

2008

Assessment of Tumour Growth in Murine Cancer Models with Three-Dimensional High-Frequency Ultrasound

Lauren Ashleigh Wirtzfeld
Western University

Follow this and additional works at: <https://ir.lib.uwo.ca/digitizedtheses>

Recommended Citation

Wirtzfeld, Lauren Ashleigh, "Assessment of Tumour Growth in Murine Cancer Models with Three-Dimensional High-Frequency Ultrasound" (2008). *Digitized Theses*. 4215.
<https://ir.lib.uwo.ca/digitizedtheses/4215>

This Thesis is brought to you for free and open access by the Digitized Special Collections at Scholarship@Western. It has been accepted for inclusion in Digitized Theses by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

**Assessment of Tumour Growth in Murine Cancer Models with
Three-Dimensional High-Frequency Ultrasound**

(Spine title: High-Frequency Ultrasound Assessment of Murine Tumour Growth)
(Thesis format: Integrated-Article)

by

Lauren Ashleigh Wirtzfeld

2
1
Biomedical Engineering Graduate Program

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

December 15, 2008

© Lauren Ashleigh Wirtzfeld, 2008

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Chief Advisor

Dr. Aaron Fenster

Dr. James C. Lacefield

Advisory Committee

Dr. Paula Foster

Dr. Ann Chambers

Examining Board

Dr. David Goertz

Dr. Eugene Wong

Dr. Peter Canham

Dr. Terry Peters

The thesis by

Lauren Ashleigh Wirtzfeld

entitled

**Assessment of Tumour Growth in Murine Cancer Models with
Three-Dimensional High-Frequency Ultrasound**

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Chair of Examination Board

Dated this 15th day of December, 2008.

Dr. Giles Santyr

Abstract

Preclinical cancer research could benefit from quantitative, non-invasive measurements of tumour growth provided by three-dimensional high-frequency ultrasound imaging. High-frequency ultrasound has been shown to be appropriate for tracking experimental liver metastases from a variety of cell lines without exogenous contrast agents. Tumour growth over time can be monitored on an individual tumour basis, allowing a growth curve to be constructed and the tumour to act as its own control in a treatment study.

In order to quantify tumour volume and growth, the measurement variability must be known. Inter- and intra-observer variability was determined for tumours in four size ranges with average volume from 0.43 mm³ to 60.42 mm³. Intra-observer variability was as low as 4% for mid-sized tumours averaging 2.39 mm³, while the inter-observer variability for the smallest and largest tumours measured had the highest variability at 25% and 15%, respectively. Breathing motion did not significantly effect the volume measurements, however, having the region of interest beyond the geometric focus resulted in significantly different measured volumes.

Measurement variability is one factor that influences how well growth data can be characterized mathematically through curve fitting. Simulations of tumour growth were performed to relate experimental imaging parameters, such as intervals between acquiring images, minimum and maximum volume recorded and length of time over which data is acquired, to the quality of curve fitting results. Simulations show that improving the ability of the ultrasound system to image small (<1 mm diameter) tumours would improve the ability to draw conclusions from growth parameters.

The spatially variant point-spread function influences lesion-size measurement variability and consequently growth curve fitting. The transducer employed is tightly focused, so spatial image resolution is high at the focus but rapidly degrades away from the focus. Synthetic aperture focusing was employed with a variety of weighting techniques to retrospectively focus the images through a range of depths. The

improvement in focusing was measured using point-like targets and the effect on measurement variability was evaluated using lesion phantom images. Synthetic aperture focusing did not produce a significant reduction in lesion-size measurement variability but did diminish the sensitivity of the measured size to lesion depth.

Keywords: high-frequency ultrasound, three-dimensional ultrasound, mouse models, preclinical cancer, measurement variability, tumour growth, synthetic aperture focusing

Co-Authorship

Chapter 2 was published in *Cancer Research* as: K. C. Graham, L. A. Wirtzfeld, L. T. MacKenzie, C. O. Postenka, A. C. Groom, I. C. MacDonald, A. Fenster, J. C. Lacefield, and A. F. Chambers, "Three-dimensional high-frequency ultrasound imaging for longitudinal evaluation of liver metastases in pre-clinical models," *Cancer Res.*, vol. 65 2005. All authors contributed to editing of the paper and was written by Kevin Graham and myself. This paper was a joint publication between Drs Chambers, MacDonald and Groom's laboratory and Drs Lacefield and Fenster's laboratory. Kevin Graham, a PhD student in Medical Biophysics at the time, was responsible for animal models, including animal handling, animal care, cell culture, coordinating the necessary surgeries for creating the mouse models and the histology. I was responsible for the imaging setup, scanning and image acquisition. I was responsible for the segmentation for the initial growth studies and the segmentation for the treatment study was shared between myself and Kevin Graham. Post-analysis, including statistical comparisons of groups were performed jointly. Kevin worked on this project under the supervision of Alan Groom, Ian MacDonald and Ann Chambers. I worked on this project under the supervision of Aaron Fenster and James Lacefield.

Chapter 3 was published in *Physics in Medicine and Biology* as: L. A. Wirtzfeld, K. C. Graham, A. C. Groom, I. C. Macdonald, A. F. Chambers, A. Fenster, and J. C. Lacefield, "Volume measurement variability in three-dimensional high-frequency ultrasound images of murine liver metastases," *Phys Med Biol*, vol. 51, no. 10, pp. 2367-81, 2006. All authors contributed to the editing of the paper. Kevin Graham was responsible for the animals and the animal model that were imaged for the variability studies. The multi-observer segmentation study was coordinated between Kevin Graham and myself.. I segmented the tumours for the rest of the studies and performed the statistical analyses. Kevin worked on this project under the supervision of Alan Groom, Ian MacDonald and Ann Chambers. I worked on this project under the supervision of Aaron Fenster and James Lacefield.

Chapter 4 is in preparation to be submitted with authors: L. A. Wirtzfeld, A. Fenster and J. C. Lacefield. The range of coefficients used for the simulated growth data came from growth curves obtained from work for the paper in chapter 2, which was worked carried out jointly between myself and Kevin Graham. I wrote the code to simulate the growth curves and perform the curve fitting and subsequently analysed the data. All work was carried out under the supervision of Aaron Fenster and James Lacefield.

Chapter 5 is in preparation to be submitted with authors: L. A. Wirtzfeld, A. F. Chambers, A. Fenster and J. C. Lacefield. For this study, I designed and constructed the phantom required and set up all the imaging experiments. I wrote the code to perform the synthetic aperture focusing and weighting techniques. I made all the measurements presented in the chapter and have done the analyses of the results. All work was carried out under the supervision of Aaron Fenster and James Lacefield.

This thesis is dedicated to my husband, Michael, and my daughter, Katarina.

You have transformed my life since I began this degree.

I love you both with everything that I am.

Acknowledgments

It has been a privilege to work with so many fantastic people during my time at Robarts and UWO. All those that I've met and worked with are what made these last five years more than just school but life.

First of all, I'd like to acknowledge my supervisors, Jim Lacefield and Aaron Fenster. Two supervisors makes everything just a little more interesting! I have benefited enormously from having a chance to work with both of them and I do not think I could have found better supervisors anywhere else. Their two styles complement each other nicely and this truly added to my experience as a graduate student. Aaron has an incredible ability to calm down any situation, it does not matter if the world is coming to an end, Aaron will have a way to make the situation manageable. Jim truly has a passion for teaching what he knows best, ultrasound. As a results we all have learned a great deal beyond our own projects which is incredibly valuable.

Thank you to my advisory committee, Ann Chambers and Paula Foster. You have both been so approachable it has made things that bit easier. I appreciate that you were available to answer questions when needed and I never felt I had to wait for an official meeting.

I also have to acknowledge the Lacefield lab, everyone still here and those who have moved on. Adam Waspe, Ethan Shen, Stephen Pinter, Mohammad Daoud, Zamir Khan and James Qui, you have shared the joys, frustrations, late nights and early mornings of research with me. Especially to Adam Waspe who walked through the doors to Robarts for the first time on the same day as me. We've been through a lot together these past five years and I wouldn't have made it through with my sanity (what's left of it) without you. To Ethan, congratulations to you on finishing, you beat me by a couple of days! Stephen and Mohammad, it has been great getting to know you, your expertise have been a welcome help. Zamir and James headed off to bigger and brighter things before us, but are not forgotten. Your code for the RF

acquisition allowed me to do the work in chapter 5, for that I am very grateful. Thank you to all the summer students who have brightened up the lab with their enthusiasm and especially the two high-school co-op students who worked with me, Megan Gregor and Imane Bekri; you put your enthusiasm into your work and accomplished more than anyone expected a high school student to be able to do. I wish you both luck in your future careers.

I need to give a huge thank you to Mike Bygrave, he has been the glue that has stuck the ultrasound micro-imaging group together. He is the one who will offer to join you for experiments on a long weekend to give you a hand, the one who will be there until 1am because he thought he could get a bit better image and the one who always had time to help everyone out and make their day a bit better, even if it meant he had little time for himself. For all the help in the lab and all the personal support you've offered me, thank you.

A thank you to our collaborators and co-authors from Ann Chambers, Ian MacDonald and Alan Groom's lab(s)! This project only exists because of the collaboration between our labs and has resulted in some very interesting projects and results. Thanks for your time and enthusiasm that you put into this work. In particular, many members of the lab have been directly involved, Lisa MacKenzie offered a phenomenal amount of support while she was here and is still sorely missed, Carl Postenka contributed a lot to the histological confirmation. A huge thank you to all the members of the Chambers lab who pitched in and helped out over the years, especially to Jason Townson and Mike Lizarde who were suckered in to participating in ultrasound experiments with your animal models, I am truly grateful and I thoroughly enjoyed working with you both. And to those who were co-opted into segmenting for one of our studies, most who are named above, but also Jason Carikoglu and Jackie Graham.

I need to especially thank and acknowledge Kevin Graham, he is the co-author for the first half of my thesis. This work was truly joint work between the two of us; we spent a ridiculous number of hours in the lab together imaging and worrying about

our isoflurane exposure. We spent so much time together I think Kevin was the first person to guess I was pregnant before anyone was told, and of course he quietly said nothing. This thesis would not exist if it was not for Kevin. I do not know how to properly express my gratitude and appreciation for all Kevin's skills, help and overall wonderful personality that made working on these projects together a joy.

Thank you to Kim Perry and Darlene Goodine for all the administrative help and guidance at Robarts. To Laura Groom with the CIHR-STP program who has been a wonderful person to talk to and is great at tracking down missing money!

An enormous thank you to Diana Timmermans. I am at a loss for words about what to say about Diana that would sum it up on a piece of paper. The BME grad program is what it is because of Diana. She makes the program run. She is there for all of us students no matter when or for what. Who else do you know that goes out of their way to make sure you can get hold of her by phone, email, messenger, you name it, just in case you needed her. Diana has been a fantastic person to work with and an even better person to get to know as a friend.

I would like to add a thank you to Ann Least; we've spent the last years tucked into the same corner in Robarts and she has always been there to make things run just a bit smoother or to add a bit of cheer to the day.

I would like to acknowledge my funding sources over the years. I was funded by the Ontario Graduate Scholarships for 2004-2005 and by the National Science and Engineering Research Council from 2005 to 2008. I would also like to acknowledge funding and support from the UWO CIHR Strategic Training Program in Cancer Research and Technology Transforms and the Western Graduate Research Scholarship.

To my family, thank you for being their through all the frustrations, long days living in the lab and failed experiments. To my parents, Don and Julie Hastie, you've always been there to support my every decision and to help me take the next step. I truly appreciate that you have always been behind me and supported my decisions in life and especially in my education. The love of learning starts at home and you

always encouraged to enjoy learning, from the piles of books we had when we were little, to science fair projects, to building robots in the basement all night. Thank you for your support through all the years and especially your support of Katarina, Michael and myself this past year.

To my sister and brother-in-law, Kendra Hastie and Chris Sinal, you have played the parts of the doting aunt and uncle incredibly well these last two months and helped me more than you can imagine to get this thesis finished. Both of you truly understand what it is like struggling in school and the importance of continuing on the educational path. If you're still up for the challenge, let's see which couple has more degrees at the end! Kendra, you've have been a fantastic sister to me, I couldn't have picked a better one myself, thank you for your support and let me know when you need me to help you get your next degree finished!

Lastly, but certainly not least, to Michael and Katarina. Michael, you've changed my life over these past five years in London. I moved to London to get a graduate degree and have found so much more, first a husband and now we have a daughter too. How many people are lucky enough to find the one right person, but also find a parter to discuss Fourier Transforms over pancakes and random processes as you drift off to sleep with? Your patience and understanding these last months especially is appreciated.

And to Katarina, who's help has been a little less tangible - she still likes to type randomly in the middle of my paragraphs, is fantastically fast at resetting the computer and will willing sort any pile of nicely arranged journal articles - you've been my sun and my moon since before you were born. Finishing thesis without you would, I'm sure, have been easier and a smoother ride, but I'm also certain it would not have been as meaningful nor as enjoyable. I'm finishing my formal education and you haven't quite started - just remember learning just keeps on going, have fun, learn something new every day and when you're trying to finish your degrees and need a snack or someone to talk to at 4am, I'll be there!

Contents

Certificate of Examination	ii
Abstract	iii
Co-Authorship	v
Acknowledgements	viii
List of Tables	xvi
List of Figures	xvii
List of Symbols	xix
List of Acronyms & Abbreviations	xx
1 Introduction	1
1.1 Overview	1
1.2 Preclinical Cancer Studies and Animal Models	2
1.2.1 Failure of cancer drugs at the clinical trial level	2
1.2.2 Limitations of current animal models	3
1.2.3 Current Endpoints for Analysis	5
1.2.4 Opportunity to Improve Preclinical Studies	7
1.3 <i>In vivo</i> Micro-Imaging	8
1.4 Ultrasound	9
1.4.1 Comparison of High-Frequency Ultrasound to Other Micro-Imaging Modalities	9
1.4.2 High-Frequency Compared to Clinical Ultrasound	10
1.4.3 Ultrasound Arrays and Focusing	11
1.4.4 Early Applications of High-Frequency Ultrasound to Cancer	13
1.5 Hypothesis and Objectives	13
1.6 Thesis outline	14

1.6.1	Chapter 2	14
1.6.2	Chapter 3	15
1.6.3	Chapter 4	15
1.6.4	Chapter 5	16
References		17
2	Three-Dimensional High-Frequency Ultrasound Imaging For Longitudinal Evaluation Of Liver Metastases In Preclinical Models	21
2.1	Introduction	22
2.2	Materials and Methods	24
2.2.1	Cell Culture and Experimental Metastasis Models	24
2.2.2	Ultrasound Imaging	25
2.2.3	Volume Measurements	26
2.2.4	Longitudinal Growth Measurements	27
2.2.5	Treatment Protocols	27
2.2.6	Histology	28
2.3	Results	28
2.3.1	Identification of murine liver metastases using high-frequency ultrasound	28
2.3.2	High-frequency ultrasound can identify areas of liquefactive necrosis within metastases	30
2.3.3	Tracking the growth of individual liver metastases by noninvasive ultrasound imaging	30
2.3.4	2D measurement provides frequent overestimation or underestimation of tumour volume as compared to 3D measurement	31
2.3.5	3D ultrasound can be used to monitor therapeutic response of individual metastases	34
2.4	Discussion	34
References		40
3	Volume Measurement Variability In Three-Dimensional High-Frequency Ultrasound Images Of Murine Liver Metastases	44
3.1	Introduction	44
3.2	Methods	47
3.2.1	Animal Model	47
3.2.2	Image Acquisition	47
3.2.3	Tumour Volume Measurement	48
3.2.4	Analysis of Volume Measurement Variability	50
3.3	Results	54
3.3.1	Inter- and Intra-Observer Variability	54
3.3.2	Variability due to Segmentation Inter-Slice Distance	57
3.3.3	Variability due to Experimental Factors	57

3.4	Discussion	59
3.4.1	Inter- and Intra-Observer Variability	59
3.4.2	Variability due to Segmentation Inter-Slice Distance	61
3.4.3	Variability due to Experimental Factors	62
3.4.4	Design of Longitudinal Imaging Studies	62
3.5	Conclusions	65
	References	67
4	Monte Carlo Growth Curve Simulations for Planning Longitudinal Imaging Experiments with Mouse Cancer Models	71
4.1	Introduction	71
4.2	Materials and Methods	76
4.2.1	Simulations	76
4.2.2	Statistical Methods	78
4.2.3	Gompertz Versus Exponential Functions	82
4.2.4	Comparison of Two Different Growth Curves	82
4.2.5	Comparison of Early and Late Time Growth Curves	83
4.3	Results	84
4.3.1	Gompertz Versus Exponential Functions	84
4.3.2	Comparison of Two Different Growth Curves	87
4.3.3	Comparison of Early and Late Time Growth Curves	89
4.4	Discussion	92
4.5	Conclusion	95
	References	96
5	Synthetic Aperture Focusing of a Single Element 40 MHz System	99
5.1	Introduction:	99
5.1.1	Outline of chapter	103
5.2	Materials and Methods:	104
5.2.1	Synthetic Aperture Focusing Technique	104
5.2.2	Experimental data - RF acquisition	105
5.2.3	Phantom Experiments	106
5.2.4	Liver metastasis imaging	108
5.3	Results:	108
5.3.1	Phantom validation	108
5.3.2	Liver metastasis imaging	115
5.4	Discussion:	115
5.4.1	Air channels	115
5.4.2	Gelatin lesion phantoms	117
5.4.3	Liver metastasis imaging	117
5.5	Conclusion:	118
	References	119

6 Summary and Future work	122
6.1 Summary	122
6.1.1 Chapter 2: Three-Dimensional High-Frequency Ultrasound Imaging For Longitudinal Evaluation Of Liver Metastases In Preclinical Models	122
6.1.2 Chapter 3: Volume Measurement Variability In Three-Dimensional High-Frequency Ultrasound Images Of Murine Liver Metastases	123
6.1.3 Chapter 4: Monte Carlo Growth Curve Simulations for Planning Longitudinal Imaging Experiments with Mouse Cancer Models	124
6.1.4 Chapter 5: Synthetic Aperture Focusing of a Single Element 40 MHz System	125
6.1.5 Current Progress	125
6.2 Future Work	126
6.3 Conclusions	127
References	129
A Copyright agreements	131
B Ethics Approval for Animal Subjects	132
Vita	134

List of Tables

3.1	Tumour images included in variability analysis.	51
3.2	Inter- and intra-observer variability measures.	55
3.3	Tumour volume measurement comparison for varied breathing patterns and location within the field of view.	59
4.1	Parameters used in data simulation and curve fitting.	80
5.1	Minimum measured cross-sectional area of the air channels with different SAFT weighting methods.	109
5.2	Measurement variability for lesion phantom areas.	115

List of Figures

2.1	Examples of B16F1, HT-29, MDA-MB-435/HAL and PAP2 liver metastases.	29
2.2	Tracking the growth of individual liver metastases by noninvasive ultrasound imaging.	32
2.3	Discrepancy between tumour volumes obtained from 3D segmentation or diameter measurement with the assumption of an ellipsoid shape.	33
2.4	A longitudinal ultrasound study shows that doxorubicin (DXR) decreases tumour growth rate and tumour volume in the B16F1 liver metastasis model.	35
3.1	Steps in tumour segmentation.	49
3.2	Standard deviation versus volume of repeated tumour measurements.	55
3.3	Inter- and intra-observer COV versus volume bin	56
3.4	Average normalized volume versus inter-slice distance for different tumour size ranges.	58
3.5	47.7 mm ³ tumour.	60
3.6	(a) Minimum detectable change in volume versus current volume. (b) Time required to observe a detectable change in volume.	64
4.1	Example of an exponential and Gompertz function with the same growth rate at time zero.	74
4.2	Graphs showing the simulation steps to general data and perform curve fitting.	79
4.3	Example of a simulated tumour growth function with a Gompertz and exponential function fitted to the data.	84
4.4	Graphs showing the marginal means of the Akaike weight comparing a Gompertz function to an exponential function.	86
4.5	Example of two different simulated tumour growth curves with a Gompertz function fit to each individual or constrained to be parallel.	88

4.6	Graphs showing the marginal means of the Akaike weight comparing individual Gompertz functions fit to the two datasets compared to fitting parallel functions.	90
4.7	Example of a simulated tumour growth curve with the first and second halves of the data fitted separately with Gompertz functions or with parallel Gompertz functions	91
4.8	Graphs showing the marginal means of the Akaike weight comparing individual Gompertz functions fit to both the first and second half of the graph compared to fitting parallel functions.	93
5.1	Virtual Source	102
5.2	Air channels on diagonal in ultrasound images with different weighting functions for the SAFT.	110
5.3	Average measured cross-sectional area of the air channels versus image depth.	111
5.4	Lesion phantom images from original and static weighted SAFT data.	113
5.5	Lesion phantom images from original and adaptive weighted SAFT data.	114
5.6	Example liver metastasis image with different SAFT weighting techniques applied.	116

List of Symbols

α	Parameter in generalized logistic function for point of inflection
Δv_{min}	Minimum detectable change in volume
λ	Ultrasound wavelength
θ	Carrying capacity of generalized logistic function
a	Growth rate of exponential function, initial growth rate in Gompertz and generalized logistic functions
b	Rate at which growth slows down in the Gompertz function
K	Number of coefficients in a function plus one
N	Number of data points
R	Autocovariance
t	Time
$V(t)$	Volume as a function of time
V_{∞}	The maximum volume the Gompertz function asymptotically approaches
V_0	Initial volume of a growth function

List of Acronyms & Abbreviations

3D	Three-dimensional
2D	Two-dimensional
α MEM	alpha minimum essential medium
AW	Akaike weight
AIC	Akaike information criterion
AIC _c	Corrected AIC
DMEM	Dulbecco's Modified Eagle Medium
CT	X-ray computed tomography
CF	Coherence factor
COV	Coefficient of variation (standard deviation divided by the mean)
DXR	Doxorubicin
GAFF	Generalized coherence factor
FIBS	fetal bovine serum
ISO	Inter-slice distance
LR	Lateral resolution
MRI	Magnetic resonance imaging
MV	Minimum variance
RF	Radio frequency
SAFT	Synthetic aperture focusing technique
s.c.	Sub-cutaneous
SCID	Severe combined immunodeficiency

SEM Standard error of measurement

SS Sum-of-squares, a measure of goodness of fit for curve fitting

Computer 1

Instructions

Introduction

The purpose of this manual is to provide a detailed description of the computer program. The program is designed to calculate the standard error of measurement (SEM) and the sum-of-squares (SS) for a set of data. The user is required to input the data and the program will output the results.

