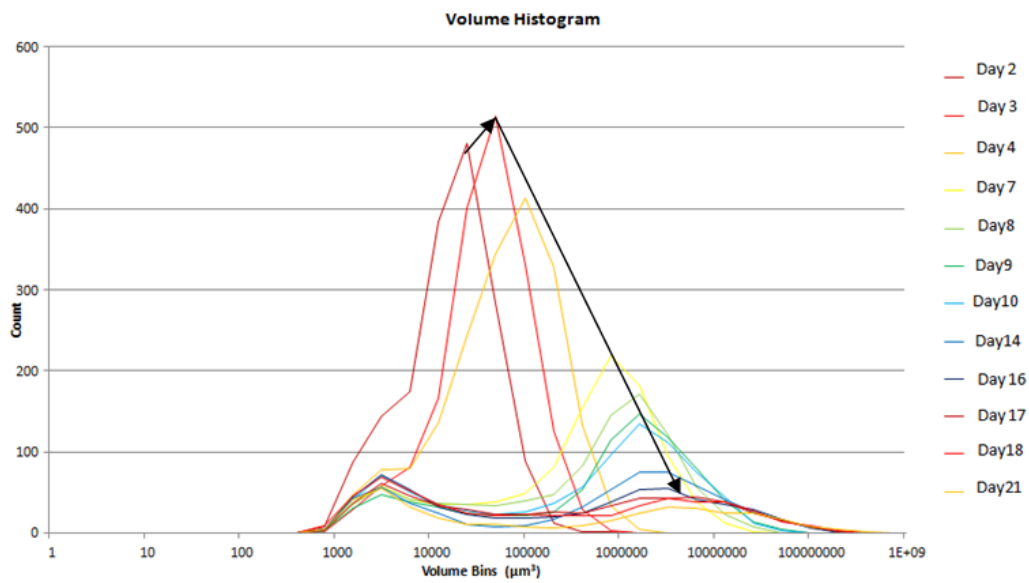
'histogram' (a matrix of multiple histograms at different times of volume measurements) created by VolDataGather.m was exported to Excel. The aggregation behavior of each mode was described by looking at the average of each mode (the first and second peak of each histogram) as highlighted in Table 1 below and shown in Figures 14 and 15. Figure 15 also demonstrates the increasing trend of the mean nodule volume and the bimodal nodule volume distribution.

Histogram bin values in units of *micrometers*^{3}	# on Day 2	# on Day 3	# on Day 4	# on Day 7	# on Day 8	# on Day 9	# on Day 10	# on Day 14	# on Day 16	# on Day 17	# on Day 18	# on Day 21
200	0	0	0	0	0	0	0	0	0	0	0	0
400	0	0	0	0	0	0	0	0	0	0	0	0
800	14	3	5	7	6	7	7	9	8	7	17	9
1600	160	54	90	67	65	33	77	78	82	65	76	75
3200	128	61	66	48	49	42	63	36	62	57	63	36
6400	220	100	92	36	33	34	38	36	45	36	40	27
12800	550	235	179	35	41	30	31	13	20	31	31	11
25600	412	566	308	35	30	18	25	7	25	28	18	10
51200	153	463	381	42	38	26	22	8	12	19	26	11
102400	23	199	446	57	42	18	30	9	24	26	21	3
204800	2	52	208	105	54	33	42	24	17	26	22	8
409600	2	7	59	205	110	74	71	41	32	22	21	10
819200	0	0	8	252	182	157	122	66	45	44	21	21
1638400	1	0	0	133	162	136	146	84	63	43	45	27
3276800	0	0	0	56	79	99	78	67	47	44	42	36
6553600	0	0	0	21	34	61	72	48	35	45	35	26
13107200	0	0	0	3	13	19	18	36	34	30	41	23
26214400	0	0	0	0	1	8	6	17	23	19	13	25
52428800	0	0	0	0	0	0	0	14	10	11	14	10
104857600	0	0	0	0	0	0	0	1	2	3	4	6
209715200	0	0	0	0	0	0	0	0	1	0	1	3
419430400	0	0	0	0	0	0	0	0	0	0	0	0

(Table 1: The results of 'histogram'. Each histogram on each day has two modes. The first mode maxima is highlighted in blue. The second mode maxima is highlighted in red.)

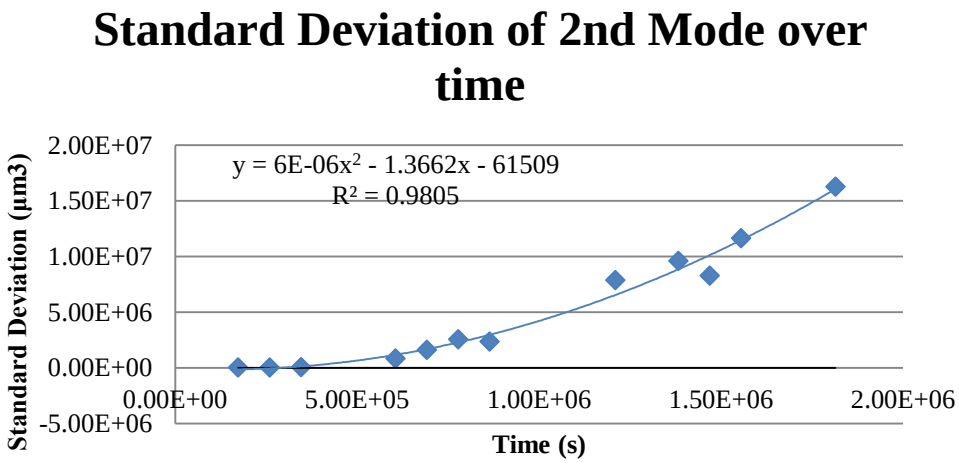


(Figure 14: This is the resulting graph of all volume histograms at all times t for Plates 1, 2, and 3.)