The program is written in a high-level programming language and is easy to use. It requires a computer with a minimum of 100K of memory. The user should refer to the manual for more details on the program's operation.

The program is designed to be used on a variety of computers. It is compatible with the following systems: IBM PC, IBM compatible, and other systems with a similar architecture.

The program is a self-contained package. It includes the source code, the compiled program, and the manual. The user should refer to the manual for more details on the program's operation.

The program is a valuable tool for researchers and students alike. It provides a quick and easy way to calculate the SEM and SS for a set of data. The user should refer to the manual for more details on the program's operation.

Chapter 1

Introduction

1.1 Overview

Clinical trials for therapeutic oncology drugs suffer from one of the highest rates of drug failure of any disease [1, 2]. Additionally, conventional endpoint analyses in preclinical cancer research provide limited information about tumour growth, due to limitations in measurement precision or an inability to measure tumours within some organs. These limitations suggest additional information may be required to assess the efficacy of novel therapeutics in the preclinical models. Three-dimensional (3D) high-frequency ultrasound imaging offers a non-invasive method to monitor tumour growth through time and provide quantitative information on tumour growth.

This thesis examines the use of three-dimensional (3D) high-frequency ultrasound imaging to measure tumour volumes in preclinical cancer models over time. To ensure that the data can be appropriately interpreted and changes in volume quantified, it is imperative to quantify measurement variability. Within this thesis, a liver metastasis model is imaged to validate that tumours could be imaged with ultrasound and show sufficient contrast to make measurements. The measurement variability of the tumour volumes was evaluated for multiple observers and varied experimental factors. The measurement variability is a function of tumour size and therefore will have an effect

on curve fitting of growth data. Therefore, this information was subsequently used in a study on theoretical growth curves and curve fitting to aid in the planning of imaging experiments. Lastly, since the measurement variability was shown to be a function of lesion depth within the image, a technique to improve the image focusing was implemented to determine the effects on the measurement variability and the spatial variance of the variability within the image.

In order to provide some context for the subsequent chapters, this chapter first discusses the shortcomings of clinical cancer drug trials and some of the limitations of the current animal models. Current techniques for measuring outcomes, including caliper and endpoint analysis, are discussed including how precise, longitudinal measurements could improve upon the current endpoints. Options for *in vivo* imaging are briefly presented followed by more detailed information on the use of high-frequency ultrasound for longitudinal imaging. Information on focusing of ultrasound beams and the use of transducer arrays is included as background on the limitations of the ultrasound system used. Previous work on high-frequency imaging of cancer models is presented followed by a more details outline of the thesis chapters.

1.2 Preclinical Cancer Studies and Animal Models

Preclinical cancer research employs a range of *in vivo* and *in vitro* techniques to study the fundamental biology of cancer and the response to current or new therapies. Techniques range from *in vitro* cell culture to *in vivo* animal models with implanted cells to genetically modified animals.

1.2.1 Failure of cancer drugs at the clinical trial level

All new therapeutic agents go through extensive testing in the laboratory before they ever reach a clinical trial. Although an enormous amount of time and resources are spent on researching, developing and testing new therapeutics, cancer drugs see

a high rate of failure at clinical trials (9 out of every 10 drugs fail) [1, 2]. This results in a much lower percentage of drugs that start trials making it through to the clinical compared to other disease models. In fact as few as half as many drugs compared to therapeutics for other diseases such as cardiovascular and arthritis [2]. Limitations in the preclinical models and evaluation of treatment response endpoints reduce their predictive potential in clinical trials. It is the predictive potential of these models to select effective drugs that is considered to be the most important concept when dealing with animal models [3]. The mechanism of action as well as the toxicity profiles often differ between the preclinical and clinical trials [2]. Such vast differences in results make it difficult to use preclinical results to determine an appropriate route of administration, initial dosing and type of cancer to target when these drugs reach the clinic.

1.2.2 Limitations of current animal models

To begin to understand the reasons for the discrepancies between preclinical trial outcomes and clinical trial results, it is necessary to look at the types of animal models employed in the preclinical studies and how they differ from the disease in humans. A large number of different preclinical models are employed in cancer research, as each has advantages and none have shown a superior predictive value over the others. Xenograft models, where cells are grown in culture then injected into the mouse, or small tumour masses are transplanted from one animal to another, are the most commonly used mouse model, due to the ease of use. For all xenografts, the resulting disruption in the cellular micro-environment due to creating the xenograft means that many molecular pathways are no longer intact [3], which has the potential to change the behaviour of the tumour cells and their interaction with the surrounding tissue [4].

Subcutaneous (*s.c.*) tumours are one of the most commonly used xenografts, where cancer cells are injected beneath the skin to form a tumour. Although tumours form and develop a blood supply that can be studied, they lack an organ

microenvironment and many of the constraints that arise for tumours in their natural state growing within an organ such as limited space.

Orthograft models, a xenograft where cancer cells are implanted into their native host organ (ie. melanoma cells into the skin, prostate cancer cells into the prostate), exhibit some of the tumour-host interactions more appropriately including invading the normal tissue and recruiting a blood supply from within the organ. However, it is typical to inject a large number of cancer cells in one location compared to a naturally arising tumour starting from mutations on the level of individual cells. Genetically modified mouse models offer a more realistic model of some aspects of tumour development as tumours initiate spontaneously due to genetic modifications making the mice susceptible to particular cancers. The benefit of these models is that the natural history of the tumours more closely relate to naturally occurring tumours. However, if the genetic modifications are not representative of those found in human cancers, it is difficult to interpret how the results may relate to the human form of the disease from their results [5]. In many cases, these animals exhibit a mutation in every cell in their body rather than having a mutated cell arise among normal cells and thus these models more closely resemble familial cancers, which are rare compared to the sporadically arising tumours found in humans [4, 5]. Monitoring and comparing animals in different treatment groups also poses challenges since each tumour will develop at a different time point and often develop multifocally, making them both more cumbersome as well as more expensive than xenograft models [5].

The process of cancer cells shedding from the primary tumour, travelling through the blood vessels or lymphatic system to distant organs where they can form tumours is called metastasis [6]. Metastases, not the primary tumours, are responsible for most cancer deaths, making them an important area of study. In order to study metastases, both spontaneous models and experimental models are employed. Spontaneous models use xenograft or genetically modified tumour models and allow the metastatic process to proceed naturally. In experimental models cancer cells are injected directly

into veins or lymphatics to target a particular organ. The tumour cells travel to that organ and a number of them will extravasate from the vessels into the tissue and have the potential to form metastases.

Outcomes of clinical trials may be poor due to the fact that preclinical models are not representing the diversity of the true disease of cancer. Often positive results achieved in a very small number of models, for only a small number of cell lines, are used as a positive indication to pursue the drug into clinical trials [7]. When the models showing positive results have no positive correlation with how human cancers will respond, this limits their usefulness. Rather than simply knowing that a drug has an effect on some cells, it would be more valuable to know what characteristic of the cells are important in achieving a response to the drug [1]. This would in turn allow the clinical trials to be conducted using the most appropriate type of cancer for the drug to show an effect.

Although there are limitations to the different preclinical models, they provide valuable tools for cancer research and show clinical predictive value for some cases. Voskoglou-Nomikos *et al.* [7] have shown, by meta-data analysis, that some orthotopic models show correlation with phase II clinical trial outcomes particularly when panels of xenografts were used preclinically.

1.2.3 Current Endpoints for Analysis

In order to assess experimental outcomes for growth and especially treatment studies, specified endpoints and metrics to compare groups are required. Currently, the most common analyses in cancer models are caliper measurements of subcutaneous or other superficial tumours or sacrifice at the experimental endpoint to allow gross pathological and histological comparisons [8]. Caliper measurements are subject to a high level of error and variability as there is no way to consistently measure a single dimension of a three-dimensional object. *In vivo* caliper measurements will always include the skin on top of the tumour and potentially additional layers of tissue or

fat. While this additional thickness is often assumed to be consistent across all measurements, there is no way to know if there are changes in the amount of normal tissue present in addition to the solid tumour mass. There is also the possibility that the widest part of the tumour is not at the surface but buried deeper and not being measured at all, thereby introducing more errors.

To compare tumour sizes over time and across groups, the volume is typically assumed to be an ellipsoid and estimated by measuring the longest dimension of the tumour, the perpendicular diameter to this initial measurement and setting the third depth dimension equal to the perpendicular diameter, and calculating the volume of an ellipsoid based on these three one-dimensional measurements [8]. Diameter measurements rarely provide a reliable estimate of the tumour volume as the true shape is often more complex than an ellipsoid. These assumptions and limitations severely limit the ability to compare tumour volumes over time and to construct meaningful growth curves. Average tumour measurements obtained from caliper measurements can be used to evaluate changes in growth between control and treated groups at each time point and to construct average growth curves, though the compounding effect of the measurement errors is difficult to assess. The high variability in these measurements makes comparing individual growth curves and even changes in volume of the same tumour over time unreliable.

For models which do not allow for external caliper measurements, analysis is often limited to *ex-vivo* endpoint analysis. As a result, to analyse multiple time points, multiple cohorts of animals are required to allow a cohort to be sacrificed at each time point of interest. This can be costly in terms of number of animals and work required to create that number of animal models and process the data. In addition, variations between the growth of individual tumours in individual animals increases the experimental variability and limits the ability to interpret the results and detect significant differences between treatment groups. A variety of endpoints can be used, including tumour size, tumour weight or total tumour burden. Animals that fail to

develop tumours can skew the results and limit the data interpretation, especially if a control animal does not develop a tumour.

In the case of genetically modified animals with spontaneously arising tumours that are not readily visible or palpated, it can be difficult to determine an appropriate endpoint to allow comparison across multiple animals and to time treatment experiments. As a result, there can be a large range of tumour sizes or burden at the endpoint or at the point of treatment initiation.

1.2.4 Opportunity to Improve Preclinical Studies

In order to obtain additional information regarding tumour growth, the ability to evaluate all tumours *in vivo* is necessary. This would open the possibility of obtaining time-series data to assess growth over time of different tumours, tumour models and treatments. Longitudinal monitoring could provide information on how a treatment alters tumour growth - whether tumours regress, stabilize or merely grow more slowly. This information cannot be obtained from endpoint analysis alone. Although caliper measurements can be obtained *in vivo*, the limitations discussed above mean an improved method is still required for subcutaneous tumours.

Additionally, growth data from each tumour would allow each tumour to act as its own control to compare pre- and post-treatment growth or change in size. Many cancer therapies are effective on proliferating tumours but show little effect on non-proliferating tumours. Obtaining pre-treatment growth estimates has been recommended [8] to aid in obtaining the desired data and allow an evaluation of the post-treatment results. If there is a wide range in the response to a particular treatment, it can still be observed if tumour growth is observed on a tumour by tumour basis.

Non-invasive *in vivo* imaging would allow for this type of monitoring and data acquisition and can be performed on a micro-imaging scale to allow for the visualization of tumours in small-animal models. The additional ability to visualize tumours *in vivo*

is becoming necessary as models and treatments become more sophisticated [4].

1.3 *In vivo* Micro-Imaging

Non-invasive imaging of preclinical cancer models enables tumours to be repeatedly imaged at different time points, allowing longitudinal measurements of the same tumour. For each model of interest, an appropriate imaging modality is required to perform these measurements. The criteria required of the imaging modality include high resolution to obtain accurate and precise tumour measurements. Additionally, appropriate contrast to visualize the desired tumours is required. Endogenous contrast simplifies the imaging procedure making it the first choice. If there is insufficient contrast due to the natural tissue properties, it may be possible to use an exogenous contrast agents which could provide the necessary contrast to differentiate tumours from health tissue. Being able to minimize the use of contrast agents leads to more rapid throughput for imaging and less potential for bioeffects due to contrast agents.

If repeated measurements are desired over a short period of time, it is preferable to avoid techniques with cumulative exposure effects and potential biological effects from high exposure as this may limit the scans that can be performed. Modalities that can limit the scan time per animal are also desirable as this will minimize anesthetic exposure for each animal and allow for larger cohorts of animals to be evaluated on an ongoing basis.

The ability to measure tumour volumes over time is necessary for all tumours. Measurements of tumours conducive to caliper measurements will benefit from improved measurement accuracy and variability, while other tumours will benefit from longitudinal measurement of the same tumours over time.

There are imaging modalities scaled for preclinical imaging, frequently referred to as micro-imaging, analogous to clinical imaging modalities, with the addition of fluorescence and luminescence optical imaging techniques, which are not currently

used in the clinic. As these modalities have been scaled to reduce the resolution making them appropriate for preclinical studies, small animal anatomy can be visualized with comparable detail to clinical images in patients, with typical resolution on the order of $100\ \mu\text{m}$. While there are often imaging modalities that can be used for a given experiment, there are a number of advantages and disadvantages to every modality that need to be evaluated for the particular experiments.

1.4 Ultrasound

1.4.1 Comparison of High-Frequency Ultrasound to Other Micro-Imaging Modalities

Ultrasound has inherent soft-tissue contrast, without requiring exogenous contrast agents, making it a practical option for soft-tissue tumour imaging. However, ultrasound cannot be used for lung and bone imaging therefore preventing its use for primary and metastatic tumours in those organs. Ultrasound offers many benefits for *in-vivo* imaging of mouse models, including real-time 2D imaging capabilities (up to 100 frames per second), 3D imaging capabilities, portability and comparatively low equipment and operational costs [9]. However, ultrasound has some distinct limitations compared to other modalities, including a relatively small field of view, limitations on imaging where there is an appropriate acoustic window, and inability to image bone and lung tissue.

X-ray computed tomography (CT) allows for entire body scans to be acquired, which is particularly useful when examining mouse phenotypes. CT provides excellent contrast for bone and lung imaging; however, it does not provide soft-tissue contrast without the use of exogenous contrast agents.

Magnetic resonant imaging (MRI) provides high spatial resolution and contrast in soft-tissue, as well as full-body scans. However, the equipment and operational costs

are very high for MRI [9]. Both CT and MRI scan times can vary depending on the protocol and imaging system from being fairly rapid to very long (on the order of hours for MRI). Even with fast scans, the reconstruction times can be very long so that the images of the animal are often not available until long after they have been scanned. Although the scanning and reconstruction times are long, the segmentation step often requires very little time as automated segmentation algorithms can be used. In ultrasound imaging, manual segmentation is typically required which takes a considerable amount of time, especially in 3D images [10]

Nuclear medicine scans, including positron emission tomography (PET) and single photon emission computed tomography (SPECT) provide functional information regarding tumour activity; however, they suffer from low spatial resolution and require the use of radioisotopes to provide contrast increasing cost and dose to the animal [9].

Recently high-frequency ultrasound has been used a wide variety of application, including cardiac [11, 12], embryonic [13], vascular [14] and image guided interventions [15]. The high frame rates are particularly useful in many of these applications as it allows for real time monitoring of the changes *in vivo*.

1.4.2 High-Frequency Compared to Clinical Ultrasound

There are a number of notable differences between high-frequency and clinical ultrasound imaging. Clinical imaging is typically performed in the range of 5 to 12 MHz with the high end being at 15 MHz, compared to 20 MHz to 60 MHz for high-frequency ultrasound [16]. The increased frequency improves the resolution (approximately $40 \times 80 \mu\text{m}$ at 40 MHz), allowing smaller objects to be imaged, while sacrificing the depth of penetration (approximately 10 mm at 40 MHz). High-frequency ultrasound has reduced penetration compared to clinical ultrasound due to the increased attenuation at higher frequencies which reduces the energy in the ultrasound pulse as it propagates. These trade-offs make it appropriate as a micro-imaging modality for small animal imaging where the increased resolution is required and limited field of

view is still sufficient for the subject size. Clinically, high-frequency ultrasound is restricted to several specialized areas, including intravascular, ocular and dermatological imaging.

High-frequency ultrasound produces signals from different scatterers [17] than is the case when employing lower (clinical) frequencies, meaning the ultrasound images provide information about different structural components in tissue.

1.4.3 Ultrasound Arrays and Focusing

While the axial resolution in ultrasound is determined by the transmitted pulse length, the lateral resolution (LR) depends on the focusing from the transducer. The lateral resolution at the focus for a transducer with a rectangular aperture is defined as $LR = \lambda \frac{F}{D}$ where F is the focal distance, D is the transducer aperture diameter and λ is the ultrasound wavelength. In the case of single-element transducers, they can be either unfocused or focused at a fixed point in space. The lateral resolution at the focus can be improved by moving the fixed focus towards the transducer or increasing the transducer diameter. The focal depth is typically chosen based on the application and penetration depth at the imaging frequency, which constrains how much it can be varied. The width of the transducer can be varied, but there are obvious practical limitations on the size of a transducer to make it practical for imaging. Even if the lateral resolution at the focal depth is acceptable, with tightly focused transducers the lateral resolution rapidly degrades away from the fixed focus. To overcome these limitations, transducer arrays can be employed and are used clinically for imaging.

Arrays use multiple small elements to send and receive sound. There is a range of designs of array transducers; however, they all allow the ultrasound beam to be focused and in some cases also steered. Since there are many elements in an array (typically 128 to 256 individual elements), it is possible to use a varying number of elements to send and receive thereby changing the transducer aperture width. The subset of elements used to send and receive are referred to as the subaperture. Having

multiple elements additionally allows each element to send pulses at different times, which can be used to focus the ultrasound beam at a desired point in space. This is achieved by calculating how long it will take sound to travel from each element to the desired focal point. These delays are then used to time when each element sends its pulse so that all pulses arrive at the desired focal point at the same time. The same technique can be applied when the returned signals reach the transducer, appropriate delays can be applied to the received signals to align all the signals from the desired focal point. Since the focal depth can be chosen and changed, it is possible to focus at multiple depths to keep the image in focus instead of having the image in focus at only one depth. Being able to change both the focal depth and the subaperture size allows for the f-number to be varied more readily than in the case of fixed focus transducers, thereby allowing for the lateral resolution to be maintained over a larger region of the image.

Although arrays provide great improvement to image quality over single element transducers, due to a number of physical limitations, arrays are not yet commonly available for high-frequency ultrasound imaging. Linear and phased arrays require an element spacing of less than λ , the ultrasound wavelength, and $\lambda/2$ respectively, to prevent energy from the grating lobes being imaged [18]. If the beamsteering is restricted to small angles, the element spacing can be relaxed slightly. As an example, for a 40 MHz array an element spacing of less than $38.5 \mu\text{m}$ for a linear array and $19.3 \mu\text{m}$ for a phased array would be required. Cutting the piezoelectric material alone can be technically difficult as elements this small are prone to falling off the backing. Bonding an electronic channel to the back of each array element can also be very technically challenging.

There are several groups currently working on developing arrays for high-frequency ultrasound including work on linear arrays [19, 20] as well as annular arrays [21-23] which allow for depth focusing, but still require the transducer to be swept to produce an image.