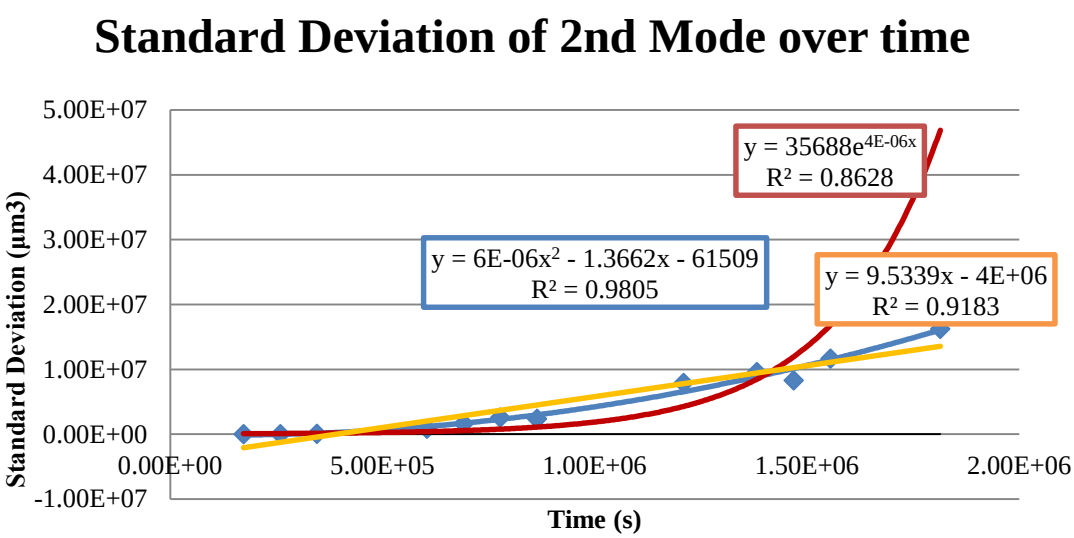


(Figure 15: This is the resulting graph of all volume histograms at all times t. Black arrows have been added to point in the direction of change/time. For Plates 1, 2, and 3.)

Figure 16 and 17 show the resulting curve fits of the standard deviation of the second mode over time of the long term Plates 1, 2, and 3. A square quadratic equation fits the increasing value of Standard Deviation vs. Time. The square quadratic fit had an R^2value of .9805, while an exponential fit had an R^2value of .8628, and a linear fit had an R^2 value of .9183.



(Figure 16: This is the resulting curve fit of. the Standard Deviation of the 2nd mode over time of the long term Plates 1, 2, and 3.)



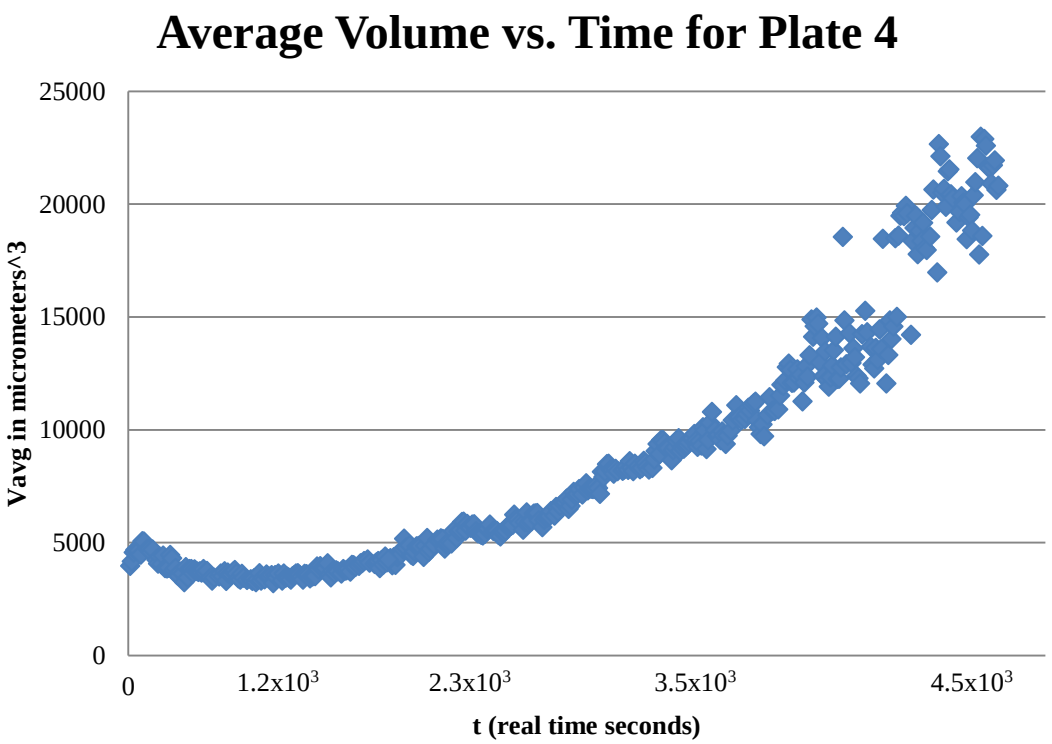
(Figure 17: This figure shows attempts at curve fitting of. the Standard Deviation of the 2nd mode over time of the long term Plates 1, 2, and 3. The yellow fit is the linear attempt, the red fit is the exponential attempt, and the blue fit is the quadratic attempt.)

CHAPTER 5

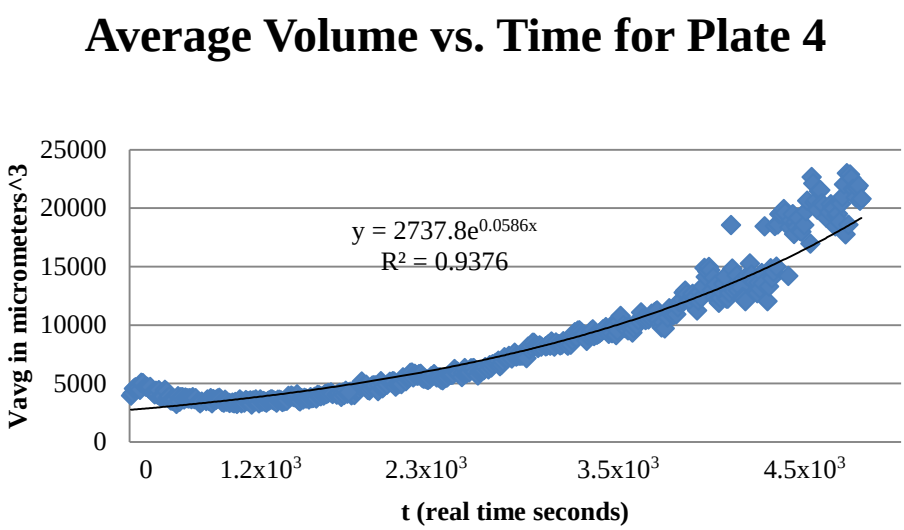
INSIGHTS FROM TIME LAPSE VIDEO ANALYSIS

Volume Changes at Short Times

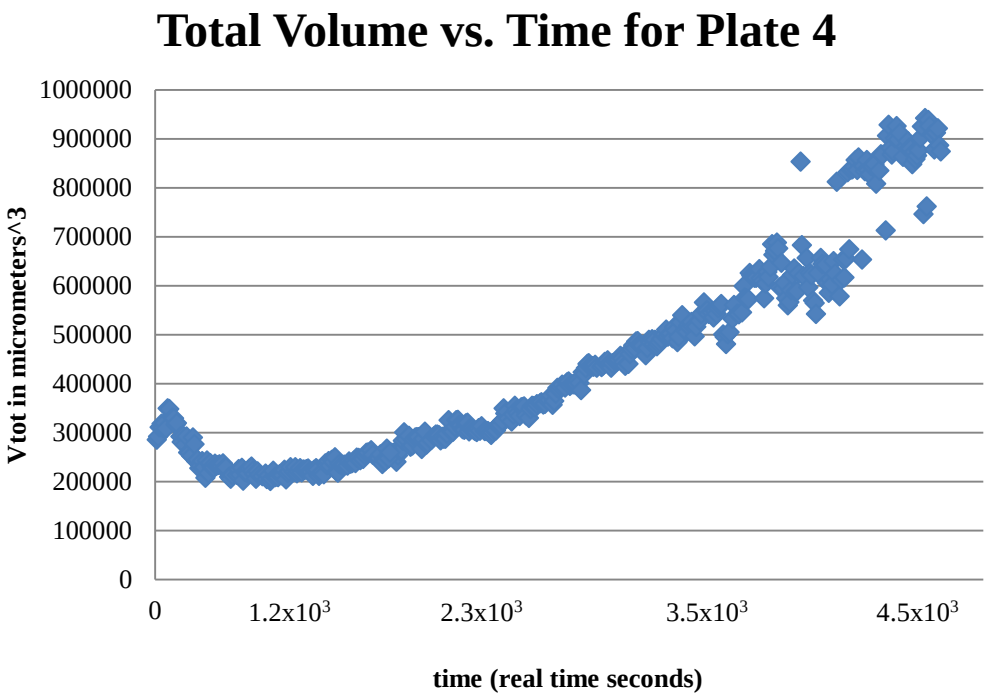
From Figure 18, it is apparent that the average volume increased over time for the short term behavior recorded on Plate 4. From Figure 19, it is apparent that the average volume increased over time for the short term behavior recorded on Plate 4.



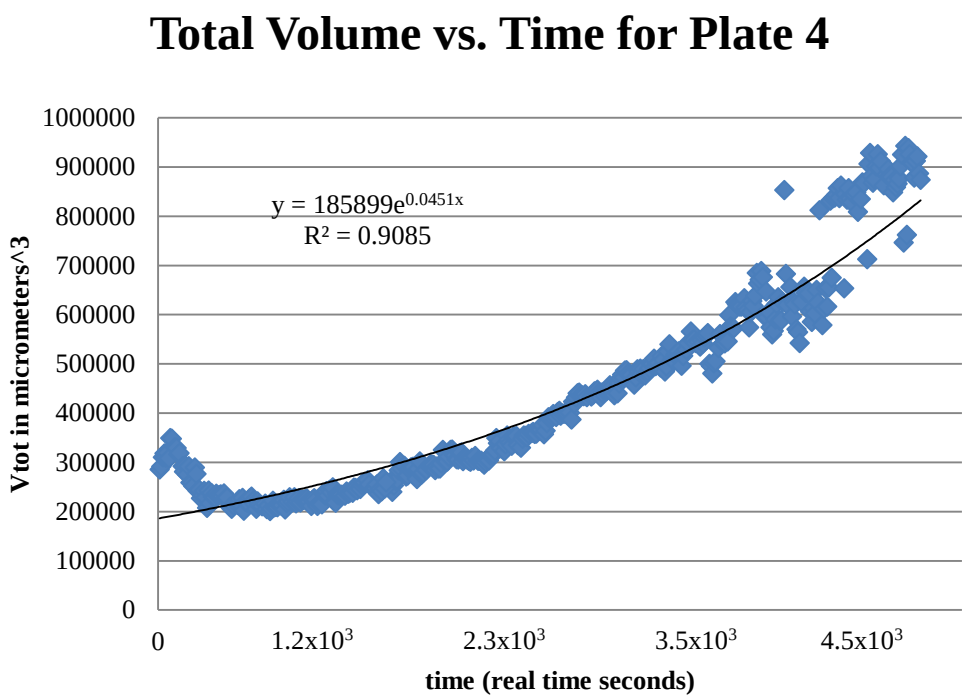
(Figure 18: This graph shows that the Average Volume of the Volumes recorded at short term times at Plate 4 increase vs. time.)



(Figure 19: This is the resulting curve fit of the Average Volume vs. time for the short term Plate 4. An exponential equation was fit to the data. A decrease in average volume at early times was seen.)