1.4.4 Early Applications of High-Frequency Ultrasound to Cancer

One of the earliest applications of high-frequency ultrasound imaging was cancer. In 1987, Sherar *et al.* [24] demonstrated the ability to differentiate the internal structure of tumour spheroids (less than 1.5 mm diameter) using a 100 MHz transducer. Subsequently, Turnbull *et al.* [10] were able to monitor the growth of murine melanoma over time using 50 MHz and 62 MHz ultrasound. Tumours ranging in size from 1 to 4 mm diameter were imaged in this study. Two-dimensional measurements were used to measure tumour growth and several 3D images were reconstructed to provide proof of principle [10].

Recently high-frequency ultrasound has been employed in many preclinical oncology studies including growth, treatment and blood flow studies. In mouse models a variety of soft-tissue tumours have been imaged including melanoma [25] and prostate carcinoma [26]. Other animal models have also been successfully imaged using high-frequency ultrasound, including a chemically induced liver metastases model in Zebrafish [27].

1.5 Hypothesis and Objectives

It is hypothesized that the refinement of experimental design and focusing methods for high-frequency ultrasound will enable precise longitudinal analysis of tumour growth ranging from 0.1 to 10 mm diameter in a mouse model of liver metastasis. To test this hypothesis, several objectives were set:

1. To determine if longitudinal imaging of soft-tissue tumours with high-frequency ultrasound can be performed without the use of exogenous contrast agents, based solely on the inherent tissue contrast.
2. To quantify the volume measurement variability and use this information to help

plan longitudinal imaging experiment timing to ensure statistically significant changes in volume will be observed.

3. To employ simulations of tumour growth as a method to determine appropriate imaging parameters that will enable a function more complex than an exponential to be fitted and allow changes in growth rates to be observed.
4. To apply synthetic aperture focusing techniques retrospectively to acquired data in order to decrease the lesion-size measurement variability and determine whether adaptive or conventional apodization* is more effective at reducing variability.

1.6 Thesis outline

This thesis examines how 3D high-frequency ultrasound can be employed to provide quantitative information on tumour growth in mouse models. In particular, the above objectives are addressed, one in each of the following chapters.

1.6.1 Chapter 2

Chapter 2 demonstrates the feasibility of using high-frequency ultrasound to monitor the growth of tumours in soft tissue by imaging using endogenous contrast in an experimental liver metastasis model. A liver metastasis model was chosen as metastases cannot be tracked without non-invasive imaging due to their location. Liver metastases also tend to be multifocal, thereby providing multiple tumours to track growth on an individual tumour basis. A syngeneic melanoma model, as well as a human colon carcinoma, human breast carcinoma and mouse fibroblast model were

*Literally meaning to 'remove the foot.' The amplitude weighting of the aperture to reduce the side-lobes in the ultrasound beam. Typically a smooth function with the edges weighted less than the centre.

used to establish that a range of tumours can be imaged with ultrasound. Even in this multifocal model, individual tumours could be identified on different days to allow each tumour to be tracked independently. Where it was possible to obtain volume measurements, growth curves over time were plotted and an example treatment study was conducted using the melanoma model.

1.6.2 Chapter 3

Chapter 3 evaluates the measurement variability of the 3D tumour volume to enable more rigorous interpretation of the growth curves as obtained in Chapter 2. An intra- and inter-observer variability study was conducted to evaluate the measurement variability as a function of tumour size. The effects of breathing motion and the limited depth of field in the image were evaluated. The results are used to determine the minimum detectable change in volume and consequently the limitations on observing treatment effects. As a first approximation, the length of time required to wait between imaging sessions to obtain a change in volume that can be measured was plotted as a function of tumour volume and doubling time.

1.6.3 Chapter 4

To extend the planning of imaging sessions from the end of Chapter 3, Chapter 4 presents simulations of tumour growth based on analysis of experimentally obtained growth data, commonly used imaging parameters and the measurement variability. The ability of an exponential or Gompertz functions to fit the data was analyzed as a function of a variety of parameters. The parameters examined fit into different categories including biological characteristics of the animal model (ie. tumour growth rate), user selectable imaging parameters (ie. imaging schedule) and imaging system parameters that could be improved upon (ie. measurement variability, minimum size of detection, maximum size from field of view.) The results provide

information on how to structure experiments to obtain the most data and indicate which improvements would provide the most benefit.

1.6.4 Chapter 5

To improve curve fitting of tumour growth, one approach is to reduce measurement variability. In chapter 5, synthetic aperture focusing techniques were applied to 2D ultrasound images to assess the effect on measurement variability. Both conventional and adaptive weighting techniques were assessed, with adaptive techniques requiring more computational power. Improved focusing has the potential to decrease the measurement variability and increase tumour conspicuity for small tumours.

References

- [1] A. Kamb, "What's wrong with our cancer models?," *Nat. Rev. Drug Discov.*, vol. 4, pp. 161–165, 2005.
- [2] I. Kola and J. Landis, "Can the pharmaceutical industry reduce attrition rates?," *Nat Rev Drug Discov*, vol. 3, no. 8, pp. 711–6, 2004.
- [3] K. P. Olive and D. A. Tuveson, "The use of targeted mouse models for preclinical testing of novel cancer therapeutics," *Clin Cancer Res*, vol. 12, no. 18, pp. 5277–87, 2006.
- [4] T. Van Dyke and T. Jacks, "Cancer modeling in the modern era: progress and challenges," *Cell*, vol. 108, no. 2, pp. 135–44, 2002.
- [5] B. Weiss and K. Shannon, "Mouse cancer models as a platform for performing preclinical therapeutic trials," *Current Opinion in Genetics and Development*, vol. 13, pp. 84–89, 2003.
- [6] A. F. Chambers, A. C. Groom, and I. C. MacDonald, "Dissemination and growth of cancer cells in metastatic sites," *Nat Rev Cancer*, vol. 2, no. 8, pp. 563–72, 2002.
- [7] T. Voskoglou-Nomikos, J. L. Pater, and L. Seymour, "Clinical predictive value of the in vitro cell line, human xenograft, and mouse allograft preclinical cancer models," *Clin Cancer Res*, vol. 9, no. 11, pp. 4227–39, 2003.
- [8] , "
- [9] R. Weissleder, "Scaling down imaging: molecular mapping of cancer in mice," *Nat Rev Cancer*, vol. 2, no. 1, pp. 11–8, 2002.

- [10] D. H. Turnbull, J. A. Ramsay, G. S. Shivji, T. S. Bloomfield, L. From, D. N. Sauder, and F. S. Foster, "Ultrasound backscatter microscope analysis of mouse melanoma progression," *Ultrasound Med Biol*, vol. 22, no. 7, pp. 845-53, 1996.
- [11] Yu-Qing Zhou, Yonghong Zhu, Jonathan Bishop, Lorinda Davidson, R. Mark Henkelman, Benoit G. Bruneau, and F. Stuart Foster, "Abnormal cardiac in-flow patterns during postnatal development in a mouse model of holt-oram syndrome," *Am J Physiol Heart Circ Physiol*, vol. 289, no. 3, pp. H992-1001, 2005.
- [12] R. Williams, A. Needles, E. Cherin, Y.-Q. Zhou, R. M. Henkelman, S. L. Adamson, and F. S. Foster, "Noninvasive ultrasonic measurement of regional and local pulse-wave velocity in mice," *Ultrasound in Medicine & Biology*, vol. 33, no. 9, pp. 1368-75, 2007.
- [13] Rui Ping Ji, Colin K. L. Phoon, Orlando Aristizabal, Kathleen E. McGrath, James Palis, and Daniel H. Turnbull, "Onset of cardiac function during early mouse embryogenesis coincides with entry of primitive erythroblasts into the embryo proper," *Circ Res*, vol. 92, no. 2, pp. 133-135, 2003.
- [14] D. E. Goertz, J. L. Yu, R. S. Kerbel, P. N. Burns, and F. S. Foster, "High-frequency doppler ultrasound monitors the effects of antivascular therapy on tumor blood flow," *Cancer Res*, vol. 62, no. 22, pp. 6371-5, 2002.
- [15] C. Waspe Adam, H. Jason Cakiroglu, C. Lacefield James, and Fenster Aaron, "Design, calibration and evaluation of a robotic needle-positioning system for small animal imaging applications," *Physics in Medicine and Biology*, , no. 7, pp. 1863, 2007.
- [16] F. S. Foster, M. Y. Zhang, Y. Q. Zhou, G. Liu, J. Mehi, E. Cherin, K. A. Harasiewicz, B. G. Starkoski, L. Zan, D. A. Knapik, and S. L. Adamson, "A new ultrasound instrument for in vivo microimaging of mice," *Ultrasound Med Biol*, vol. 28, no. 9, pp. 1165-72, 2002.

- [17] Linda R. Taggart, Ralph E. Baddour, Anoja Giles, Gregory J. Czarnota, and Michael C. Kolios, "Ultrasonic characterization of whole cells and isolated nuclei," *Ultrasound in Medicine & Biology*, vol. 33, no. 3, pp. 389–401, 2007.
- [18] K. K. Shung, "Diagnostic ultrasound imaging and blood flow measurements," CRC Taylor & Francis Group, New York, 2006.
- [19] J. M. Cannata, J. A. Williams, Zhou Qifa, T. A. Ritter, and K. K. Shung, "Development of a 35-MHz piezo-composite ultrasound array for medical imaging," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 53, no. 1, pp. 224–236, 2006.
- [20] M. Lukacs, Yin Jianhua, Pang Guofeng, R. C. Garcia, E. Cherin, R. Williams, J. Mehi, and F. S. Foster, "Performance and characterization of new micromachined high-frequency linear arrays," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 53, no. 10, pp. 1719–29, 2006.
- [21] J. A. Brown and G. R. Lockwood, "A digital beamformer for high-frequency annular arrays," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 52, no. 8, pp. 1262–1269, 2005.
- [22] J. A. Ketterling, S. Ramachandran, and O. Aristizabal, "Operational verification of a 40-MHz annular array transducer," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 53, no. 3, pp. 623–630, 2006.
- [23] E. J. Gottlieb, J. M. Cannata, Hu Chang-Hong, and K. K. Shung, "Development of a high-frequency (> 50 MHz) copolymer annular-array, ultrasound transducer," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 53, no. 5, pp. 1037–1045, 2006.
- [24] M. D. Sherar, M. B. Noss, and F. S. Foster, "Ultrasound backscatter microscopy

- images the internal structure of living tumour spheroids," *Nature*, vol. 330, no. 6147, pp. 493-5, 1987.
- [25] A. M. Cheung, A. S. Brown, L. A. Hastie, V. Cucevic, M. Roy, J. C. Lacefield, A. Fenster, and F. S. Foster, "Three-dimensional ultrasound biomicroscopy for xenograft growth analysis," *Ultrasound Med Biol*, vol. 31, no. 6, pp. 865-70, 2005.
- [26] L. A. Wirtzfeld, G. Wu, M. Bygrave, Y. Yamasaki, H. Sakai, M. Moussa, J. I. Izawa, D. B. Downey, N. M. Greenberg, A. Fenster, J. W. Xuan, and J. C. Lacefield, "A new three-dimensional ultrasound microimaging technology for preclinical studies using a transgenic prostate cancer mouse model," *Cancer Res*, vol. 65, no. 14, pp. 6337-45, 2005.
- [27] W. Goessling, T. E. North, and L. I. Zon, "Ultrasound biomicroscopy permits in vivo characterization of zebrafish liver tumors," *Nat Meth*, vol. 4, no. 7, pp. 551-3, 2007.

Chapter 2

Three-Dimensional High-Frequency Ultrasound Imaging For Longitudinal Evaluation Of Liver Metastases In Preclinical Models

The content of this chapter has been adapted from: "Three-dimensional high-frequency ultrasound imaging for longitudinal evaluation of liver metastases in pre-clinical models", published in Cancer Research, vol 65 (12) June 15 2005, by Kevin C. Graham, Lauren A. Wirtzfeld, Lisa T. MacKenzie, Carl O. Postenka, Alan C. Groom, Ian C. MacDonald, Aaron Fenster, James C. Lacefield, and Ann F. Chambers. This chapter has been included as it contains the background information around which the subsequent chapters are built. The liver metastasis images in this chapter are subsequently used in chapter 3 where the measurement variability is analysed. Chapter 4 bases the growth of the modeled tumours on the growth rates obtained from

*the tumours within this paper and uses the measurement from chapter 3 to make the values more realistic. Chapter 5 has some data from the liver metastasis model as well as phantoms made to have contrast similar to the in-vivo liver metastasis model with the aim of improving the images within the liver. Although, this is a second author paper and also appears in Kevin Graham's thesis *, the Biomedical Engineering Graduate program approved the inclusion of this material in this thesis.*

2.1 Introduction

Metastasis, the dissemination and growth of cancer cells in a secondary organ, is the leading cause of cancer mortality. The liver is a frequent metastatic site for melanoma, colon and breast cancer, and therefore an important area of metastasis research. Preclinical animal models, such as the mouse, are essential to the study of liver metastasis, yet their utility has been limited by difficulty in tracking the progression of metastases through time. Noninvasive longitudinal imaging would decrease experimental variability, provide a more accurate assessment of metastatic progression and the efficacy of therapeutic interventions, and allow the study of dynamic or time varying processes such as tumour vascularization and dormancy.

Many preclinical imaging modalities are under development, including magnetic resonance imaging (MRI), x-ray computed tomography (CT), positron emission tomography (PET), and fluorescent and bioluminescent imaging, yet no single modality should be considered a comprehensive solution for cancer micro-imaging applications. Each modality possesses a unique combination of strengths and weaknesses that impact their selection for use in a particular study. In general, desirable characteristics in a noninvasive imaging modality would be high-resolution to allow detection of min-

*Graham, Kevin C. *Noninvasive imaging for the study of preclinical liver metastasis models* [Ph.D. dissertation]. Canada: The University of Western Ontario (Canada); 2007.

imal disease, cost effectiveness and rapid image acquisition to facilitate throughput, inherent contrast between the liver parenchyma and tumour to avoid genetically encoded or endogenously administered contrast agents, and applicability to a range of liver metastasis models. MRI offers high-resolution imaging yet may be time consuming and relatively expensive to purchase and operate [1]. X-ray CT also offers high resolution, but poor soft tissue contrast necessitates the use of radio-opaque contrast agents and radiation dosage may limit longitudinal imaging [1]. The resolution of PET does not match that of MRI or CT, and the requirement for production and containment of radionuclides can make costs prohibitive [1]. Fluorescent and bioluminescent imaging offer a relatively cost effective way to study liver metastases, but suffer from poor resolution and the requirement to transfect endogenous reporter genes into the cell line of interest [1, 2]. The expression of foreign reporter proteins may lead to increased immunogenicity, and thus must be carefully examined for their impact on the metastatic model being studied [3-5].

Ultrasound is an attractive option for preclinical imaging due to the cost and time efficiencies of this modality. Previous studies using high-frequency ultrasound imaging of murine cancer models demonstrated the feasibility of this technique to track subcutaneous tumour progression [6]. That study concluded that further application of ultrasound imaging would require a fast method for generating 3D images. A new high-frequency scanner that employs 3D image acquisition methods and reconstruction software developed in our laboratory has addressed this limitation [7]. These developments raised the possibility of using high-frequency ultrasound in the evaluation of clinically relevant metastatic models, which are difficult to study in a noninvasive fashion.

We report here the application of a high-frequency (40 MHz) ultrasound system with three-dimensional imaging capabilities to the study of murine liver metastasis.

We demonstrated that the resolution of high-frequency ultrasound allowed detection of liver metastases at a minimum size that compared favourably to that of MRI, CT or optical methods [8–11]. The applicability of this technique was demonstrated by identifying and tracking liver metastases from four tumour cell lines of different tumour origins. The importance of 3D volumetric imaging to reduce uncertainty in volume determination was established by comparison of 3D segmented volumes with the commonly assumed ellipsoidal volume calculated from diameter measurements. The utility of 3D high-frequency ultrasound in the evaluation of chemotherapeutics was shown in a preclinical trial with doxorubicin. These results demonstrate that the cost and time efficiencies of traditional ultrasound coupled with the 3D capabilities and high resolution of this high-frequency ultrasound system make this modality particularly well-suited to the study of liver metastasis in a wide range of preclinical models.

2.2 Materials and Methods

2.2.1 Cell Culture and Experimental Metastasis Models

B16F1 [12] and HT-29 cells (Cat# CRL-6323 and HTB-38, respectively; American Type Culture Collection, Manassas, VA, USA) were maintained in α MEM with L-glutamine, ribonucleosides and deoxyribonucleosides (Invitrogen, Carlsbad, CA, USA) supplemented with 10% FBS (Sigma, Mississauga, ON, Canada). PAP2 cells [13, 14] were maintained in DMEM (Invitrogen) supplemented with 10% FCS (Invitrogen). MDA-MB-435/HAL cells were maintained in EMEM (Invitrogen) supplemented with 2 mM L-glutamine, 100 μ M non-essential amino acids, 25 mM HEPES buffer, 1 mM sodium pyruvate (Invitrogen) and 1X MEM vitamin solution (Sigma). The MDA-MB-435/HAL line was derived from the MDA-MB-435 cell line by an *in vivo* selection procedure for increased metastatic potential [15]. All animals were

cared for in accordance to the Canadian Council on Animal Care guidelines, under a protocol approved by the University of Western Ontario Council on Animal Care. For experimental metastasis assays, mice were anesthetized with an intraperitoneal injection of xylazine/ketamine (2.6 mg ketamine and 0.13 mg xylazine per 20 g body mass). As described previously, anesthetized mice received mesenteric injections of 0.1 mL of cells suspended in their respective growth media to target the liver [16]. For B16F1 cells, a suspension of 3×10^5 cells was injected into C57BL/6 mice (Harlan, Indianapolis, IN, USA). For HT-29 cells, a suspension of 2×10^6 cells was injected into NIH III mice (Charles River, Wilmington, MA, USA). For PAP2 cells, a suspension of 2×10^5 cells was injected into SCID mice (Charles River). For MDA-MB-435/HAL cells, a suspension of 3×10^5 cells was injected into SCID mice (Charles River). All mice were female and 7-11 weeks old at the time of cell injection.

2.2.2 Ultrasound Imaging

For ultrasound imaging the Vevo 660 high-frequency ultrasound system (Visual-Sonics Inc. Toronto, ON, Canada) was used. The Vevo 660 is the second generation of a system described previously [17]. The Vevo 660 ultrasound probe has a 40 MHz centre frequency with a 6 mm focal depth. The spatial resolution at the focus is $40 \times 80 \times 80 \mu\text{m}^3$. Prior to the first imaging session the mouse's abdomen was depilated with commercial hair removal cream. During imaging the mouse was kept under anesthesia with 1.5% isoflurane in oxygen and restrained on a heated stage. During imaging with the immunodeficient NIH III and SCID mice the animals were handled and imaged in a HEPA filtered workstation (Microzone Corp. Ottawa, ON, Canada). Ultrasound is strongly reflected by the ribcage, which hinders imaging of any tissue located beneath the ribs, such as the lungs and a portion of the liver. Thus, the volume of liver tissue accessible for ultrasound imaging may vary among animals and between imaging sessions for the same animals. In general, we found that a significant volume of the left lateral, left medial and right medial liver lobes were routinely

accessible for imaging. During imaging two-dimensional images were acquired in the sagittal plane after ultrasound contact gel was applied to the abdomen. For three-dimensional imaging, parallel two-dimensional images were acquired by stepping the transducer in 30 μm intervals in the out of plane dimension. Using software developed in our laboratory, two-dimensional images were interpolated and reconstructed on-line to create a 3D volumetric image [7]. The system can acquire and produce a typical three-dimensional image in less than 20 seconds. The 3D reconstruction software is available through VisualSonics Inc., or through the authors for research purposes.