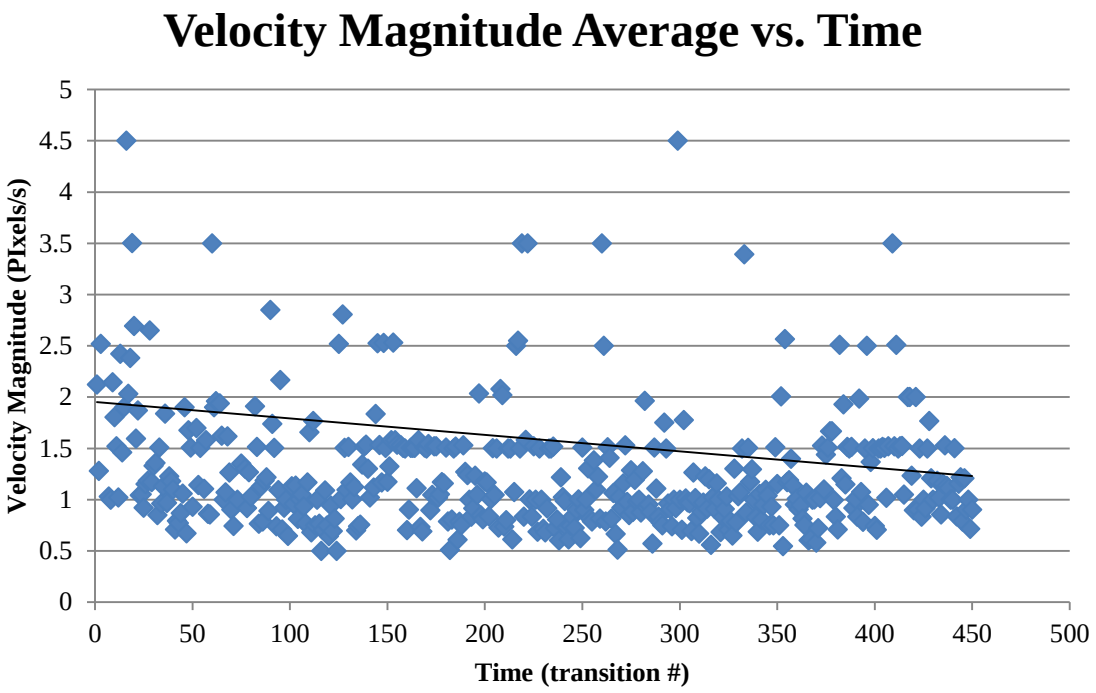
(Figure 20: The Total Volume, the sum of all volumes recorded at each time, for Plate 4 increases near small times t. A decrease in total volume was seen at early times during this period, and is considered to represent the incidence of the combination and compression of multiple cells into a nodule that is equal or less than the sum of the individual cells' original volumes.)



(Figure 21: This is the resulting curve fit of the Total Volume. An exponential curve fit was guessed in order to resemble the early times t of the long term V_{total} data.)

Visualization of Aggregation Events

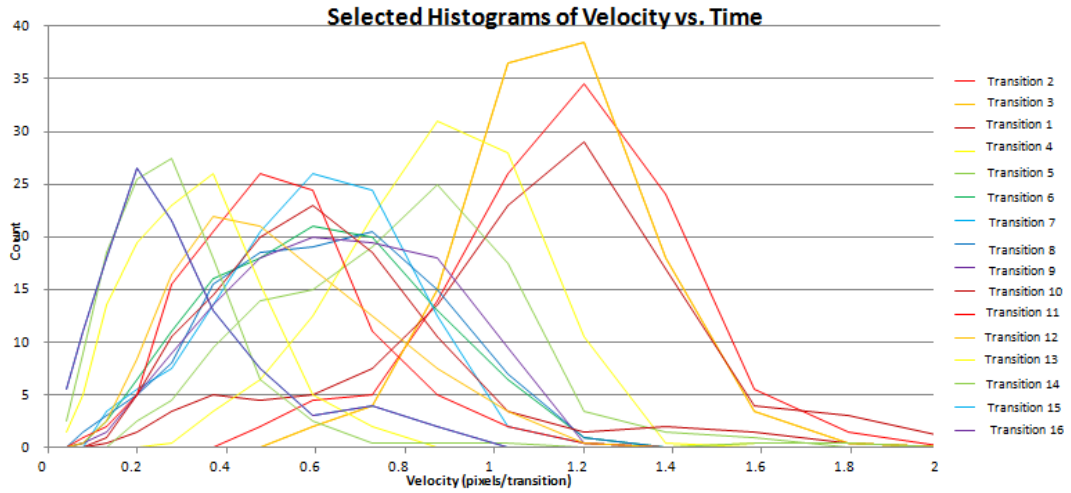
Figure 22 shows that the average velocity vs. time for short term plate 4 is noisy.



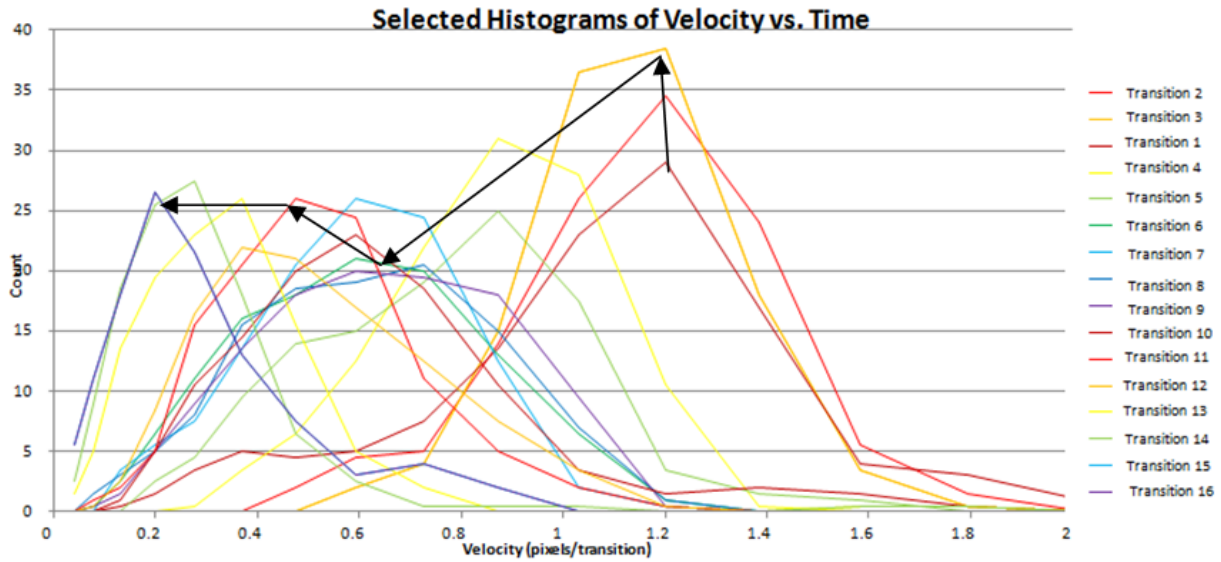
(Figure 22: Average Lateral Speed vs. Time for all cells and nodules on Plate 4 shows a general decrease in speed vs. time.)

Figure 22 also shows the curve-fit average velocity of acini decreasing over time. Figure 23 shows select histogram distributions of average lateral speed vs. time for short term plate 4. Figure 24 makes this behavior a bit clearer with the addition of arrows following the average peak of each histogram as it changes over time. This further indicates that as the second mode range increases and the average of the second mode

increases in size, the velocity of those acini decrease.



(Figure 23: Looking at select histograms of lateral speed vs. time for Plate 4 shows a decrease in average lateral speed over time. The analysis attempts to correct for possible "Brownian motion" behavior and look for a general trend.)



(Figure 24: Looking at select histograms of lateral speed vs. time for Plate 4 shows a decrease in average lateral speed over time. The black arrows have been added to follow the histogram peaks over time.)

CHAPTER 6

DISCUSSION

Relationships between measured values

Velocity decreases linearly with time, while the Average Volume increases exponentially. This could follow the thought that the bigger cells are slower. However, it seems as if the velocity vs. time graphical representations might be noisy to curve fit properly and it may be that velocity is consistently variable throughout time while Average Volume increases exponentially independently of the cell or nodule speeds.

Velocity decreases with time while the Total Volume increases as a Gompertzian function. This could follow the thought that the bigger cells are slower. However, the velocity vs. time graphical representations might be too noisy to curve fit properly and it may be that velocity is consistently variable throughout time while Total Volume increases and then levels off independently of the cell or nodule speeds.

Average Velocity decreases at a slow rate while the Standard Deviation of the 2nd Population Mode increases quadratically. This could follow the thought that the bigger cells are slower, but the average velocity stays relatively variable because the range of cell sizes increases as time increases.

Total Volume increases according to a Gompertzian according to the curve fitting done in Excel, while the Average Volume increases exponentially. Simplifying the comparison, where

$$\sum V = V_{total}$$

$\frac{\sum V}{N} = V_{avg}$, where N = number of Volumes

and thus:

V_{total}=$V_{avg} x N$,

it seems that both values would be suited to a Gompertzian fit at large times t, if not for aggregation behavior that affects the Average Volume and not the Total Volume values.

Total Volume vs. Time follows a Gompertzian curve, while the Standard Deviation of the 2nd Mode follows a quadratic equation. Thus, as the range of nodule sizes increases, it might be said that the total volume increases to a point before nodules begin shedding and recombining with nearby cells instead of continuing to divide. thus creates the leveling off effect seen at higher times on the Gompertzian curve.

Average Volume vs. Time increases exponentially, while the Standard Deviation of the 2nd Mode follows a quadratic equation. The maxima of the 2nd mode also increases at a quadratic pace. Thus, as the range of nodule sizes increases, it might be said that the average volume increases quickly and consistently until the histogram of Volumes recorded at each photo becomes very flat to accommodate the larger spread as well as the exponentially increasing maxima.

Error and Reproducibility

The main sources of error in this exercise were rounding errors due to finite precision arithmetic, truncation errors due to discretization and truncation of series, experimental errors due to inaccuracy in data values, and inefficiently written MATLAB code that can be improved upon in the future to reduce calculation times and efficiency for larger data sets. Possible sources of inaccuracies in the calculations would be inaccurate automatic thresholding by MATLAB for all images considered. In addition, if a nodule is not perfectly spherical, the Volume calculation using the measured Area would be inaccurate.

The MATLAB programs used in this project could be edited to more fully automate the image analysis and data analysis process to allow for use on photographs with subjects more detailed than cancer cells. One possible idea is using a simple image analysis program like was used in this project plus a satellite picture of a field of cattle to update the number of cattle in the field every morning and alert the owner of that field if a cow had given birth to a calf (which would have increased the count of creatures in the field).

Looking back closer to the goals of this project, the MATLAB programs could be reworked to handle larger batches of photographs of cancer cells in the hope that the patterns shown here would hold true for ovarian cancer cells grown into spherical tumors in a 3D environment.

Potential Implications and Conclusions

The Standard Deviation of the Second Mode of the Volume Histogram over a long time period did increase dramatically over time. A quadratic equation was fit to this increase. If this indicates a larger range of nodule sizes over time, several physical processes could be happening. One guess is that, as ovarian cancer cells do *in vivo*, that ovarian cancer cells *in vitro* in a 3D environment shed or seed off individual cancer cells when the nodules get too big. The ovarian cancer nodule development cycle may follow the pattern demonstrated by this data set. Initial single cells divide and combine with other single dividing cells before getting too big and beginning to shed cells. In turn those cells are beginning to divide and recombine with other cells, creating a diverse population of nodule sizes.

The Average Volume increased exponentially over time. This hints at cell division that results in a larger nodule instead of cell division that results in two individual daughter cells that separate completely. This could also be indicative of a steady increase in the pattern of either cells combining with nodules or other cells or a steady increase in the pattern of cells dividing within nodules over and over to increase the size of the larger nodules, maybe faster than other behaviors like seeding or shedding cells.

For cells in a 2D petri dish, the Gompertzian curve makes sense because the cells grow flat along the bottom exponentially until they fill the 2D flat surface and hit the walls, thus suddenly leveling out the Total Volume vs. Time curve. For cells in a 3D petri dish, made up of layers of biologically friendly gel or stroma, they are free to grow up as well through the gel. They have been found naturally to form spheroids. If the spheroid

growth leveled off at a long term time, it seems to indicate that ovarian cancer nodules have limitations on their size dependent on something about their growing environment besides simply running up against walls. If the Gompertzian curve accurately describes the ovarian cancer cells growing in the 3D environment, it might be an indicator that cells grown in an 3D *in vitro* environment have even more in common with those in an *in vivo* environment, where nodule size might be limited by the nodules metastasizing and shedding cells.

The velocity vs. time graph was too messy for us to truly have confidence in the possible decrease in velocity vs. time. The noise may have been due to the way the PIV program measured the pathways of cells within the short term Plate 4. In the future, I would aim to test out PIVlab and other PIV programs extensively before using them to process image data. However, the general trend of the velocity was decreasing over time. Perhaps the velocity data could be filtered somehow before comparing the Average Velocities vs. Time to one another.

In conclusion, this project met its goals of exploring the behavior of the second mode of the Volume histogram over time, examining the changes in the sum of all volumes measureable at each time over time, and examining the changes in the average of all volumes measureable at each time over time. In the future, I would look forward to working with more photographic data, and with MATLAB programs edited to handle more data.