2.2.3 Volume Measurements

To determine tumour volume, the boundaries of a metastasis were outlined within parallel planes separated by 50 μm in the volumetric image. The total metastasis volume was calculated by summing the outlined areas and multiplying by the interslice distance [18]. Segmented volumes were compared to ellipsoid volumes estimated using the formula $V = (\pi abc)/6$. The measurements for diameters 'a', 'b', and 'c', were obtained from the 3D volumetric images. The sagittal plane showing the greatest tumour diameter was selected, and the greatest diameter 'a' measured. The diameter 'b', perpendicular to 'a' was then measured. The volume was then rotated and the transverse plane showing the largest tumour diameter was selected. The diameter 'c', perpendicular to both 'a' and 'b' was then measured. To determine the % difference between the ellipsoid and 3D segmented volumes the following formula was used

$$100\% \times \frac{(\text{ellipsoid volume}) - (\text{segmented volume})}{(\text{ellipsoid volume} + \text{segmented volume})/2} \quad (2.1)$$

2.2.4 Longitudinal Growth Measurements

For longitudinal imaging, the initial imaging timepoint was based on previous indications of when micrometastases could first be detected by ultrasound. Individual liver metastases were identified and a 3D image recorded. Individual liver metastases were identified on successive imaging dates by their particular liver lobe location, tumour shape, and proximity to landmark structures such as major blood vessels or the liver edges. Landmarks were all internal to the liver, since the liver lobes move in relation to any external landmarks, such as the ribs. Animals were sacrificed due to escalating tumour burden, as assessed by ultrasound imaging, or when at least four imaging time points had been acquired to construct a growth curve. Approximately 5 minutes were required to locate, identify and image the liver metastases of each mouse. If the time spent on set-up, animal handling, anesthesia, and recovery is included, the average duration of an imaging session was 15 minutes per mouse.

2.2.5 Treatment Protocols

The B16F1 liver metastasis model was used to assess the ability of high-frequency ultrasound to detect response to cytotoxic chemotherapeutic agents. At day 7 post cell injection the first treatment with doxorubicin was administered. Doxorubicin (Pharmacia, Mississauga, ON, Canada) was given at a previously described treatment schedule (1 mg/kg, 0.1 mL, i.p.) every second day until day 17 post cell injection, for a total of 6 treatments [19]. Control animals received saline control injections (0.1 mL, i.p.). Ultrasound imaging was performed from day 8 post cell injection, the earliest time B16F1 liver metastases could be detected in the images, until the end of the experiment.

2.2.6 Histology

At sacrifice the mouse liver was excised and fixed in 10% neutral buffered formalin. Visual inspection validated the tumour size and location depicted by ultrasound imaging. For histological confirmation of tumour size, formalin fixed paraffin embedded livers were sectioned (4 μm slice thickness) and stained with hematoxylin and eosin.

2.3 Results

2.3.1 Identification of murine liver metastases using high-frequency ultrasound

To validate ultrasound imaging for the detection of murine liver metastases, mice were noninvasively imaged and, once suspected metastases were detected with ultrasound, the animal was sacrificed. Gross pathology and histological sections verified the presence of a tumour, its size and location. Ultrasound reliably detected murine liver metastases from the four tumour cell lines tested, B16F1, HT-29, MDA-MB-435/HAL and PAP2, with excellent agreement between ultrasound images, gross pathology and histological sections (Fig. 2.1). Ultrasound imaging proved highly sensitive to small metastases with a minimum detection size (maximum diameter \rightarrow segmented volume) of 0.22 mm \rightarrow 0.01 mm³, 0.47 mm \rightarrow 0.03 mm³, 0.66 mm \rightarrow 0.08 mm³, and 0.78 mm \rightarrow 0.17 mm³ for B16F1, HT 29, MDA-MB-435/HAL and PAP2 tumours respectively. As a point of reference, a volume of 0.01 mm³ would be produced by approximately 6000 cells, based on the assumption of a spherical cell volume and a cell diameter of 15 μm .

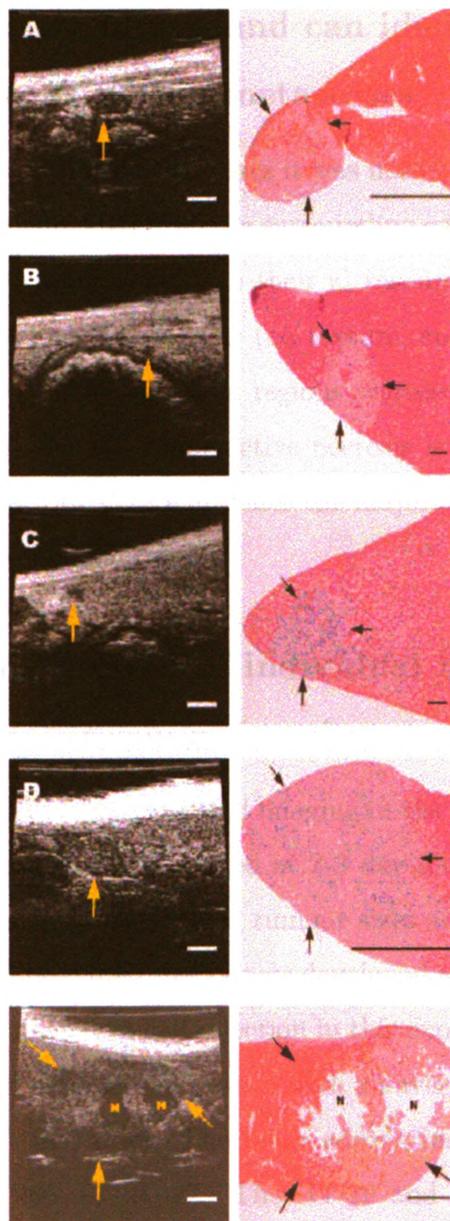


Fig. 2.1: Identification of B16F1 (A, E), HT-29 (B), MDA-MB-435/HAL (C), and PAP2 (D) liver metastases by noninvasive ultrasound imaging. Two dimensional ultrasound images are shown with corresponding histological sections. The location of each metastasis in the ultrasound image is denoted with yellow arrows, in the histological section with black arrows. The maximum diameter of the metastasis shown in the ultrasound image is 1.56 mm (A), 0.47 mm (B), 0.66 mm (C), 2.33 mm (D), and 5.01 mm (E). The scale bars on the ultrasound images are 1.00 mm. The scale bars on the histological sections are 1.00 mm (A, D, E) or 0.10 mm (B, C). In panel E, necrotic areas (N) depicted by ultrasound imaging and confirmed by histology are labelled.

2.3.2 High-frequency ultrasound can identify areas of liquefactive necrosis within metastases

During imaging of B16F1 liver metastases it was noted that, although the metastases were always clearly delineated from the surrounding parenchyma, the metastases showed a large amount of heterogeneity in their ultrasound backscatter. In a number of metastases distinct anechoic regions (no texture, appears dark) were evident. Histological examination of these anechoic regions revealed that they are regions of liquefactive necrosis (Fig. 2.1 E). Liquefactive necrosis is expected to be anechoic, since the breakdown of necrotic cells eliminates the majority of ultrasound scattering structures from that region of the tumour.

2.3.3 Tracking the growth of individual liver metastases by noninvasive ultrasound imaging

To demonstrate the utility of ultrasound imaging in the longitudinal study of liver metastases, mice were noninvasively imaged at 2-3 day intervals. At sacrifice, gross pathology and histological sections verified tumour sizes and locations depicted during ultrasound imaging. The B16F1 metastases developed rapidly, forming detectable metastases as early as 10 days post cell injection in this experiment (Fig. 2.2 A). The B16F1 metastases showed exponential growth with an average volume doubling time of 1.2 ± 0.2 (mean \pm SD) days. The mean correlation coefficient for fitting an exponential curve was 0.966 ± 0.047 . The HT-29 and MDA-MB-435/HAL liver metastases were much slower to develop, forming detectable metastases at a minimum of 33 days post cell injection (Fig 2.2 B). The HT-29 and MDA-MB-435/HAL metastases also showed exponential growth with doubling times ranging from 3.7 - 4.8 days for the HT-29 metastases and 5.7 - 10.4 days for MDA-MB-435/HAL metastases. The correlation coefficients ranged between 0.972 - 0.993 for HT-29, and 0.635 - 0.889 for MDA-MB-435/HAL. The metastasis HT-29 - 4 was not included in the

range of doubling times because its volume did not increase over the 10-day interval that it was imaged. Representative two-dimensional images from the longitudinal imaging of an individual B16F1 metastasis, B16F1 - E, are shown (Fig 2.2 C). Individual PAP2 liver metastases could not be evaluated for longitudinal growth because in this highly aggressive model numerous metastases form and quickly fuse. In such cases, ultrasound could be used to monitor increasing tumour burden, instead of the growth of individual metastases, as an indicator of tumour progression.

2.3.4 2D measurement provides frequent overestimation or underestimation of tumour volume as compared to 3D measurement

Tumours are often assumed to have an ellipsoid shape, which allows a volume to be calculated from the maximum widths and length in 2D images. Tumour volumes calculated from this 2D method and from 3D segmentation were compared to determine if the assumption of an ellipsoid shape was valid for liver metastases. In the liver metastasis models examined here, there were large differences in the measured tumour volumes between the 3D and 2D methods. The mean percent difference for B16F1 liver metastases was $-8.8 \pm 23.5\%$ (range 90.1% to 53.2%) (Fig 2.3 A). The negative mean indicates that the ellipsoid volume was on average smaller than the 3D segmented volume. For the MDA MB-435/HAL metastases the mean percent difference was $-15.0 \pm 25.3\%$ (range -45.2% to 23.1%) and for the HT-29 metastases $-7.9 \pm 43.8\%$ (range 106.5% to 80.5%). Three-dimensional surface rendering of the liver metastases allowed visualization of the irregular shapes of some tumours (Fig 2.3 B-C).

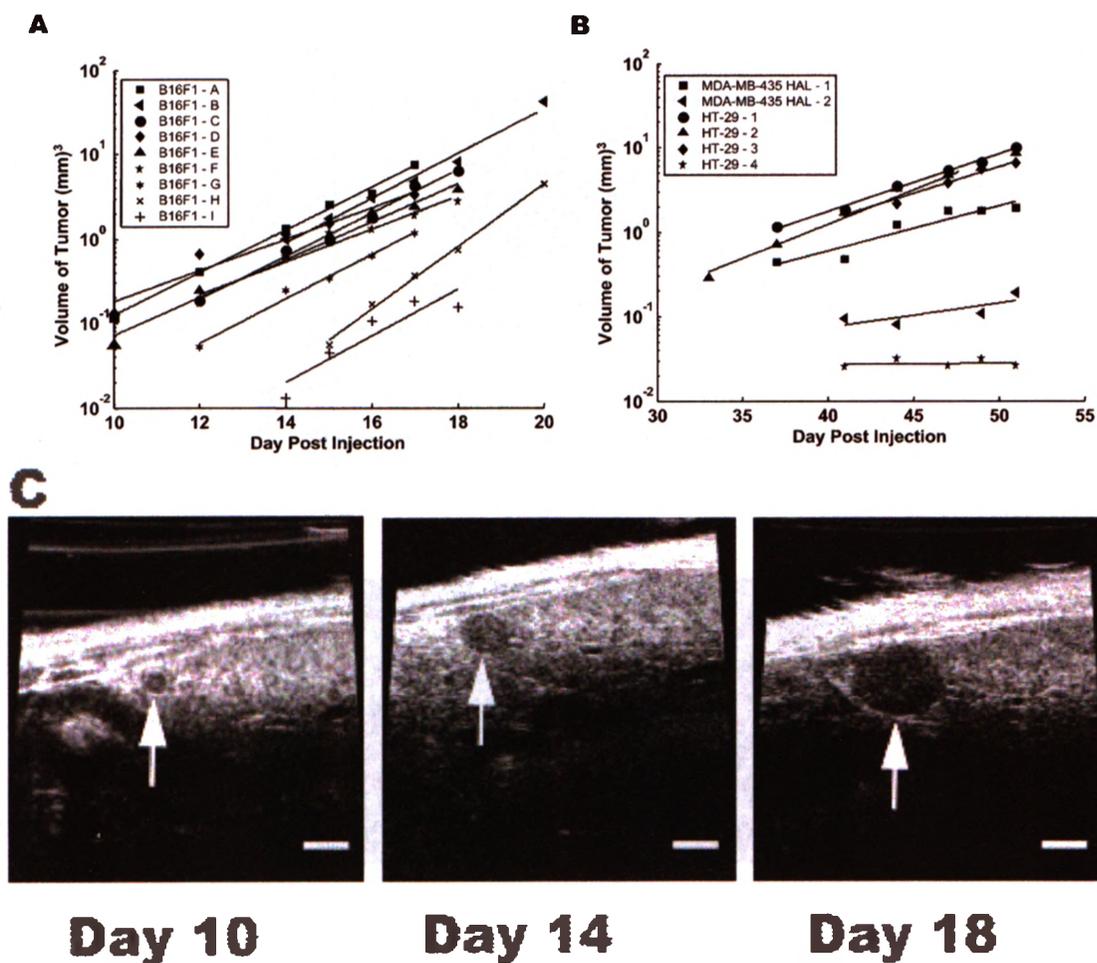


Fig. 2.2: Tracking the growth of individual liver metastases by noninvasive ultrasound imaging. (A) Growth curves of B16F1 liver metastases, plotted on a semi-logarithmic scale. (B) Growth curves of HT-29 and MDA-MB-435/HAL liver metastases, plotted on a semi-logarithmic scale. (C) Representative two dimensional ultrasound images of B16F1 - E. Size of B16F1 - E (maximum diameter \rightarrow segmented volume) is 0.50 mm \rightarrow 0.06 mm^3 (Day 10), 1.07 mm \rightarrow 0.61 mm^3 (Day 14), and 2.09 mm \rightarrow 3.79 mm^3 (Day 18). The scale bars on the ultrasound images are 1.00 mm.

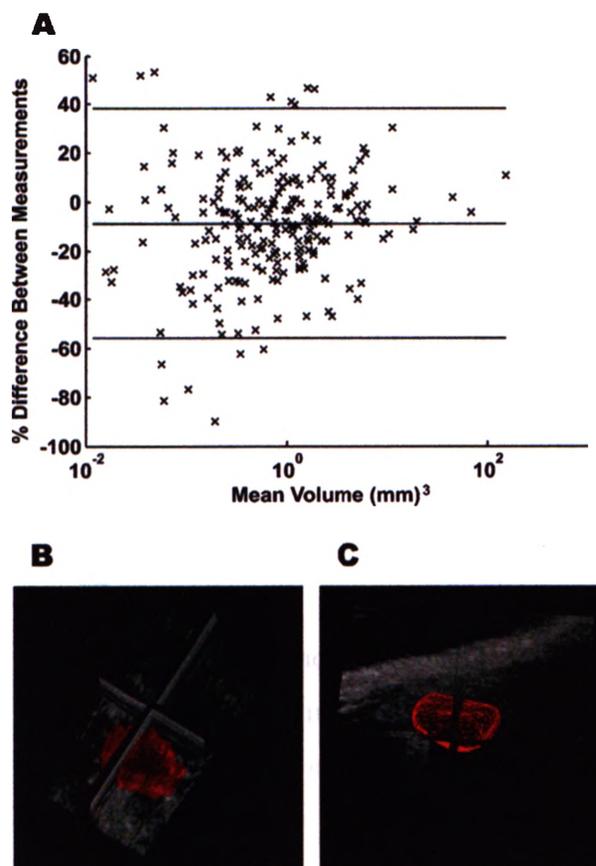


Fig. 2.3: Discrepancy between tumour volumes obtained from 3D segmentation or diameter measurement with the assumption of an ellipsoid shape. (A) The % difference between the ellipsoid volume and the 3D segmented volume (equation 2.1) for individual B16F1 metastases is plotted versus the mean volume of those two measurements. The solid bars indicate the mean \pm 2 SD. (B) 3D surface rendering of a B16F1 metastasis in which the ellipsoid and 3D segmented volumes are in close agreement (ellipsoid = 5.58 mm^3 , 3D = 5.79 mm^3 , % difference = -3.70%). (C) 3D surface rendering of a B16F1 metastasis in which the ellipsoid and 3D segmented volumes are not in close agreement (ellipsoid = 3.51 mm^3 , 3D = 5.03 mm^3 , % difference = -35.6%).

2.3.5 3D ultrasound can be used to monitor therapeutic response of individual metastases

To assess the ability of high-frequency ultrasound to evaluate the efficacy of cytotoxic chemotherapeutic agents, the B16F1 liver metastasis model was used. By noninvasively tracking the development of individual liver metastases it was shown that doxorubicin significantly increased the doubling time of B16F1 metastases from (mean \pm SD) 1.4 ± 0.4 days to 1.7 ± 0.4 days (t-test, $p = 0.038$) (Fig 2.4 A-B). The increased doubling time is apparent in the significant difference between the average tumour volumes of doxorubicin treated metastases and the control metastases (Fig 2.4 C).

2.4 Discussion

The importance of preclinical animal models in oncological research has supported the development of small animal imaging modalities including MRI, x-ray CT, PET, and fluorescent and bioluminescent based systems. Each of these modalities occupies a niche in noninvasive imaging, based on the unique requirements and constraints of a particular research study. Important factors in choosing the appropriate imaging modality for a particular study may include the anatomical site being imaged, the desired resolution and animal throughput, availability of targeted contrast agents and cost. In this report we describe the use of high-frequency ultrasound imaging for the detection and longitudinal tracking of murine liver metastases. High-frequency ultrasound offers distinct advantages as a cost effective, rapid, high-resolution and noninvasive imaging system. The ultrasound imaging described here was performed without exogenous contrast agents or genetic manipulation of the cell lines being studied. This offers significant advantages both in terms of animal throughput and in the number of animal models able to be studied.

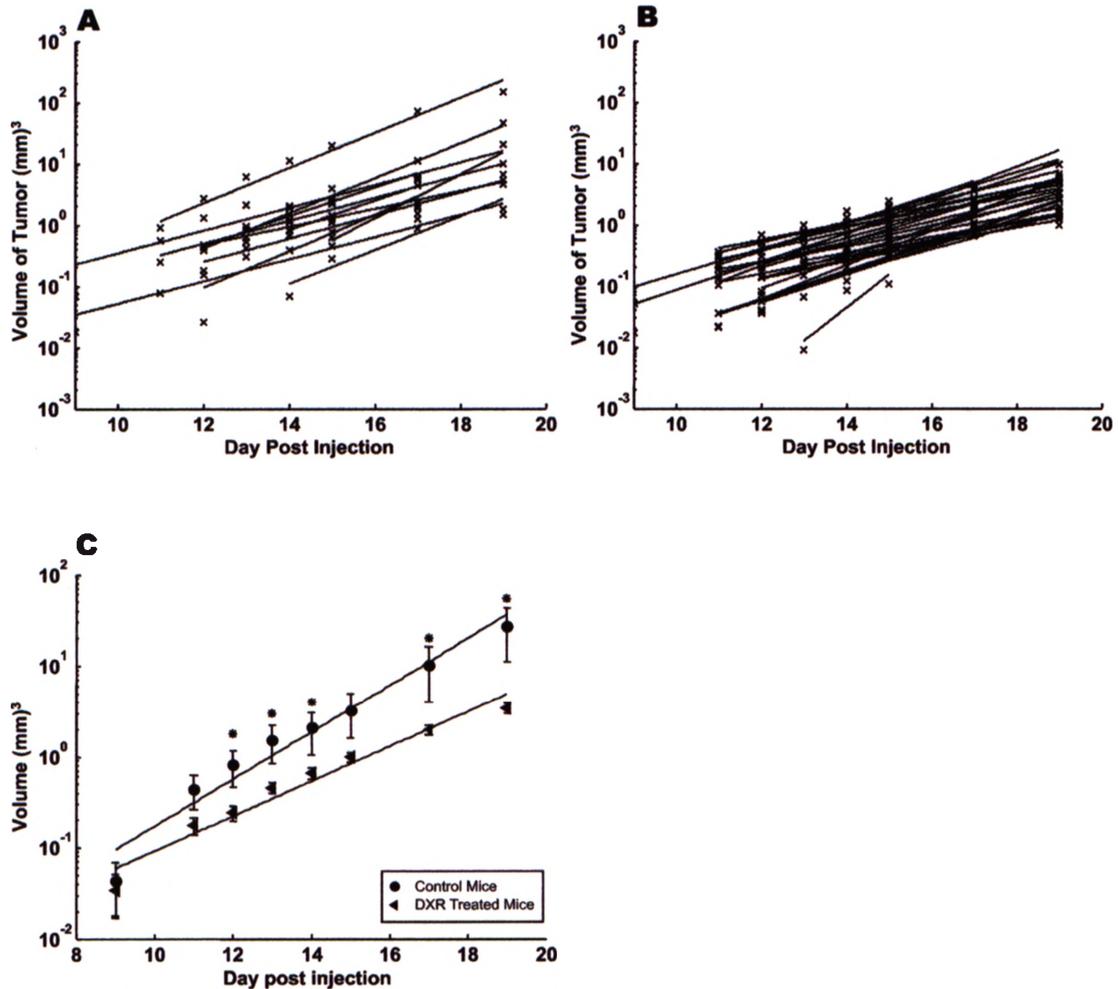


Fig. 2.4: A longitudinal ultrasound study shows that doxorubicin (DXR) decreases tumour growth rate and tumour volume in the B16F1 liver metastasis model. (A) The growth of 9 B16F1 liver metastases, from 5 mice, was tracked longitudinally in the control group. (B) The growth of 24 B16F1 liver metastases, from 8 mice, was tracked longitudinally in the doxorubicin treatment group. (C) The mean metastasis volume (mean \pm SE) for each imaging timepoint is shown for the control and doxorubicin treatment groups. The asterisk (*) denotes a significant difference in tumour volume (rank sum test, $p < 0.05$) at the indicated timepoint.

As demonstrated in this report, all four cell lines tested in experimental liver metastasis models, B16F1, HT-29, MDA-MB-435/HAL and PAP2, showed inherent ultrasound contrast relative to the surrounding liver parenchyma. These cell lines are derived from different primary tumour types, murine melanoma, human colon carcinoma, human breast carcinoma and oncogene-transformed murine fibroblasts, which demonstrates the wide applicability of this technique. In all cases the liver metastases were hypoechoic, appearing darker than the surrounding tissue on ultrasound images. Ultrasound imaging was shown to be highly sensitive, with metastases from all four tumour models detected with maximum diameters less than 0.78 mm. The consistent background image texture of normal liver parenchyma likely contributed to the detection of small tumours, so ultrasound appears particularly well-suited for imaging liver metastasis models. To determine the absolute detection limit for any particular cell line, more frequent imaging on a greater number of mice would need to be performed.

The ability to track the growth of individual liver metastases over time was demonstrated for B16F1, HT-29 and MDA-MB-435/HAL tumour cell lines. The use of the human cell lines HT-29 and MDA-MB-435/HAL is particularly noteworthy since immunodeficient animals, which must be protected from infection, are used for these metastasis models. The ultrasound system was easily adapted to this requirement by restricting the mice and ultrasound probe to a HEPA filtered environment. This process would not be possible with larger, less portable imaging modalities.

The longitudinal imaging trials demonstrated the importance of noninvasive imaging in allowing analysis on a per metastasis instead of a per mouse basis. For example, in contrast with the exponential growth seen with other liver metastases in the same animal, the metastasis HT-29 - 4 did not show a significant increase in

tumour volume during the 10 days it was imaged. The volume of this metastasis was constant at 0.03 mm^3 (0.3 mm diameter). The identification of metastases with variable growth patterns would allow further investigation, such as microdissection and microarray analysis, to elucidate the molecular basis of such variations. The monitoring of dormant metastases would also permit the study of host-tumour interactions, such as the role of angiogenesis in the switch from a dormant to progressive tumour phenotype [20]. The detection sensitivity of high-frequency ultrasound allows the investigation of processes, such as tumour dormancy and angiogenesis, which may occur very early in the metastatic process.

The longitudinal imaging of B16F1 liver metastases revealed striking changes in ultrasound image texture during tumour development. The most obvious of these changes was the development of distinct anechoic regions that were shown to be areas of liquefactive necrosis. The ability to detect the formation of necrosis may be useful in the assessment of vascular targeting agents and anti-angiogenic compounds [21, 22]. Incorporating Doppler blood flow imaging into studies utilizing the high-frequency ultrasound system may further enhance the assessment of tumour vasculature and hemodynamics [23]. The cellular or structural characteristics that cause the more subtle changes in ultrasound backscatter are currently under investigation in several laboratories [24-26].

These studies have focused on imaging the development of individual metastases because of the unique research opportunities that this approach presents. In time, the models used here will form large coalescing metastases, which can no longer be monitored for individual growth characteristics. At this stage, ultrasound imaging could be used to measure relative tumour burden between animals. This application will require further study to determine how tumour burden in acoustically accessible liver areas represent the state of the entire liver.

Since subcutaneous tumour growth is frequently monitored by caliper measurement and calculation of an ellipsoid volume, we sought to determine if the approximation of an ellipsoid volume was sufficient for monitoring the growth of liver metastases. It was demonstrated that tumour volumes calculated from 3D and 2D methods yielded vastly different results for many metastases. For B16F1 metastases the mean percent volume difference between the two methods was $-8.8 \pm 23.5\%$. The large standard deviation indicates that the 2D method often gives large over- or underestimations of tumour volume when compared to the 3D method. Since the true volume of the metastases could not be determined it cannot be definitely stated that one method is more accurate than the other. However, it is reasonable to suggest that the 3D method is more accurate since there is no assumption of a defined shape, and a 3D image allows the operator greater time and control when defining tumour borders. Definition of tumour borders can be done off-line with a 3D image, while the operator of a 2D system must identify maximum tumour diameters during imaging. Furthermore, previous work with a clinical ultrasound system has shown that the 3D method is more accurate than the 2D method when measuring the volume of regular and irregular shaped phantoms [18, 27]. The inaccuracy and variability brought about by assuming a defined shape could hinder the ability to track volume changes in slowly growing metastatic models, the ability to track subtle responses to therapeutic treatment, and the ability to determine if a metastasis is going through a period of dormancy. The elimination of this uncertainty presents a compelling case for using an imaging modality with 3D imaging capabilities.

The utility of longitudinal ultrasound imaging in preclinical trials was demonstrated with the anthracycline chemotherapeutic, doxorubicin. Longitudinal assessment of individual liver metastases showed that doxorubicin significantly decreased tumour growth rate and tumour volume in the B16F1 liver metastasis model. Signifi-

cant differences in tumour volume were evident at day 12 post cell injection, after only three doxorubicin treatments, when the average tumour volume in the control group was less than 1.00 mm^3 , and in the treated group less than 0.25 mm^3 . The ability of high-frequency ultrasound to track the progression of micrometastases noninvasively allows the evaluation of therapeutic efficacy on sequential stages of tumour development, from early formation to the development of large vascularized metastases, in a single experiment.

In summary, this report is the first to describe the use of 3D high-frequency (40 MHz) ultrasound imaging in the noninvasive detection and longitudinal evaluation of murine liver metastases. This development is significant in that ultrasound offers rapid, cost effective, high-resolution imaging that can be applied to a wide range of liver metastasis models without the requirement for contrast agents. Compared to traditional histological methods, ultrasound imaging may provide a more accurate assessment of tumour progression and chemotherapeutic response, while opening new avenues of investigation into dynamic processes such as tumour vascularization and tumour dormancy.

References

- [1] R. Weissleder, "Scaling down imaging: molecular mapping of cancer in mice," *Nat Rev Cancer*, vol. 2, no. 1, pp. 11–8, 2002.
- [2] G. Choy, P. Choyke, and S. K. Libutti, "Current advances in molecular imaging: noninvasive in vivo bioluminescent and fluorescent optical imaging in cancer research," *Mol Imaging*, vol. 2, no. 4, pp. 303–12, 2003.
- [3] R. Stripecke, M. Carmen Villacres, D. Skelton, N. Satake, S. Halene, and D. Kohn, "Immune response to green fluorescent protein: implications for gene therapy," *Gene Ther*, vol. 6, no. 7, pp. 1305–12, 1999.
- [4] F. Re, R. Srinivasan, T. Igarashi, F. Marincola, and R. Childs, "Green fluorescent protein expression in dendritic cells enhances their immunogenicity and elicits specific cytotoxic t-cell responses in humans," *Exp Hematol*, vol. 32, no. 2, pp. 210–7, 2004.
- [5] C. Khanna and K. Hunter, "Modeling metastasis *in vivo*," *Carcinogenesis*, vol. 26
- [6] D. H. Turnbull, J. A. Ramsay, G. S. Shivji, T. S. Bloomfield, L. From, D. N. Sauder, and F. S. Foster, "Ultrasound backscatter microscope analysis of mouse melanoma progression," *Ultrasound Med Biol*, vol. 22, no. 7, pp. 845–53, 1996.
- [7] A. Fenster, D. B. Downey, and H. N. Cardinal, "Three-dimensional ultrasound imaging," *Phys Med Biol*, vol. 46, no. 5, pp. R67–99, 2001.
- [8] H. Kobayashi, T. Saga, S. Kawamoto, N. Sato, A. Hiraga, T. Ishimori, J. Konishi, K. Togashi, and M. W. Brechbiel, "Dynamic micro-magnetic resonance imaging of liver micrometastasis in mice with a novel liver macromolecular magnetic resonance contrast agent dab-am64-(1b4m-gd)(64)," *Cancer Res*, vol. 61, no. 13, pp. 4966–70, 2001.

- [9] M. Yang, E. Baranov, P. Jiang, F. X. Sun, X. M. Li, L. Li, S. Hasegawa, M. Bouvet, M. Al-Tuwaijri, T. Chishima, H. Shimada, A. R. Moossa, S. Penman, and R. M. Hoffman, "Whole-body optical imaging of green fluorescent protein-expressing tumors and metastases," *Proc Natl Acad Sci U S A*, vol. 97, no. 3, pp. 1206-11, 2000.
- [10] N. Smakman, A. Martens, O. Kranenburg, and I. H. Borel Rinkes, "Validation of bioluminescence imaging of colorectal liver metastases in the mouse," *J Surg Res*, vol. 122, no. 2, pp. 225-30, 2004.
- [11] S. M. Weber, K. A. Peterson, B. Durkee, C. Qi, M. Longino, T. Warner, Jr. Lee, F. T., and J. P. Weichert, "Imaging of murine liver tumor using microCT with a hepatocyte-selective contrast agent: accuracy is dependent on adequate contrast enhancement," *J Surg Res*, vol. 119, no. 1, pp. 41-5, 2004.
- [12] I. J. Fidler, "Selection of successive tumour lines for metastasis," *Nat New Biol*, vol. 242, no. 118, pp. 148-9, 1973.
- [13] S. A. Hill, S. Wilson, and A. F. Chambers, "Clonal heterogeneity, experimental metastatic ability, and p21 expression in h-ras-transformed nih 3t3 cells," *J Natl Cancer Inst*, vol. 80, no. 7, pp. 484-90, 1988.
- [14] H. J. Varghese, M. T. Davidson, I. C. MacDonald, S. M. Wilson, K. V. Nadkarni, A. C. Groom, and A. F. Chambers, "Activated ras regulates the proliferation/apoptosis balance and early survival of developing micrometastases," *Cancer Res*, vol. 62, no. 3, pp. 887-91, 2002.
- [15] C. M. Schmidt, S. L. Settle, J. L. Keene, W. F. Westlin, G. A. Nickols, and D. W. Griggs, "Characterization of spontaneous metastasis in an aggressive breast carcinoma model using flow cytometry," *Clin Exp Metastasis*, vol. 17, no. 6, pp. 537-44, 1999.

- [16] V. L. Morris, I. C. MacDonald, S. Koop, E. E. Schmidt, A. F. Chambers, and A. C. Groom, "Early interactions of cancer cells with the microvasculature in mouse liver and muscle during hematogenous metastasis: videomicroscopic analysis," *Clin Exp Metastasis*, vol. 11, no. 5, pp. 377-90, 1993.
- [17] F. S. Foster, M. Y. Zhang, Y. Q. Zhou, G. Liu, J. Mehi, E. Cherin, K. A. Harasiewicz, B. G. Starkoski, L. Zan, D. A. Knapik, and S. L. Adamson, "A new ultrasound instrument for in vivo microimaging of mice," *Ultrasound Med Biol*, vol. 28, no. 9, pp. 1165-72, 2002.
- [18] Tong Shidong, H. Neale Cardinal, F. McLoughlin Raymond, B. Downey Donal, and Fenster Aaron, "Intra- and inter-observer variability and reliability of prostate volume measurement via two-dimensional and three-dimensional ultrasound imaging," *Ultrasound Med. Biol.*, vol. 24 1998.
- [19] G. N. Naumov, J. L. Townson, I. C. MacDonald, S. M. Wilson, V. H. Bramwell, A. C. Groom, and A. F. Chambers, "Ineffectiveness of doxorubicin treatment on solitary dormant mammary carcinoma cells or late-developing metastases," *Breast Cancer Res Treat*, vol. 82, no. 3, pp. 199-206, 2003.
- [20] T. Udagawa, A. Fernandez, E. G. Achilles, J. Folkman, and R. J. D'Amato, "Persistence of microscopic human cancers in mice: alterations in the angiogenic balance accompanies loss of tumor dormancy," *Faseb J*, vol. 16, no. 11, pp. 1361-70, 2002.
- [21] H. Varghese, L. Mackenzie, A. Groom, C. Ellis, A. Ryan, I. Macdonald, and A. Chambers, "In vivo videomicroscopy reveals differential effects of the vascular-targeting," *Angiogenesis*, vol. 7, no. 2, pp. 157-64, 2004.
- [22] P. E. Thorpe, "Vascular targeting agents as cancer therapeutics," *Clin Cancer Res*, vol. 10, no. 2, pp. 415-27, 2004.

- [23] D. E. Goertz, J. L. Yu, R. S. Kerbel, P. N. Burns, and F. S. Foster, "High-frequency doppler ultrasound monitors the effects of antivasular therapy on tumor blood flow," *Cancer Res*, vol. 62, no. 22, pp. 6371-5, 2002.
- [24] G. J. Czarnota, M. C. Kolios, J. Abraham, M. Portnoy, F. P. Ottensmeyer, J. W. Hunt, and M. D. Sherar, "Ultrasound imaging of apoptosis: high-resolution non-invasive monitoring of programmed cell death in vitro, in situ and in vivo," *Br J Cancer*, vol. 81, no. 3, pp. 520-7, 1999.
- [25] J. W. Hunt, A. E. Worthington, A. Xuan, M. C. Kolios, G. J. Czarnota, and M. D. Sherar, "A model based upon pseudo regular spacing of cells combined with the randomisation of the nuclei can explain the significant changes in high-frequency ultrasound signals during apoptosis," *Ultrasound Med Biol*, vol. 28, no. 2, pp. 217-26, 2002.
- [26] M. L. Oelze, Jr. O'Brien, W. D., J. P. Blue, and J. F. Zachary, "Differentiation and characterization of rat mammary fibroadenomas and 4t1 mouse carcinomas using quantitative ultrasound imaging," *IEEE Trans Med Imaging*, vol. 23, no. 6, pp. 764-71, 2004.
- [27] H. X. Xu, X. Y. Yin, M. D. Lu, G. J. Liu, and Z. F. Xu, "Estimation of liver tumor volume using a three-dimensional ultrasound volumetric system," *Ultrasound Med Biol*, vol. 29, no. 6, pp. 839-46, 2003.

Chapter 3

Volume Measurement Variability In Three-Dimensional High-Frequency Ultrasound Images Of Murine Liver Metastases

The content of this chapter has been adapted from: "Volume measurement variability in three-dimensional high-frequency ultrasound images of murine liver metastases", published in Physics in Medicine and Biology, vol 51 (10) May 21 2006, by Lauren A. Wirtzfeld, Kevin C. Graham, Alan C. Groom, Ian C. Macdonald, Ann F. Chambers, Aaron Fenster and James C. Lucefield.

3.1 Introduction

Pre-clinical mouse models of cancer have proven valuable for research and development of new cancer therapies. They serve as an intermediate step between *in vitro* cell culture experiments and clinical trials in humans [1], for example, allowing for the evaluation of different stages of cancer including primary or metastatic tumours.

The mouse models provide a 3D geometry and a realistic micro-environment in which to study tumour growth and treatment response before translation to humans.

Historically, most small animal cancer research has relied on subcutaneous tumour models. These models allow for external measurements of tumour diameter using calipers to determine tumour growth and treatment response. However, caliper measurements of tumour diameter may not be an accurate indicator of tumour growth because the skin and inflammation are included in the measurements [2]. Furthermore, there is a growing realisation that orthotopic, metastatic, and spontaneously developing tumours in transgenic models may more closely approximate human disease [3]. For these models, the development of tumours in internal organs precludes the use of caliper measurement and often necessitates the use of endpoint analysis of tumour burden and structure to evaluate tumour development. This approach is costly and is susceptible to the inherent variability in the rate of tumour development in animal models. For endpoint analysis, the researcher is unaware of the current stage of tumour development when initiating a therapeutic regimen or when considering sacrificing the animal for further analysis. The development of *in vivo* imaging for pre-clinical models has allowed for non-invasive longitudinal study of tumour progression and assessment of tumour response to therapy.

High-frequency ultrasound is an attractive option for non-invasive longitudinal imaging of pre-clinical models [4, 5]. In particular, high-frequency ultrasound can be used to detect and longitudinally evaluate tumour progression in pre-clinical soft-tissue tumour models, including liver metastases [6], transgenic prostate tumours [7] and intra-dermal tumours [8, 9]. Ultrasound is also useful for other tumour models, as demonstrated by results obtained at lower frequencies in orthotopic bladder tumours [10], spontaneous mammary tumours [11] and orthotopic hepatocellular carcinomas [12].

While ultrasound does not pose a risk to the animal with repeated imaging, every time point imaged requires an investment of machine and technician time and a

risk to the animal from anaesthesia, which should be minimized. These factors add to the cost of the study. Repeated imaging before a significant change in volume occurs provides no additional information about tumour growth; therefore, imaging time points should be carefully chosen to maximize the utility of the data. Tumour volume measurement variability, which determines the minimum observable change in size, must be known to plan these longitudinal experiments. Thus, it is critical to both quantify the magnitude and identify sources of measurement variability, in order to reduce their impact during image acquisition and analysis.

This study focuses on murine melanoma liver metastases as an example model because tumours growing in internal organs, such as the liver, cannot be evaluated longitudinally without the use of non-invasive imaging. Experimental metastasis models, in which cells are injected directly into the circulation, enable the investigator to control the numbers of potential tumours in an organ and the time frame in which they will develop [3]. In this paper, sources of variability in the volume measurements for this model are evaluated and quantified. First, intra-observer variability, *i.e.*, the variability of repeated volume measurements of a single tumour performed by the same observer, and inter-observer variability, *i.e.*, the variability of repeated measurements by different observers, are analyzed as functions of tumour volume. Second, since the segmentation is performed in parallel planes through the 3D image, the distance between the segmented planes will affect the measurement variability and estimated volume. Therefore, we determined the most appropriate spacing between segmentation planes to minimize the time required to segment a tumour without changing the volume estimate by more than the intra-observer variability. Finally, the impact on the variability of experimental factors including breathing motion, the depth of field of the image, and the repeatability of imaging a specific tumour is also assessed. These observations enable minimization of the measurement variability from sources that can be controlled. The procedures described here for measuring variability will be applicable to other pre-clinical cancer models.

3.2 Methods

3.2.1 Animal Model

An experimental liver metastasis model, previously described in [13] and [6], was used. B16F1 murine melanoma cells [14] (Cat# CRL-6323; American Type Culture Collection, Manassas, VA, USA) were maintained in α +MEM (Invitrogen, Carlsbad, CA, USA) supplemented with 10% FBS (Sigma, Mississauga, ON, Canada). All animals were cared for in accordance to the Canadian Council on Animal Care under a protocol approved by the University of Western Ontario Council on Animal Care. To produce experimental liver metastases, mice were anaesthetized with an intraperitoneal injection of xylazine/ketamine (2.6 mg ketamine and 0.13 mg xylazine per 20 g body mass). As described previously [13], anaesthetized female C57BL/6 mice (Harlan, Indianapolis, IN, USA) received mesenteric vein injections of 3×10^5 cells suspended in 0.1 mL of growth media to target the liver. All mice were 7 to 11 weeks old at the time of cell injection.

Prior to the first imaging session, the mouse's abdomen was depilated with commercial hair removal cream. During imaging, the mouse was kept under anaesthesia with 1.5% isoflurane in oxygen and restrained on a heated stage. For longitudinal experiments, animals were imaged every two to three days.

3.2.2 Image Acquisition

All images were acquired with a Vevo 660 (VisualSonics, Toronto, Canada) high-frequency ultrasound system. A 40 MHz centre frequency transducer with a mechanically scanned, single-element aperture was employed. At the 6-mm focal distance, the resolution was $40 \times 80 \times 80 \mu\text{m}^3$ with a 1.5 mm depth of field (manufacturer's specifications). Two-dimensional (2D) images were acquired with a $8 \times 8 \text{ mm}^2$ field of view at 30 frames per second. Three-dimensional (3D) images were produced by

translating the transducer in the elevation dimension and acquiring 2D images every $30\ \mu\text{m}$ using the algorithm described in [15]. Three-dimensional images of $8 \times 8 \times 6\ \text{mm}^3$ regions of interest were acquired in approximately 20 seconds.