APPENDIX A

MATLAB PROGRAM: “VolDataGather.m”

```
function VolDataGather(calfactor,minsize,time_scale,bins)
 %This .m file is meant to be placed in a folder of .tif images taken of
%different fields at the same time from start of different cell
%preparations. When it is run in MATLAB by entering
%VolDataGather(calfactor,minsize,time_scale,bins) into the MATLAB command
line
%[example with example input values: VolDataGather(1.41,100,1,22)], the result
%is a .mat file that includes data tables of 'Vlist','histogram','Vavg','Vtot'.
%
%'Vlist' is a 2 columned table where the first column is the measured size of
individual nodules and the second
%column records the time at which the picture of the nodule was taken, as
%determined by the input value of 'time'.
%
%'histogram' is a 2x22 matrix where the first column is a list of the minimum
edge of each bin and the
%second column lists the number of organisms whose volume (in units of
%microm^2) is greater than or equal to the minimum edge of each bin, but less
%than the minimum edge of the next biggest bin minimum edge. The number of
```

```
%bins is determined manually by determining the largest nodule volume at
%all times of data and is entered as a dependent variable.
%
% 'Vavg' is a single value for each time. It is the average volume of all
% volumes of organisms measured for each measurement time. It is calculated
% after every image at that specific time has been analyzed by using the
% built-in MATLAB function avg. Vavg=avg(Vlist) calculated the average
% volume of all organisms detected by the MATLAB image analysis tools at
% that specific time.
%
% 'Vtot' is a single value for each day. It is the total volume of all
% volumes of organisms measured for each day. It is calculated after every
% image at that specific time has been analyzed by using the built-in
% MATLAB function sum. Vtot=sum(Vlist) calculated the average volume of all
% organisms detected by the MATLAB image analysis tools at that specific
% time.
%
%
% Dependent Variables:
% calfactor = calibration sizing factor difference
% between picture and microscope
%
```

```
% minsize = minimum allowable cell size, in Gwen Deger's case, 100 microm^2
was chosen.
%
% time_scale = the time conversion factor from days to the desired time scale. For
example, ‘time_scale’ is
% equal to 1 if the time is recorded in the photograph file names in units
% of days (following the convention described above) and the desired units
% of time in the results .mat file is days.

% bins = initially is a guess of
% how many bins will be appropriate in the histogram, and then once the
% data is collected for all groups of data at each time t, a script is used
% to calculate the exact bin number necessary, and then the program is
% re-run on all the data using the exact bin manually calculated number.

%How to calculate the value = bins:
% 1) Run VolDataGather(calfactor,minsize,time_scale,bins) on all groups of
% of .tif files at all times t, with bins = a guess.
% 2) Load an individual sizeinfo<day#>.mat file and run
% MaxSize(Day#,1)=max(Vlist(:,1)); on each before loading the next .mat
% file.
% 3) Repeat step 2 on all sizeinfo<day#>.mat files.
% 4) Use the script [a,b]=find(max(MaxSize)); DayofMax=a;
```

```
% binMax=max(MaxSize); on the resulting MaxSize list.
% 5) Solve the following equation for i: binMax =
%minVol  + baseStep * 2^(i-1)
% 6) Round the bin number up to the nearest whole number. This number =
% bins.
% 7) Run VolDataGather(calfactor,minsize,time_scale,bins) again on all groups
of
% of .tif files at all times t, with bins = the accurate calculated number.

%Example bins calculation results:
%The maximum volume data value was at time = 18 days in the data set this
%program was originally written for and was equal to 160098749.340446
%(microm)^3. In this case: 160098749.340446 (microm)^3=
752.2528(microm)^3  + 100*2^(i-1)
%i=21.6105. As bin number must be a whole number, 21.6105 was rounded up to
22 bins.

% The values Gwen Deger used for her data were: VolDataGather(1.41,100,1,22)
dirname = uigetdir; % User specifies input folder containing .tif files by clicking.
Vlist=[];
% Finds image files
frames=dir('*.tif*'); %list of .tif files in input directory
c=char(frames.name); %makes a list c of the names of the tif files
```

```
S=size(char(frames.name)); %S(1)=# of tif files
timeS=c(1,end-5:end-4);
timeS=str2num(timeS).*time_scale;
   for frame = 1:S(1)
         progressIndication = sprintf('Analyzed frame %4d of %d.', frame, S(1)); %
Defines indicator of frame analysis progress
         filename2=c(frame,:);
         [acini_area]=get_area_acini_info(filename2,calfactor,minsize); %output =
acini_area column vector
            acini_Vol=(pi*(4/3)).*((acini_area./pi).^(.5)).^3;
            Vlist=[Vlist; acini_Vol, (timeS.*ones(length(acini_Vol),1))];
            disp(progressIndication); %displays as each frame is analyzed
   end
Vtot=sum(Vlist(:,1));
Vavg=mean(Vlist(:,1));

[histogram]=createVolHistogram(Vlist(:,1),100,bins)
outfile=strcat(dirname, '\', 'sizeinfo',c(1,end-5:end-4),'.mat');
save(outfile,'Vlist','histogram','Vavg','Vtot'); % write the output file to the input
directory
end

function [acini_area]=get_area_acini_info(filename,calfactor,minsize)
```

```
% code taken from (Celli, 2010) file
    I=imread(filename); %reads each frame individually
    level=graythresh(I); %turns to b&w + autothresholds
    Ibw=im2bw(I,level); %make the original image binary based on threshold from
graythresh
    Ibw=imfill(Ibw, 'holes'); %fill in holes so that objects are solid (not donuts)
    Ibw=imclearborder(Ibw, 8); %clear the edges - remove acini that are partially
in the field of view
    [labeled] = bwlabel(Ibw, 8);
    acinidata = regionprops(labeled, 'Area');
    calfactorsqrd=calfactor^2;
    acini_area = cat(1, acinidata.Area)*calfactorsqrd;
    acini_area=acini_area(acini_area>minsize); %removes outliers
end

function [histogram, histogramVector] =
createVolHistogram(values,minsize,bins)
%code taken from (Celli, 2010) file
minVol = (pi*(4/3)).*((minsize./pi).^(.5)).^3;
values = values;
bins = bins;

histogram = zeros(bins, 2);
```

```
baseStep = 100;
histogramVector = [];


binMax = minVol;
for i=1:size(histogram, 1)
    binMax = binMax + baseStep * 2^(i-1);
    histogram(i,2) = binMax;
    histogram(i,1) = length(values(values<=binMax));
    for j=1:length(values(values<=binMax));
        histogramVector = cat(1, histogramVector, i);
    end
    areas(areas<=binMax) = [];
end
end
```