3.2.3 Tumour Volume Measurement

Tumour volumes were measured by manual segmentation of parallel planes through the 3D images using software developed in our laboratory [16–18]. The observer was able to rotate the 3D image and cut through the volume in any plane to familiarize themselves with the 3D image. Once familiar with the image, the observer began manual segmentation on an initial plane, typically near the middle of the tumour, that was pre-selected by one of the investigators (L.A.W.). After the observer outlined the tumour in the visible plane by placing points around the tumour edge, the segmentation plane was moved by a $50\ \mu\text{m}$ step and the tumour was outlined on the new plane. Once the edge of the tumour was reached in one direction, the observer returned to the centre of the image and segmented the other half of the tumour. Any observed errors could be corrected by moving, adding or deleting selected points on the outline. Figure 3.1 shows an example of the stages in segmentation of a 4.7-mm^3 tumour. Once the tumour was completely segmented, the volume of the tumour was automatically calculated using a method analogous to the familiar trapezoidal rule for 1D integration, where the volume between two slices is calculated as the average of the two outlined areas multiplied by the inter-slice distance, and the inter-slice volumes summed to obtain the total tumour volume [16]. All observers' tumour outlines were saved for later analysis.

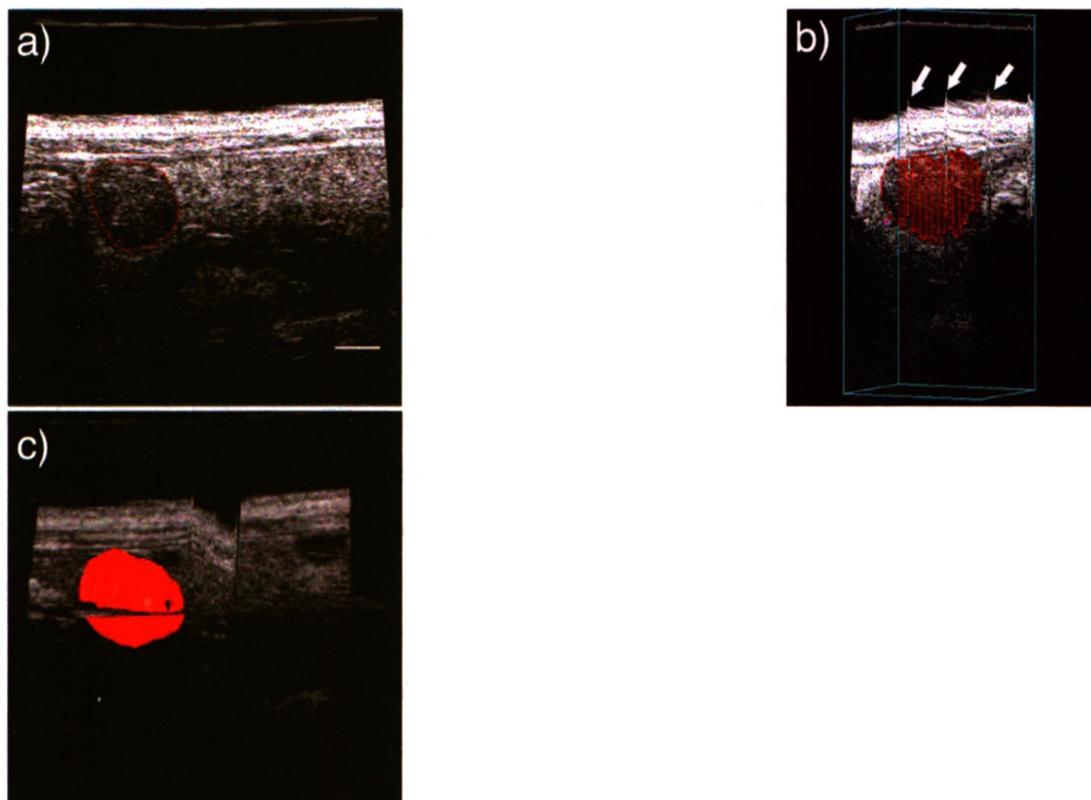


Fig. 3.1: Stages in segmentation of a 4.7-mm³ tumour. a) The red outline of the tumour shows the manual segmentation of a 2D plane from a 3D image. b) A 3D image showing the locations of segmented planes in red. The arrows indicate breathing artifacts. c) A surface rendered view of the segmented tumour with perpendicular planes through the image volume. Scale bar in (a) is 1 mm.

3.2.4 Analysis of Volume Measurement Variability

3.2.4.1 Inter- and Intra-Observer Variability

A multi-observer study was performed to assess the inter- and intra-observer variability of tumour volume measurement. The desired number of repeat measurements per image (n) was determined by a conventional sample size calculation [19] with significance level 0.05 and power 0.80. The calculation used an estimate of the standard deviation of repeated volume measurements ($\sigma = 0.04 \text{ mm}^3$) that was obtained from a previous study [20] with a single observer and tumours with a volume less than 2 mm^3 . The calculation indicated that 16 repeated measurements per image were necessary to obtain an observable difference in volume of 0.04 mm^3 , which was slightly smaller than the smallest tumour in the multi-observer study. A minimum of three repeated measurements per person enabled calculation of statistics for each observer, so three repeated measurements by seven observers ($n = 3 \times 7 = 21$) were chosen to ensure sufficient significance and power if the observer standard deviations were higher than estimated.

Seven trained observers segmented thirteen 3D ultrasound images of B16F1 liver metastases from a longitudinal study [6]. As outlined in table 3.1, 13 images were included in the study, comprised of 10 unique tumours, and 3 cases where images of the same tumour, but at different time points over the longitudinal study, were included. Each observer segmented each image three times, at one-week intervals, to minimize the effects of the observers remembering images. Images were randomized for each repeated segmentation and numbered, then the observers performed the segmentation on the images in sequential order. Two viewing locations were used for image segmentation. All other environmental conditions remained the same between all observers and images. The computer, monitor (including the settings) and lighting conditions were kept constant for all images and observers. The image magnification, image contrast and brightness settings, and starting plane for segmentation were kept

constant for each image. Predetermined start planes were used for all observers.

Table 3.1: Tumour images analyzed. The tumour volume listed is the average volume over all 21 measurements from the multi-observer study. The mean volume column is the average volume of the tumours included in the tumour volume range bin. The day column lists the numbers of days after injections of the B16F1 cells that the image was acquired. The same tumours are noted with *,** and ***

Bin	Mean Volume (mm ³)	Day	Volume (mm ³)
A	0.43	10	0.07
		12	0.45
		18	0.76
B	2.39	15	1.15***
		15	1.58
		18	2.85***
		18	3.98
C	06.26	17	4.71*
		17	04.90**
		18	7.46**
		18	7.98*
D	60.42	24	47.68
		22	73.16

The standard deviations of the repeated volume measurements were calculated for each tumour and for each observer. The pooled measurements of all the observers were used to determine the relationship between the standard deviation and volume by linear regression of log-scaled data (GraphPad Prism version 4.00 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com).

Intra- and inter-observer reliability coefficients were calculated using the method described in [21]. The reliability coefficients range from 0 to 1 and indicate the agreement within or between groups. Based on a preliminary segmentation, the tumours were grouped by size as shown in table 3.1 and a two-way analysis of variance (ANOVA) was performed on each subset of the data. The inter- and intra-observer standard error of measurements (SEM_{inter} and SEM_{intra} respectively) were calculated using a two-way ANOVA following the procedure described by [21] and [22].

The inter- and intra-observer minimum detectable change in volume (Δv_{min}) was calculated from the SEM using equation 3.1 [21],

$$\Delta v_{min} \geq z_{\alpha} \cdot \sqrt{2} \cdot SEM, \quad (3.1)$$

where a significance level of $\alpha = 0.05$ was used, giving $z_{\alpha} = 1.96$.

3.2.4.2 Variability Due to Segmentation Inter-Slice Distance

The tumour outlines from the three segmentations of each tumour from one observer were used to determine the effect on volume estimation of inter-slice distance (ISD), *i.e.*, the spacing between parallel segmentation planes. Since manual segmentation was performed with a 50 μm ISD, a volume could be calculated for an ISD equal to any integer multiple of 50 μm by applying the trapezoid rule to outlined contours from a subset of the segmented planes. As the start plane was constrained and kept the same for each repeated tumour measurement, the start plane was also kept constant when analyzing the effect of ISD on volume estimation.

The coefficient of variation (COV, (SEM/mean volume)) for the repeated segmentations at each ISD for each tumour was calculated and used in a one-way ANOVA. ISD over the range from 50 to 450 μm were used as this was the range where volumes for all tumours could be calculated. If there was a significant result from the ANOVA ($p < 0.05$), Dunnett post-hoc tests were performed. The Dunnett tests compared all data sets to the 50 μm data as a control to determine if there were any group differences in the COV obtained at different ISD.

For each segmentation of each tumour, the volumes obtained at all ISD were normalized by dividing by the volumes obtained at 50 μm . The average and standard deviation of the normalized volumes for all tumours in each bin were then calculated. For each bin, the normalized volume and standard deviation were plotted against the ISD. In order to keep the effects of ISD lower than the intra-observer measurement variability, the largest ISD where the mean plus or minus the standard deviation did

not extend beyond the mean plus or minus the intra-observer SEM for that volume bin was chosen as the preferred ISD.

3.2.4.3 Variability Due to Experimental Factors

A separate cohort of mice was used to evaluate imaging repeatability and the effects of breathing motion and depth of field on volume estimation. Three mice with a total of 13 isolated metastases were imaged on day 15 after injection of B16F1 cells. To determine the scan-rescan repeatability, each tumour was imaged three times without moving the mouse and a fourth time after removing and replacing the mouse on the imaging stage.

Due to the limited (1.5 mm) ultrasound depth of field, the tumour can extend outside of the focal zone. The variability in the volume estimate due to the depth of field was evaluated by imaging each tumour after it had been raised or lowered 1 mm within the image, resulting in images with the centre of the tumour above and below the ultrasound focal zone.

Breathing artifacts result from motion between the acquisition of parallel planes in the 3D image. Four images of each tumour were acquired asynchronously with the respiratory cycle. To obtain an image with no breathing artifacts, each mouse underwent a tracheotomy. Mice were ventilated (Harvard Model 687 Small Animal Ventilator, Harvard Apparatus, MA) at 119 breaths per minute with 0.65 mL tidal volume. An image was acquired with ventilated breathing and a second image with the breathing halted for the duration of the image acquisition.

All the images were randomized and segmented by a single blinded observer as described above. This analysis yielded eight sets of volume data obtained under different experimental conditions. Pearson's correlation coefficients were calculated to determine if the volume measurements from the different imaging conditions were correlated. An ANOVA was performed to determine if there were any differences in volume means between the different imaging conditions. If the ANOVA identified

a difference in means at the $\alpha = 0.05$ significance level, post-hoc paired t-test tests with a Bonferroni correction were performed. As the Bonferroni correction was used to avoid *Type I* errors and there were a total of 28 pairs, $p < 0.05/28 = 0.002$ was considered statistically significant.

3.3 Results

3.3.1 Inter- and Intra-Observer Variability

The standard deviation (SD) of each observer's measurement of each tumour versus volume (V) is shown in figure 3.2. Linear regression was performed on log scaled data, which gave the following relationship:

$$\log_{10}SD = 0.925\log_{10}V - 1.40. \quad (3.2)$$

The range of tumour volumes were not normally distributed; however, no transformation was applied as volumes were analyzed in small number for each volume bin which did not allow for evaluation of the distribution. The reliability coefficients for each bin of tumour sizes are shown in table 3.2. High reliability coefficients are seen for all bins, except bin D (table 3.1) for the inter-observer variability, which was slightly less than 0.80. The high value for the coefficients indicates good agreement between repeated measurements by one observer and slightly lower reliability between different observers.

The intra-observer standard errors of measurement (SEM_{intra}) for the four bins were lower than the inter-observer (SEM_{inter}) and are summarized in table 3.2. The coefficients of variation (COV, *ie.*, SEM divided by the average tumour volume) for each volume bin are plotted in figure 3.3. The lowest COV values are seen for the medium volume range of tumours (bins B and C) and is as low as 4% for the intra-observer variability in the 1 to 4 mm³ bin. The largest (bin D) and smallest (bin

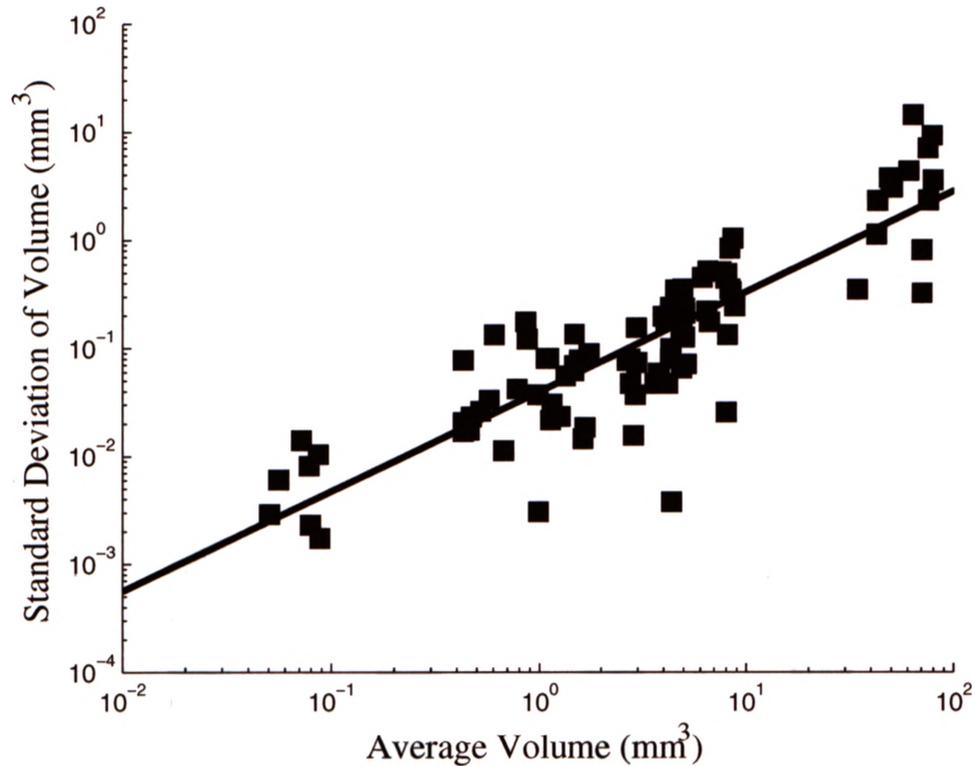


Fig. 3.2: Standard deviation of three measurements made by an observer for each tumour plotted against the mean volume of the three measurements. Data from all seven observers are presented. The solid line indicates the linear regression of the logged data. Data are plotted on a log-log scale.

Table 3.2: Inter- and intra-observer reliability coefficients, standard errors of measurement (SEM) and minimum detectable change in volume (Δv_{min}) for each volume bin.

Volume Bin (mean volume (mm ³))	Reliability Coefficient		SEM (mm ³)		Δv_{min} (mm ³)	
	Intra-	Inter-	Intra-	Inter-	Intra-	Inter-
A (0.43)	0.97	0.92	0.06	0.10	0.17	0.29
B (2.39)	0.99	0.98	0.09	0.18	0.26	0.49
C (6.26)	0.96	0.85	0.37	0.72	1.03	2.00
D (60.42)	0.93	0.79	5.46	9.13	15.16	25.32
All	0.99	0.97				

A) tumours have the highest COV at 9% and 14% respectively. The highest COV, approximately 24%, was obtained for the inter-observer results in the under 1 mm³ tumour bin.

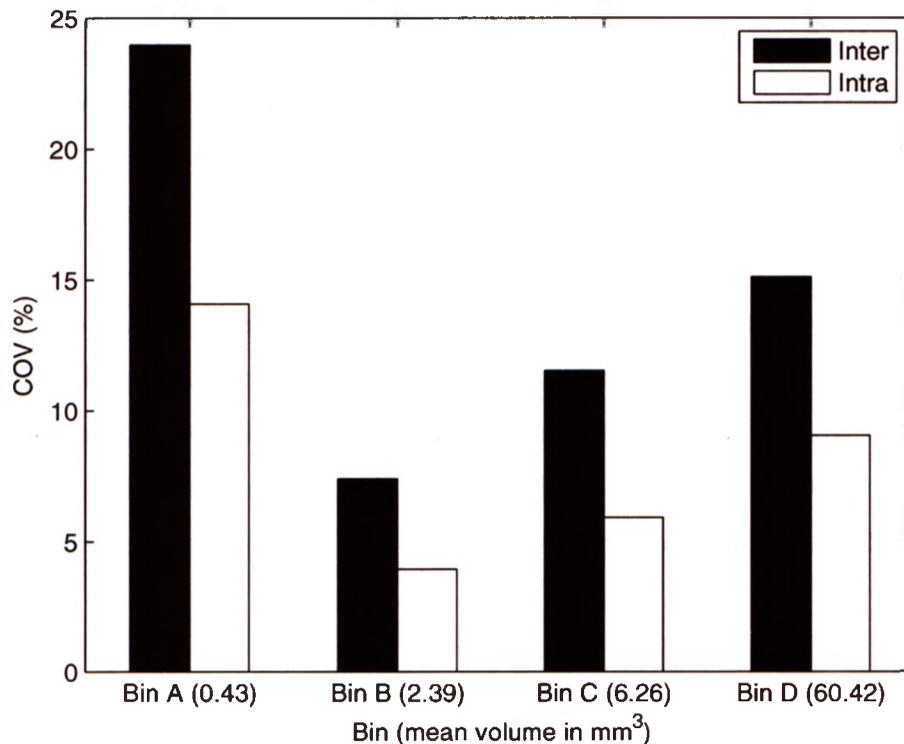


Fig. 3.3: The coefficient of variation (COV: standard error of measurement divided by the average volume in each bin) is plotted for the intra- and inter-observer variability over four different size ranges of tumours.

Minimum detectable changes in volume were calculated from the SEM using equation 3.1 and are shown in table 3.2. The inter-observer Δv_{min} is larger than the intra-observer value due to the higher SEM for all size bins. In all cases, the Δv_{min} is smaller than the average tumour volume being assessed. For the smallest tumours (bin A), Δv_{min} is approximately 40% of the average tumour volume, whereas for larger tumours, Δv_{min} is a smaller percentage of the tumour volume. This arises because small variations in measurement for small tumours can correspond to a large

percentage change in volume.

3.3.2 Variability due to Segmentation Inter-Slice Distance

The one-way ANOVA of the COV at different ISD was significant ($p = 0.001$). Dunnett's post-hoc tests showed significant differences between COV at ISD of 50 μm and the COV at all ISD greater than or equal to 300 μm ($p < 0.05$).

Figure 3.4 shows the volumes obtained with different slice thicknesses normalized by the volume determined with a 50- μm inter-slice distance for each tumour volume bin. As the ISD is increased, the volume calculated begins to deviate from the estimate obtained with the smallest ISD. The largest ISD that yielded a mean \pm standard deviation volume within the $\text{SEM}_{\text{intra}}$ bounds were 100, 150, 200 and 600 μm for bin A, B, C and D respectively. The maximum desired ISD monotonically increases with increasing tumour size. As the tumours in bin D are substantially larger than the rest, there is also a considerable jump in maximum ISD up to 600 μm .

3.3.3 Variability due to Experimental Factors

The imaging repeatability and the effects of respiratory motion and ultrasound depth of field on volume measurement were evaluated using a second set of tumours that ranged in size from 0.38 mm^3 to 7.0 mm^3 . Since the distributions of tumour volumes analyzed were not normally distributed, the volumes were log transformed to give normal distributions as judged by the Shapiro-Wilk test for normality ($p > 0.29$ for all data sets). The transformed data were used for all statistical analyses. A significant result ($p < 0.01$) was obtained from the one-way ANOVA and therefore the post-hoc tests were performed. The results of the correlation and significance of the t-tests are summarized in table 3.3 and show a comparison of the volumes of the 12 tumours measured in focus with halted breathing compared to each of the other

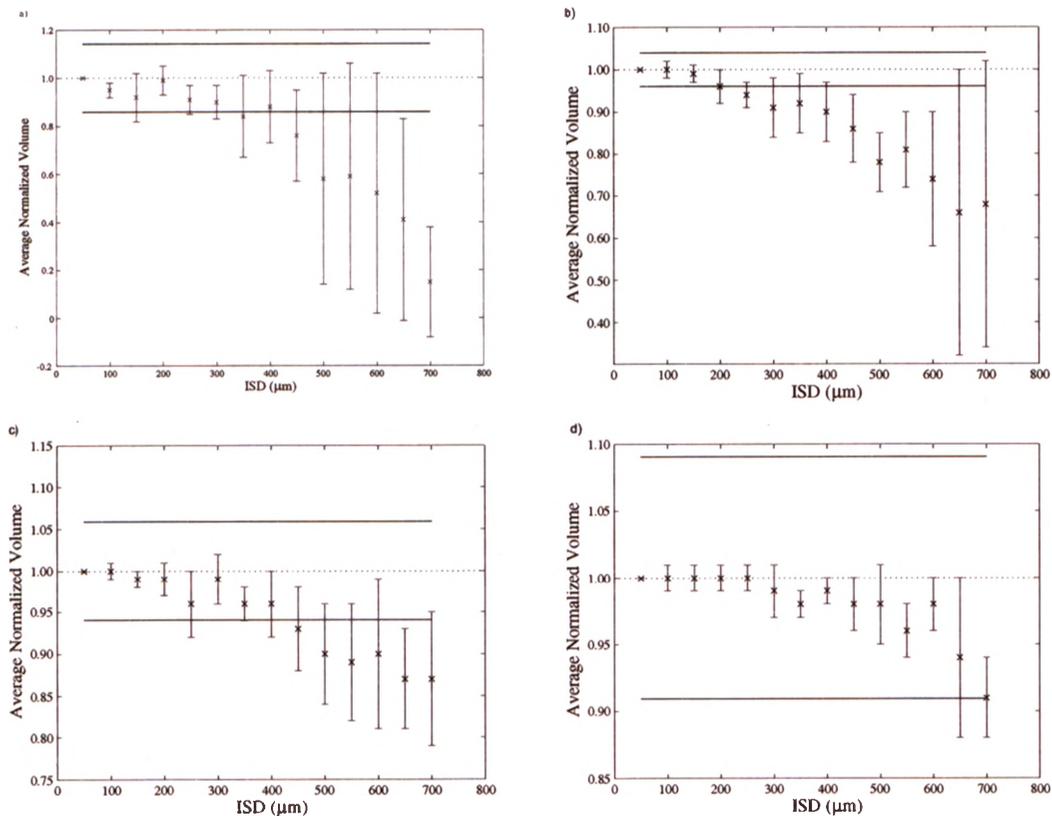


Fig. 3.4: Average normalized volume versus inter-slice distance (ISD) for tumour sizes of (a) less than 1 mm^3 , (b) 1 to 4 mm^3 , (c) 4 to 8 mm^3 and (d) greater than 8 mm^3 , up to 70 mm^3 . The solid lines indicate the relative intra-observer standard error of measurement (SEM_{intra}) for each of the volume ranges and the dotted line indicates a relative volume of 1. Data points are the mean volume for three measurements \pm the standard deviation.

Table 3.3: Comparison of image volumes obtained from 12 images acquired repeatedly while animal was breathing freely, while ventilated and with halted breathing and also for images with the tumour above and below the focus and after moving the animal. The correlation coefficients and t-tests are performed against the ideal case with the tumour at the focus and the breathing halted.

	Correlation Coefficient (ρ)	Paired t-test p -value
Free breathing 1	0.986	0.191
Free breathing 2	0.988	0.055
Free breathing 3	0.990	0.149
Re-scan	0.975	0.247
Tumour above focus	0.976	0.103
Tumour below focus	0.980	0.001
Ventilated breathing	0.996	0.325

conditions. Mean volumes were not statistically different from the in-focus halted-breathing data with the exception of the data from free-breathing mice imaged with the tumour below the focus. The halted breathing case gave the highest average volume and the tumour below the focus gave the lowest average volume. All the correlation coefficients were 0.975 or higher. All correlations were significant with p less than 0.0001.

3.4 Discussion

3.4.1 Inter- and Intra-Observer Variability

The trends shown in the intra- and inter- observer variability follow those seen in clinical studies as demonstrated by Xiong *et al.* [23] in liver tumours, in that the inter-observer variability is higher than the intra-observer variability. One difference between clinical ultrasound data and the high-frequency system is the relationship between the variability and the volume. It has been shown using data at clinical ultrasound frequencies [24] that, as the volume increases, the coefficient of variation (COV) decreases monotonically. For volumes measured in high-frequency images, the COV decreases with increasing volume for small tumours, but increases as a function

of volume for larger tumours. One potential reason for this increase at larger volumes is that these tumours now extend outside of the depth of field of the fixed-focus imaging system. Therefore, much of the boundary of the tumour is no longer in focus, leading to difficulty in identifying the location of the boundary. This effect can be seen in figure 3.5.

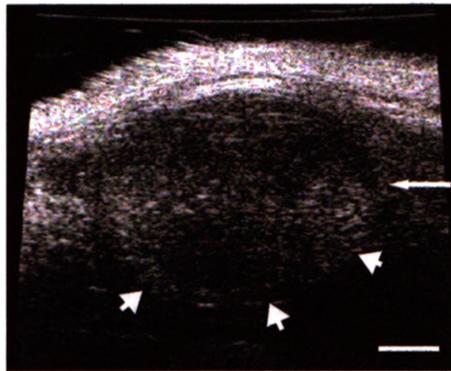


Fig. 3.5: 47.7 mm³ tumour. Image shows depth of focus artefact (narrow arrow), out of focus boundary (wide arrows) and rapid attenuation beyond the focus. Scale bar indicates 1 mm.

The intra-observer values for the SEM, COV, and Δv_{min} are all lower than the inter-observer values. This is the same trend that has been reported in clinical imaging studies, as repeated measurements by the same person are more likely to agree than measurements by different individuals. The highest COV is just under 24% for the inter-observer variability and just under 14% for the intra-observer variability for the smallest tumours. Since the typical volume doubling time for B16F1 liver metastasis is 1.2 days, [6] the intra-observer and inter-observer COV are both sufficiently low to permit daily tumour volume measurements in this model.

The measurement variability is most important for longitudinal imaging studies as it will dictate the detectable change in volume as the tumours grow. The high inter-observer reliability coefficients indicate that different observers could be used during a longitudinal study. To be conservative, if the same observer always segmented the

same tumour at each imaging time point, any bias introduced by the observer would apply to the entire growth curve. This would allow the segmentation to be shared between several observers.

In addition, the use of semi-automated segmentation techniques would significantly reduce the time required to measure tumour volumes. The data presented on the variability in manual segmentation could be used for the comparison of any semi-automated segmentation results to ensure the volume variability is the same or improved with the semi-automated technique.

3.4.2 Variability due to Segmentation Inter-Slice Distance

Significant differences in the COV are seen as the ISD increases, indicating that the variability in the volume estimate changes with ISD. The suggested maximum ISD increases monotonically with increasing volume. At the recommended ISD, the number of slices through the average tumour in each size range varied from 7.5 to 12. Using the maximum acceptable ISD will reduce the total number of planes segmented for each tumour and therefore reduce the amount of time required to analyze the data. For larger experiments where many tumours are followed, the time spent segmenting the tumours can be prohibitive if every acquired image plane is analyzed.

In this study, the evaluation of ISD was constrained solely to varying the distances between planes but not the initial plane. In our segmentation method, the volume outside the last segmented plane at each end of the tumour is omitted from the volume computation. With the start plane fixed, differences in the volume omitted at the ends of the tumour are the primary source of discrepancies between volumes estimated at different ISD. If the start plane is allowed to vary, the number of planes segmented can also vary, which could introduce an additional source of variability in the volume estimate.

3.4.3 Variability due to Experimental Factors

Our results demonstrate that breathing motion does not substantially change the measured tumour volume. The breathing rate of an anaesthetized mouse is approximately one to three breaths per second. The ultrasound system acquires a single 2D image plane in 1/30 second, so respiratory motion is not apparent in the 2D images. During 3D image acquisition, ten 2D planes are acquired in one second, which permits acquisition of several 2D planes per respiratory cycle. While the breathing cycle is periodic, there is a plateau at end expiration with minimal organ motion. As a result, few 2D planes show the active part of the respiratory cycle. This effect can be seen in figure 3.1b.

Since the estimated tumour volumes were smaller when the images were acquired when the tumour was located in the far field, below the focus, it is advisable to avoid having any data of interest in this region. Significant differences in tumour volume were not observed when the tumour was located in the near field, above the focus, suggesting that larger lesions that do not fit within the depth of field should preferably extend above the focal zone rather than below it. The tight focus of the single-element probe and high attenuation at 40 MHz results in rapid signal loss beyond the focal depth and consequent difficulties in segmenting these tumours.

3.4.4 Design of Longitudinal Imaging Studies

If the equation to describe tumour growth, the growth rate, and the variability of volume measurement are all known, then it is possible to compute the time needed to observe a significant change in volume. An example is presented here for the specific case of exponential growth in volume, which we have shown to be a good approximation to the growth curves over the range of volumes studied [6]. Exponential growth is a common model used in the literature.

The Δv_{min} calculated from the intra-observer SEM_{intra} (table 3.2) were plotted

versus the mean volume of each tumour size bin (table 3.1) on a log-log graph. The use of a log-log graph is justified by figure 3.2, which shows a log-log relationship between the standard deviations of the observers' measurements and the mean tumour volumes measured. Linear regression was performed to interpolate the value of Δv_{min} over the range of volumes studied, yielding the result

$$\log_{10}\Delta v_{min}(V) = 0.95\log_{10}(V) - 0.65. \quad (3.3)$$

Equation 3.3 can be used to determine the change in volume at which the tumour should next be imaged, $\Delta v_{min}(V)$, as a function of the current volume, V (figure 3.6a).

Figure 3.6a can be used to compute the time needed to observe a detectable change in volume. Exponential growth is described by the equation $V(t) = V(0)\exp[(\log_e 2/t_2)t]$, where the volume doubling time, t_2 , can be determined from two measurements that differ by more than Δv_{min} . The exponential function can be rearranged to obtain an equation for t as a function of volume and doubling time, which can be used to compute the times at which the tumour reaches sizes V and $V + \Delta v_{min}(V)$. Taking the difference of these times yields

$$\Delta t_{min} = \frac{t_2}{\log_e 2} \log_e \left[1 + \frac{\Delta v_{min}(V)}{V} \right], \quad (3.4)$$

where Δt_{min} is the time needed to observe a detectable change in volume and $\Delta v_{min}(V)$ can be read from figure 3.6a. The specific equation will change with different mathematical growth models, but can always be derived from an equation like equation 3.3.

Figure 3.6b shows Δt_{min} versus the current tumour volume from equation 3.4. A series of curves are constructed for doubling times ranging from 1 to 13 days, which correspond to our empirical data for a variety of liver metastases and a transgenic

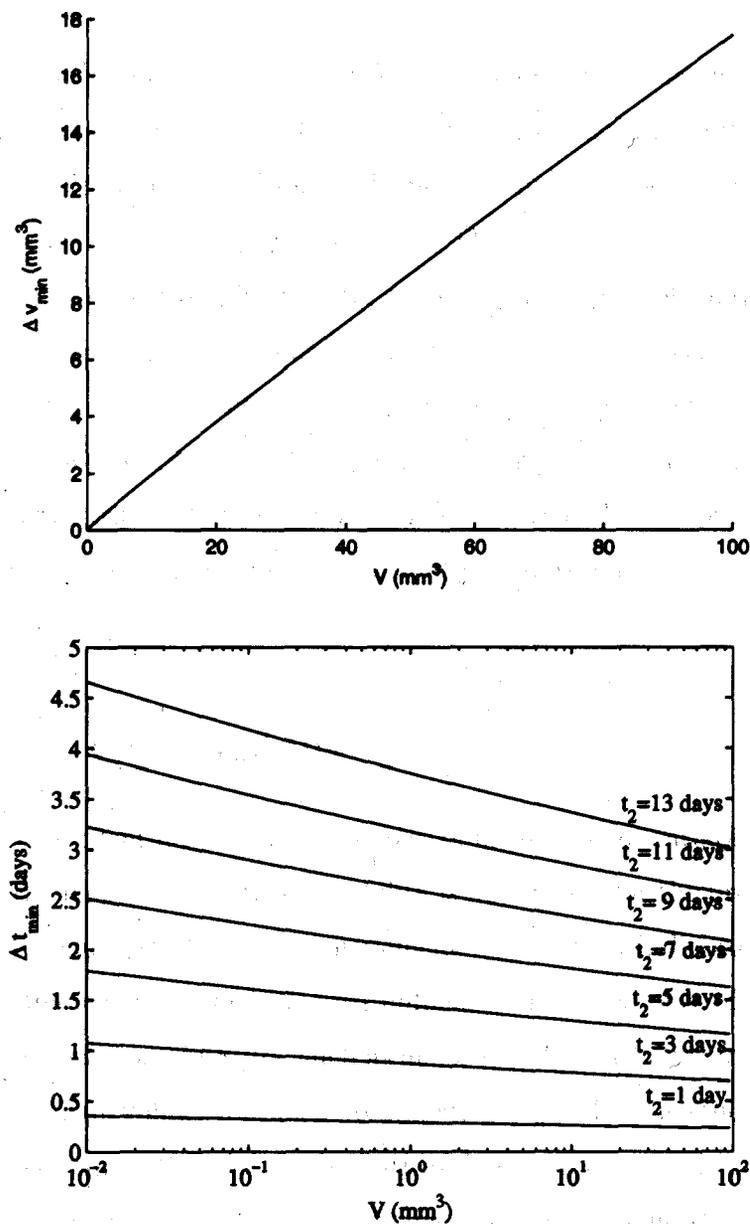


Fig. 3.6: (a) Minimum detectable change in volume (Δv_{min}) versus current volume (V). (b) Time needed to observe a detectable change in volume for different doubling times (t_2), assuming exponential growth. Doubling times range from 1 to 13 days. Both graphs are based on the intra-observer minimum detectable change in volume $\Delta v_{min}(V)$.

prostate cancer model [6, 7]. Δt_{min} increases as doubling time increases and decreases as the tumour grows.

The estimated Δt_{min} could aid in the design of imaging schedules in both growth and treatment studies. Delaying imaging until a significant change in tumour volume is expected enables resources to be used more efficiently and decreases the risks associated with anaesthetic exposure. The wide range of time intervals to wait prior to re-imaging shows that imaging as frequently as possible may not always be advisable. From the initial set of curves shown in figure 3.6b, adaptations can be made for different treatment schedules with knowledge of the growth rates of untreated tumours and how the treatment affects the tumours. To evaluate a novel therapeutic agent, a small trial could be conducted to estimate the tumour growth rate, or imaging intervals based on tumour growth in the control animals could be used for the entire study.

3.5 Conclusions

The identification of the sources of measurement variability and the quantification of the magnitude of this variability is imperative for the proper design of a longitudinal imaging study. The measurement variability will dictate the minimum detectable volume change, which in turn influences the scheduling of imaging sessions and the interpretation of results, such as an observed change in tumour volume or a difference in tumour volume between treatment groups. In this study, we have quantified the measurement variability that arises from single or multiple-observer measurement, altering the segmentation inter-slice distance, varying the position of the tumour of interest in relation to the ultrasound depth of field, and from allowing the anaesthetized animals to maintain a natural breathing cycle. There are numerous means by which measurement variability can be minimized. For each individual tumour being tracked longitudinally, one observer should be designated to measure tumour

volume throughout the experiment, although different observers can follow different tumours without impacting measurement variability. The tumour of interest should always be kept within, or above, the focal zone when acquiring three-dimensional images. When segmenting the tumour in parallel planes, the maximum inter-slice distance that yields a measurement variability within the intra-observer measurement variability should be used and will be a function of tumour size. Respiratory gating is not required to control volume measurement variability. Since the inherent contrast of liver metastases in ultrasound images is specific to the tumour type being studied, the absolute value of the measurement variabilities will change with different tumour models, yet the general guidelines outlined in this paper - designated observers, control of tumour position relative to focal depth, proper calculation of inter-slice distance for segmenting, and absence of a requirement for respiratory breathing - are likely to hold true for a variety of models. These results are necessary for the proper design of a longitudinal imaging study. The procedure outlined here will be applicable to studies of measurement variability with other pre-clinical imaging modalities.

References

- [1] A. Kamb, "What's wrong with our cancer models?," *Nat. Rev. Drug Discov.*, vol. 4, pp. 161-165, 2005.
- [2] A. Teicher Beverly, *Tumor Models in Cancer Research*, Humana Press Inc., Totowa, New Jersey, 2002.
- [3] C. Khanna and K. Hunter, "Modeling metastasis *in vivo*," *Carcinogenesis*, vol. 26
- [4] F. S. Foster, C. J. Pavlin, K. A. Harasiewicz, D. A. Christopher, and D. H. Turnbull, "Advances in ultrasound biomicroscopy," *Ultrasound Med Biol*, vol. 26, no. 1, pp. 1-27, 2000.
- [5] F. S. Foster, M. Y. Zhang, Y. Q. Zhou, G. Liu, J. Mehi, E. Cherin, K. A. Harasiewicz, B. G. Starkoski, L. Zan, D. A. Knapik, and S. L. Adamson, "A new ultrasound instrument for *in vivo* microimaging of mice," *Ultrasound Med Biol*, vol. 28, no. 9, pp. 1165-72, 2002.
- [6] K. C. Graham, L. A. Wirtzfeld, L. T. MacKenzie, C. O. Postenka, A. C. Groom, I. C. MacDonald, A. Fenster, J. C. Lacefield, and A. F. Chambers, "Three-dimensional high-frequency ultrasound imaging for longitudinal evaluation of liver metastases in preclinical models," *Cancer Res*, vol. 65, no. 12, pp. 5231-7, 2005.
- [7] L. A. Wirtzfeld, G. Wu, M. Bygrave, Y. Yamasaki, H. Sakai, M. Moussa, J. I. Izawa, D. B. Downey, N. M. Greenberg, A. Fenster, J. W. Xuan, and J. C. Lacefield, "A new three-dimensional ultrasound microimaging technology for preclinical studies using a transgenic prostate cancer mouse model," *Cancer Res*, vol. 65, no. 14, pp. 6337-45, 2005.

- [8] D. H. Turnbull, J. A. Ramsay, G. S. Shivji, T. S. Bloomfield, L. From, D. N. Sauder, and F. S. Foster, "Ultrasound backscatter microscope analysis of mouse melanoma progression," *Ultrasound Med Biol*, vol. 22, no. 7, pp. 845-53, 1996.
- [9] A. M. Cheung, A. S. Brown, L. A. Hastie, V. Cucevic, M. Roy, J. C. Lacefield, A. Fenster, and F. S. Foster, "Three-dimensional ultrasound biomicroscopy for xenograft growth analysis," *Ultrasound Med Biol*, vol. 31, no. 6, pp. 865-70, 2005.
- [10] Rooks Veronica, Beecken Wolf-Dietrich, Iordanescu Irina, and A. Taylor George, "Sonographic evaluation of orthotopic bladder tumors in mice treated with tnp-470 an angiogenic inhibitor," *Academic Radiology*, vol. 8
- [11] M. Galie, M. D'Onofrio, L. Calderan, E. Nicolato, A. Amici, C. Crescimanno, P. Marzola, and A. Sbarbati, "In vivo mapping of spontaneous mammary tumors in transgenic mice using MRI and ultrasonography," *Journal of Magnetic Resonance Imaging*, vol. 19, no. 8, pp. 570-579, 2004.
- [12] Schmitz Voker, Tirado-Ledo Lucia, Tiemann Klaus, Raskopf Esther, T. Homas Heinicke, Ziske Carsten, A. Conz 'alez-Carmona Maria, Rbae Christian, Wernert Nicolas, Prieto Jesùs, Qian Cheng, Sauerbruch Tilman, and H. Caselmann Wolfgang, "Establishment of an orthotopic tumour model for hepatocellular carcinoma and non-invasive in vivo tumour imaging by high resolution ultrasound in mice," *Journal of Hepatology*, vol. 40
- [13] V. L. Morris, I. C. MacDonald, S. Koop, E. E. Schmidt, A. F. Chambers, and A. C. Groom, "Early interactions of cancer cells with the microvasculature in mouse liver and muscle during hematogenous metastasis: videomicroscopic analysis," *Clin Exp Metastasis*, vol. 11, no. 5, pp. 377-90, 1993.
- [14] I. J. Fidler, "Selection of successive tumour lines for metastasis," *Nat New Biol*, vol. 242, no. 118, pp. 148-9, 1973.