APPENDIX B

MATLAB PROGRAM: “VolDataGatherShortTerm.m”

```
function VolDataGatherShortTerm(calfactor,minsize,time_scale)
 %This .m file is meant to be placed in a folder of .tif images taken of the
%same field at different sequential evenly spaced times from the start of a
%cell preparation. When it is run in MATLAB by entering
%VolDataGather(calfactor,minsize,time_scale) into the MATLAB command
%line [example with example input values: VolDataGather(1.41,100,1,22)],
%the result is a .mat file that includes data tables of
%'Vavg' and 'Vtot'.

%'Vavg' is a 2x452 matrix where the first column
%is the times of each .tif image and the second column is the average
%volume of all recorded nodules in the .tif image at each specific time.

%'Vtot' is a 2x452 matrix where the first column
%is the times of each .tif image and the second column is the total sum of the
%volumes of all recorded nodules in the .tif image at each specific time.
```

```
% The values Gwen Deger used VolDataGatherShortTerm.m for were:
VolDataGatherShortTerm(1.41,100,1)
dirname = uigetdir; % User specifies input folder containing .tif files by clicking.
Vlist=[];
Vavg=[];
Vtot=[];
% Finds image files
frames=dir('*.tif*'); %list of .tif files in input directory
c=char(frames.name); %makes a list c of the names of the tif files
S=size(char(frames.name)); %S(1)=# of tif files
   for frame = 1:S(1)
        timeSs=c(1,1:5);
        timeS(frame)=str2num(timeSs).*time_scale;
        progressIndication = sprintf('Analyzed frame %4d of %d.', frame, S(1)); %
Defines indicator of frame analysis progress
        filename2=c(frame,:);
        [acini_area]=get_area_acini_info(filename2,calfactor,minsize); %output =
acini_area column vector
            acini_Vol=(pi*(4/3)).*((acini_area./pi).^(.5)).^3;
            Vlist=[acini_Vol, (timeS.*ones(length(acini_Vol),1))];
            disp(progressIndication); %displays as each frame is analyzed
            Vtot(frame,:)=[timeS(frame),sum(Vlist)];
            Vavg(frame,:)=[timeS(frame),mean(Vlist)];
```

```
    end



outfile=strcat(dirname, '\', 'sizeinfoshortterm','.mat');
save(outfile,'Vavg','Vtot'); % write the output file to the input directory
end


function [acini_area]=get_area_acini_info(filename,calfactor,minsize)
% code taken from (Celli, 2010) file
    I=imread(filename); %reads each frame individually
    level=graythresh(I); %turns to b&w + autothresholds
    Ibw=im2bw(I,level); %make the original image binary based on threshold from
graythresh
    Ibw=imfill(Ibw, 'holes'); %fill in holes so that objects are solid (not donuts)
    Ibw=imclearborder(Ibw, 8); %clear the edges - remove acini that are partially
in the field of view
    [labeled] = bwlabel(Ibw, 8);
    acinidata = regionprops(labeled, 'Area');
    calfactorsqrd=calfactor^2;
    acini_area = cat(1, acinidata.Area)*calfactorsqrd;
    acini_area=acini_area(acini_area>minsize); %removes outliers
end
```

APPENDIX C

MATLAB PROGRAM: “PIVOrganizer.m”

```
function PIVOrganizer(bins)
 % Must be used in the folder of resulting files from the PIVlab analysis.
dirname = uigetdir;
m=[];
mavg=[];
mmax=[];
mmin=[];
uj=[];
vj=[];
frames=dir('*PIVlabavi_*'); %list of PIV result .mat files in input directory
c=char(frames.name); %makes a list c of the names of the PIV result .mat files
S=size(char(frames.name)); %S(1)=# of videos
timeS=c(1,end-8:end);
for i=1:450
   filename=c(i,:)
   load(filename);
   for j=1:10
```

```
        u(isnan(u))=0;
        v(isnan(v))=0;
        uj=[uj;u(:,j)];
        vj=[vj;v(:,j)];
    end
    for k=1:120
        m(k,i)=(uj(k,:).^2+vj(k,:).^2).^(.5);
    end
    uj=[];
    vj=[];
    mavg(i,:)=mean(m(:,i));
    mmax(i,:)=max(m(:,i));
    mmin(i,:)=min(m(:,i));
    [histogram, histogramVector] = createVelMagHistogram(m(:,i),bins);
    mhistogram(:,1)=histogram(:,1);
    mhistogram(:,i+1)=histogram(:,2);
end

 mavg; %should be a 450x1 matrix
 mhistogram; %should be a bin#x450 matrix
 mmax; %should be a 450x1 matrix
 mmin; %should be a 450x1 matrix
 outfile=strcat(dirname, '\', 'velinfo.mat');
```

```
 save(outfile,'mavg','mhistogram','mmax','mmin'); % write the output file to the
input directory

end

function [histogram, histogramVector] =
createVelMagHistogram(magnitudes,bins)

%bin/baseStep calculation:
%  5 = max value considered
%  0 = min value considered
%  baseStep =  ?
%  bins = 20
%  Max = bins*baseStep.*(bins-1);
%  5 = baseStep.*(20-1)
 %code taken from (Celli, 2010) file
minSize = 0;
magnitudes=magnitudes;
bins = bins;
histogram = zeros(bins, 2);
baseStep =  0.0132;
histogramVector = [];
 binMax = minSize;
```

```
for i=1:size(histogram, 1)
    binMax = binMax + baseStep.*(i-1);
    histogram(i,1) = binMax;
    histogram(i,2) = length(magnitudes(magnitudes<=binMax));
    for j=1:length(magnitudes(magnitudes<=binMax));
        histogramVector = cat(1, histogramVector, i);
    end
    magnitudes(magnitudes<=binMax) = [];
end
end
```

REFERENCE LIST

Bajzer, Ž., Vuk-Pavlović, S., & Huzak, M. (1997). Mathematical Modeling of Tumor Growth Kinetics. A Survey of Models for Tumor-Immune System Dynamics, 89-133. doi:10.1007/978-0-8176-8119-7_3

Castro, M. A., Klamt, F., Grieneisen, V. A., Grivicich, I., & Moreira, J. C. (2003). Gompertzian growth pattern correlated with phenotypic organization of colon carcinoma, malignant glioma and non-small cell lung carcinoma cell lines. Cell Proliferation, 36(2), 65-73. doi:10.1046/j.1365-2184.2003.00259.x

Celli, J. P., Rizvi, I., Evans, C. L., Abu-Yousif, A. O., & Hasan, T. (2010). Quantitative imaging reveals heterogeneous growth dynamics and treatment-dependent residual tumor distributions in a three-dimensional ovarian cancer model. Journal of Biomedical Optics, 15(5), 051603-051610. doi:10.1117/1.3483903

Chicoine, M. R., & Silbergeld, D. L. (1995). Assessment of brain tumor cell motility *in vivo* and *in vitro*. Journal of Neurosurgery, 82(4), 615-622. doi:10.3171/jns.1995.82.4.0615

DeLong, L., & Burkhart, N. W. (2008). General and oral pathology for the dental hygienist (88). Baltimore, MD: Lippincott Williams & Wilkins.

Edinger, M., Sweeney, T. J., Tucker, A. A., Olomu, A. B., Negrin, R. S., & Contag, C. H. (1999). Noninvasive Assessment of Tumor Cell Proliferation in Animal Models. *Neoplasia, 1*(4), 303-310. doi:10.1038/sj.neo.7900048

Farghaly, S. A. (2013). Anti- Metastatic Gene Therapy in Patients with Advanced Epithelial Ovarian Cancer (EOC). *Journal of Cell Science & Therapy J Cell Sci Ther, 04*(02). doi:10.4172/2157-7013.s15-001

Greenspan H. P., (1972), Models for the Growth of a Solid Tumor by Diffusion, Studies in Applied Mathematics, 51, doi: 10.1002/sapm1972514317.

Johnson, K. R., Leight, J. L., & Weaver, V. M. (2007). Demystifying the Effects of a Three-Dimensional Microenvironment in Tissue Morphogenesis. *Methods in Cell Biology Cell Mechanics,* 547-583. doi:10.1016/s0091-679x(07)83023-8

Johnston, M. D., Edwards, C. M., Bodmer, W. F., Maini, P. K., & Chapman, S. J. (2007). Mathematical modeling of cell population dynamics in the colonic crypt and in colorectal cancer. *Proceedings of the National Academy of Sciences, 104*(10), 4008-4013. doi:10.1073/pnas.0611179104

Kleinman, H. K., McGarvey, M. L., Hassell, J. R., Star, V. L., Cannon, F. B., Laurie, G. W., & Martin, G. R. (1986). Basement membrane complexes with biological activity. *Biochemistry, 25*(2), 312-318. doi:10.1021/bi00350a005

Kozusko, F., & Bajzer, Ž. (2003). Combining Gompertzian growth and cell population dynamics. *Mathematical Biosciences, 185*(2), 153-167. doi:10.1016/s0025-5564(03)00094-4

Laird, A. K. (1964). Dynamics of Tumor Growth. *Br J Cancer British Journal of Cancer, 18*(3), 490-502. doi:10.1038/bjc.1964.55

Laird, A. K., Tyler, S. A., Barton, A.D. (1965). Dynamics of normal growth. Br J Cancer, 29, 233–248.

Laird, A. K. (1965). Dynamics of Tumour Growth: Comparison of Growth Rates and Extrapolation of Growth Curve to One Cell. *Br J Cancer British Journal of Cancer, 19*(2), 278-291. doi:10.1038/bjc.1965.32

Lü, W., Zhang, L., Wu, C., Liu, Z., Lei, G., Liu, J., Gao, W., Hu, Y.R. (2014). Development of an Acellular Tumor Extracellular Matrix as a Three-Dimensional Scaffold for Tumor Engineering. *PLoS ONE, 9*(7). doi:10.1371/journal.pone.0103672

Marušić, M., B., Vuk-Pavlovic, S., & Freyer, J. (1994). Tumor growth and as multicellular spheroids compared by mathematical models. *Bulletin of Mathematical Biology, 56*(4), 617-631. doi:10.1016/s0092-8240(05)80306-4

Nelson, C. M., & Bissell, M. J. (2006). Of Extracellular Matrix, Scaffolds, and Signaling: Tissue Architecture Regulates Development, Homeostasis, and Cancer. *Annual Review of Cell and Developmental Biology Annu. Rev. Cell Dev. Biol., 22*(1), 287-309. doi:10.1146/annurev.cellbio.22.010305.104315

Norton, L. (1988). A Gompertzian model of human breast cancer growth. Cancer Research. 48 (24.1), 7067-7071.

Otsu, N. (1979). A threshold selection method from gray level histograms. IEEE Transactions on Systems. Man and Cybernetics. 9 (1), 62-66.

Davidson, M. W., Abramowitz, M. (2016). Molecular Expressions Microscopy Primer: Specialized Microscopy Techniques - Darkfield Illumination. Retrieved April 10, 2016, from http://micro.magnet.fsu.edu/primer/techniques/darkfield.html

Howlader N., Noone A.M., Krapcho M., Miller D., Bishop K., Altekruse S.F., Kosary C.L., Yu M., Ruhl J., Tatalovich Z., Mariotto A., Lewis D.R., Chen H.S., Feuer E.J., Cronin K.A. (2016). SEER Cancer Statistics Review, 1975-2013, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/statfacts/html/ovary.html, based on November 2015 SEER data submission.

Swan, G. W. (1990). Role of optimal control theory in cancer chemotherapy. *Mathematical Biosciences, 101*(2), 237-284. doi:10.1016/0025-5564(90)90021-p

Tchafa, A. M., Shah, A. D., Wang, S., Duong, M. T., & Shieh, A. C. (2012). Three-dimensional Cell Culture Model for Measuring the Effects of Interstitial Fluid Flow on Tumor Cell Invasion. *Journal of Visualized Experiments JoVE,* (65). doi:10.3791/4159

Tomlinson, I. P., & Bodmer, W. F. (1995). Failure of programmed cell death and differentiation as causes of tumors: Some simple mathematical models. *Proceedings of the National Academy of Sciences, 92*(24), 11130-11134. doi:10.1073/pnas.92.24.11130

Ward, J. (1997). Mathematical modelling of avascular-tumour growth. *Mathematical Medicine and Biology, 14*(1), 39-69. doi:10.1093/imammb/14.1.39

Zelen, M. (1966). Application of Exponential Models to Problems in Cancer Research. *Journal of the Royal Statistical Society. Series A (General), 129*(3), 368. doi:10.2307/2343503

BIOGRAPHICAL SKETCH OF AUTHOR

Gwendolyn A. Deger graduated from Grinnell College in 2011 with a BA in Physics.