- [15] S. Tong, D. B. Downey, H. N. Cardinal, and A. Fenster, "A three-dimensional ultrasound prostate imaging system," *Ultrasound Med. Biol.*, vol. 22 1996.
- [16] A. Landry, J. D. Spence, and A. Fenster, "Measurement of carotid plaque volume by 3-dimensional ultrasound," *Stroke*, vol. 35, no. 4, pp. 864-869, 2004.
- [17] A. Fenster, D. B. Downey, and H. N. Cardinal, "Three-dimensional ultrasound imaging," *Phys Med Biol*, vol. 46, no. 5, pp. R67-99, 2001.
- [18] A. Landry, J. D. Spence, and A. Fenster, "Quantification of carotid plaque volume measurements using 3D ultrasound imaging," *Ultrasound Med Biol*, vol. 31, no. 6, pp. 751-62, 2005.
- [19] P. Armitage and G. Berry, *Statistical Methods in Medical Research*, Blackwell Scientific Publications, Oxford, UK, 1987.
- [20] L. A. Hastie, K. C. Graham, A. C. Groom, I. C. MacDonald, A. F. Chambers, A. Fenster, and J. C. Lacefield, "Variability of three-dimensional high-frequency ultrasound measurements of small tumor volumes," in *Ultrasonics Symposium, 2004 IEEE*, 2004, vol. 3, pp. 2185-2188 Vol.3.
- [21] M. Eliasziw, S. L. Young, M. G. Woodbury, and K. Fryday-Field, "Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example," *Phys Ther*, vol. 74, no. 8, pp. 777-88, 1994.
- [22] J. R. Mitchell, S. J. Karlik, D. H. Lee, M. Eliasziw, G. P. Rice, and A. Fenster, "The variability of manual and computer assisted quantification of multiple sclerosis lesion volumes," *Med. Phys.*, vol. 23
- [23] H. X. Xu, X. Y. Yin, M. D. Lu, G. J. Liu, and Z. F. Xu, "Estimation of liver tumor volume using a three-dimensional ultrasound volumetric system," *Ultrasound Med Biol*, vol. 29, no. 6, pp. 839-46, 2003.

- [24] A. Landry and A. Fenster, "Theoretical and experimental quantification of carotid plaque volume measurements made by three-dimensional ultrasound using test phantoms," *Med. Phys.*, vol. 29, no. 10, pp. 2319–27, 2002.

Chapter 4

Monte Carlo Growth Curve

Simulations for Planning

Longitudinal Imaging Experiments

with Mouse Cancer Models

The contents of this chapter are in preparation to be submitted as a paper with the author list: L. A. Wirtzfeld, A. Fenster and J. C. Lacefield

4.1 Introduction

Preclinical animal studies of cancer growth and treatment frequently rely on single time-point endpoint analysis, or imprecise caliper measurements for subcutaneous tumours, to compare treatment groups and determine any effects. This method provides limited information on tumour growth or changes in tumour growth as a result of the treatments. For example, tumours that respond to therapy initially but quickly regrow and those that see a reduction in growth rate for the duration of the treatment can easily be indistinguishable at end point, yet suggest substantially different

mechanisms and consequently differences in how the treatment should be administered [1]. High-frequency ultrasound has been employed to track growth *in vivo* in mouse models of soft-tissue tumours [2, 3] and can aid in providing information on tumour growth, which could be used to evaluate treatment effectiveness and determine appropriate treatment scheduling. To obtain as much information as possible from these experiments, they must be designed appropriately to ensure that sufficient data are collected for the appropriate analysis to be applied.

Preclinical imaging will make it easier and consequently increasingly common to monitor tumour growth through time, requiring appropriate analyses of the tumour growth. There is a great deal of literature examining tumour growth, appropriate mathematical functions that can be used to describe this growth and how treatments, along with their responses, can be modelled [4, 5].

A wide range of functions are used to model tumour growth, ranging from simple exponential functions to complex growth curves including stochastic parameter changes over time to account for transitions between growing and dormant states [6, 7]. For this paper, the discussion will be constrained to the exponential and Gompertz functions as they are two of the most frequently used functions. The exponential function has been used extensively, including in the seminal paper by Skipper *et al.* [8] in a leukemia model. The exponential function is defined as:

$$V(t) = V_0 e^{at}, \quad (4.1)$$

where $V(t)$ is the volume at time t , V_0 is the initial volume at $t = 0$ and a is the growth rate which is related to the doubling time, $t_d = \frac{\ln 2}{a}$. Due to the constant growth rate, the exponential function increases without bound and is unable to take into account any growth limiting factors such as limited nutrients and space. While the work by Skipper *et al.* [8] was applied to leukemia and this model has been effective for many blood cancers, it has also been applied extensively to solid tumours [9].

The Gompertz function, which is also extensively used [5, 9–11], has an additional

parameter relative to the exponential function that gives the Gompertz an exponentially decaying rate constant which results in a decelerating growth rate over time. The Gompertz is a sigmoidal shaped function defined as:

$$V = V_0 e^{\frac{r}{b}(1-e^{-bt})}, \quad (4.2)$$

where b represents the rate at which growth slows down, $\frac{r}{b} = \ln(V_\infty/V_0)$ and V_∞ is the maximum volume that the function approaches asymptotically. There is a great deal of the literature on this model, especially by Larry Norton *et al.* [12, 13].

Figure 4.1 shows example exponential and Gompertz functions with the same instantaneous growth rate at time zero. There is a limited period of time around $t = 0$ where there is very little difference between the exponential and Gompertz functions; however, they diverge quite substantially over time. For many of the analyses performed on fitted growth data, the differences between these two functions would give different results and interpretations, including different estimated tumour initiation times and different time to reach a lethal tumour burden.

In the literature, many analyses are based on the functional form of growth being the same before and after treatment [8, 14]. Limited data availability can make it difficult to determine if two curves are the same or if they can be considered parallel with the same growth parameters but different initial volume at $t = 0$. Parallel curves would allow for an offset in the initial volume parameter, which produces an offset in time such as what might arise from a treatment experiment. In a study of rat brain tumour growth [15], exponential growth curves were fit to pre-treatment and post-treatment growth data without the requirement for the same growth rate post-treatment. Ross *et al.* [15] found the growth rate to be retarded after the treatment and backprojected this growth rate to estimate a cell-kill, while being sure to attribute the appropriate portion of the growth delay to the change in growth rate. This study clearly demonstrated that assuming the growth is unchanged after a treatment can change the outcome measures. If the growth curve post-treatment is not the same as

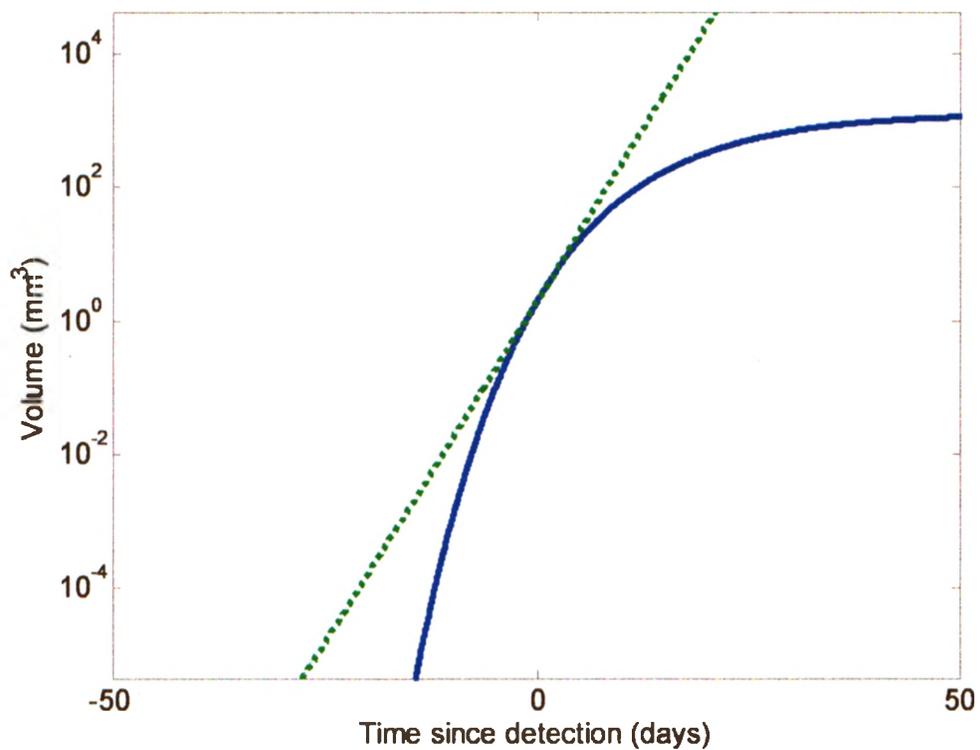


Fig. 4.1: An example of exponential (dotted line) and Gompertz (solid curve) functions with the same growth rate at time zero (the theoretical time of tumour detection). The Gompertz function has a higher growth rate prior to time zero and a decreased growth rate after time zero, compared to the exponential function which has a constant growth rate.

pre-treatment, then how it changes should influence how results are interpreted.

Simulations of growth data based on experimentally determined growth parameters enable curve fitting analyses to be performed on a large number of data sets and offer insight into how to design preclinical imaging experiments to maximize the chance of seeing the effect of a treatment. Without this information, it is difficult to know how imaging experiments should be designed as it is a challenge to determine how individual imaging and tumour-growth parameters influence the end results. If the tumour is imaged too frequently, there are risks to the animals from anesthetic, time and costs for the imaging but no additional information gained. If the imaging is too infrequent, the analysis that can be done on the growth curves will be limited and potentially result in failed experiments and wasted time, money and animals.

This paper presents Monte Carlo simulations of tumour growth based on the range of parameters obtained from curves fit to measurements of actual B16F1 liver metastases by three-dimensional high-frequency ultrasound [2]. Based on the simulated growth, a variety of parameters that contribute to an imaging experiment are varied to determine their influence the relative goodness of fit results for both exponential and Gompertz functions fitted to the data. The parameters examined include experimental design parameters (imaging frequency, initial and final imaging time and range of volume imaged), ultrasound system parameters (measurement variability, minimum and maximum size that can be imaged) and biological parameters (growth of the tumour model). Our study focuses on the experimental design and ultrasound system parameters as they are more readily controlled or improved upon.

Three sets of simulations were performed. In the first case, the fitting results of the exponential and Gompertz functions were compared to determine for what imaging parameters the Gompertz fits better than the exponential, given that it is known that the underlying function is more complex than the Gompertz or exponential.

In the second case, two different curves were generated and each fitted with Gompertz functions individually and using a constrained fit to generate parallel Gompertz

functions to the fitted two data sets. The two fitted curves could be compared to determine if there was an advantage to fitting the higher degree-of-freedom, unconstrained Gompertz functions to the data. If the individually fitted Gompertz offer no curve fitting advantage, this implies the two data sets are in fact close to parallel and consequently cannot be distinguished from each other in terms of growth parameters.

In the third case, each data set was divided in half to simulate a before and after treatment scenario. Both unconstrained Gompertz and constrained Gompertz functions, where only the initial volume can be different between two data sets, were fit to the early and late halves of the simulated data sets. Given that the two halves of the data set arise from the same growth curve, it would be expected that the constrained Gompertz function would fit the data as well as the unconstrained. If the parallel Gompertz functions offer an advantage, the implication is that the two data sets are considered close to parallel and therefore their growth parameters would be the same.

4.2 Materials and Methods

4.2.1 Simulations

All simulations and curve fitting were performed in Matlab (Mathworks, Natick, Massachusetts) using the built in statistical toolbox for non-linear least-squares optimization.

In order to simulate a growth function which can be used to generate actual tumour volume data, an appropriate growth model with an additional degree of complexity is necessary. The generalized logistic function [16] can be used to simulate the actual tumour growth. The generalized logistic allows for the point of inflection to occur anywhere along the sigmoidal curve, whereas the logistic function exhibits a symmetry about this inflection point, thus limiting its flexibility. The generalized

logistic function is defined as:

$$V = \theta \left(\frac{V_0^\alpha}{V_0^\alpha + (\theta^\alpha - V_0^\alpha) e^{-\alpha t}} \right)^{\frac{1}{\alpha}}, \quad (4.3)$$

where θ is the carrying capacity, which the tumour volume approaches asymptotically, and α determines how quickly or slowly the function reaches the carrying capacity. The generalized logistic function can be simplified to a Gompertz function in the case of $\alpha \rightarrow 0^+$ [16].

A simulated true growth function was generated by randomly selecting coefficients for the generalized logistic function. The distribution for the coefficients was based on data fitted to tumour volume data obtained from imaging B16F1 (murine melanoma) liver metastases in a syngeneic mouse model [2] to allow for reasonable values to be selected. In that study, tumours were typically imaged from approximately 0.05 mm^3 to 300 mm^3 and exponential growth curves indicated volume doubling times of $1.2 \text{ days} \pm 0.2 \text{ days}$ (mean \pm SD) and $1.4 \pm 0.4 \text{ days}$ for two experimental groups. α was selected from a uniform distribution between 0.2 days^{-1} and 0.5 days^{-1} , which correspond to doubling times of 3.7 days to 1.4 days since the B16F1 tumours are a rapidly growing tumour line. α varies between 0 and 1 and was selected on a uniform distribution to allow how rapidly the curve saturates to vary. V_0 was selected from a log-normal distribution between $10^{-5.7} \text{ mm}^3$ and $10^{0.74} \text{ mm}^3$, which corresponds to a range of 1 to 1000 cells with a diameter of $15 \mu\text{m}$. Metastasis models would be expected to start with a small number of cells, whereas orthotopic models often start with a large number of cells injected in one location. θ was selected from a log-normal distribution between 10^2 mm^3 and 10^{10} mm^3 based on values of θ or V_∞ from the Gompertz function obtained from fitted actual tumour growth data. This function was considered the true growth of the simulated tumour, which was then sampled at several different proposed imaging intervals and for several different minimum and maximum imaging time points. Sampled data were obtained to simulate imaging every 1, 2, 3, 4 and 6 days with sampling beginning between 0 and 25 days post

injection (with 5 day step size) and ending between 35 and 60 days post injection (with 5 day step size). Any sampled data sets with less than or equal to five data points were discarded as there are insufficient degrees of freedom to perform all the curve fitting. Variability, randomly selected from a normal distribution with standard deviation based on the experimentally determined intra-observer variability [17], was added to each simulated data point. A summary of the parameters selected to generate the growth function, run the simulations and sample the data is outlined in table 4.1. Figure 4.2A shows an example function with the sampled points. The same simulated data after including measurement variability are shown in figures 4.2B and 4.2C.

This set of data points was used as the simulated experimental data. Exponential and Gompertz functions were both fit to the data sets, see figure 4.2D, using nonlinear least squares regression. The residuals were also calculated and used to compute the sum of squares between the data and fitted curve. From each simulation, the chosen parameters were recorded, as well as initial and final volumes, volume range, number of data points and duration of data collection. In order to perform statistical analyses, parameters had to be binned to allow different levels of the variables to be evaluated. Individual parameters were binned based either on linear or logarithmic distributions, depending on the individual parameter distribution, to try to maintain a consistent number of cases across all bins.

4.2.2 Statistical Methods

In order to determine how well the exponential and Gompertz functions fit the simulated data, statistical analysis is required. An increase in the complexity of the function (from exponential to Gompertz) typically results in a reduction in the sum-of-squares (the sum of the squares of the volume between the data and fitted curve); however, this is insufficient to determine whether there is an improvement in fit [18]. There needs to be a sufficient improvement in the sum-of-squares to justify the additional degrees of freedom in the more complex function.

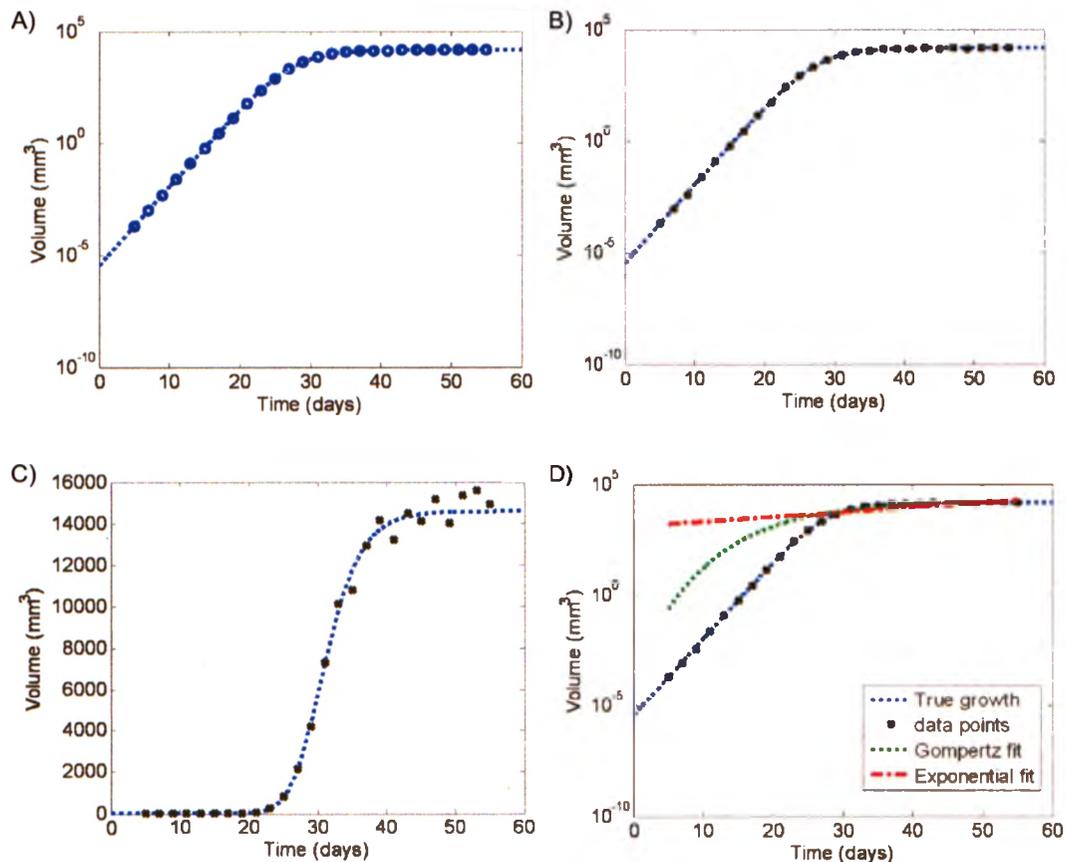


Fig. 4.2: Graphs show the steps in a simulation to generate data and perform the curve fitting. All graphs show the volume in mm^3 versus time in days where day 0 would be the beginning of the experiment. A) Shows the generalized logistic function used as the true growth function (dashed line) with the sampled data points (circles) to indicate the time points that would correspond to an imaging time point. Data points are acquired every 2 days from 5 to 55 days. B) Shows the data points (cross) that will be used in the curve fitting which were obtained by adding variability to the sample true growth. C) Shows the same curve and data points as in plot B but on a linear graph to allow the variability of the data points about the truth function to be seen. D) Shows the truth curve, data points and a fitted exponential (dash dot) and Gompertz function (dotted).

Variable (units)	Values
Randomly selected parameters for generalized logistic function	
α	uniform distribution 0 to 1
a (days ⁻¹)	uniform distribution 0.2 to 0.5
V_0 (mm ³)	logarithmic distribution $10^{-5.7}$ to $10^{0.74}$
θ (mm ³)	logarithmic distribution 10^2 to 10^{10}
Factors varied in simulations	
Imaging interval (days)	1, 2, 3, 4, 6
First imaging day (days)	0, 5, 10, 15, 20, 25
Last imaging day (days)	35, 40, 45, 50, 55, 60
Factors employed in statistical analysis	
Imaging interval (days)	1, 2, 3, 4, 6
Volume range (mm ³)	maximum volume - minimum volume
Length of time (days)	last - first imaging day
Minimum volume (mm ³)	Volume of data point on first imaging day

Table 4.1: A variety of parameters are used for different aspects of the data simulation and curve fitting. The distribution from which the parameters for the generalized logistic function are chosen is given. Three main factors were varied to give a range of values of the imaging interval and first and last day for imaging to determine how these parameters would influence curve fitting. The last four parameters listed are those which were used in the statistical analysis of the curve fitting.

Akaike developed a method for comparing fitted curves based on information theory and maximum likelihood that does not require functions to be nested [18, 19]. The Akaike method can also calculate the percentage chance each fitted function is correct, which allows a group of results to be compared since there is no hypothesis testing. The metric used is called the Akaike Information Criterion (AIC) and is calculated from the sum-of-squares of the fit (SS), the number of coefficients ($K - 1$) and the number of data points (N). For small values of N the AIC is inaccurate [18], therefore a correction has been applied to improve the results:

$$AIC_c = N \log_{10} \left(\frac{SS}{N} \right) + 2K + \frac{2K(K+1)}{N-K-1} \quad (4.4)$$

where the third term is the correction. There is a requirement to have at least two more data points (N) than the number of fitted coefficients in the function. The AIC_c values are difficult to interpret individually, but the more useful value is the difference in AIC_c values for the two fitted functions of interest. The ΔAIC is given by:

$$\Delta AIC = AIC_c^B - AIC_c^A \quad (4.5)$$

where B is a more complex model and A is a simpler model. The probability that one function is the better function to choose is also known as the Akaike weight and is calculated by:

$$AW = \frac{e^{-0.5\Delta AIC}}{1 + e^{-0.5\Delta AIC}} \quad (4.6)$$

The Akaike weight, or probability, gives the likelihood the more complex model is more appropriate out of the two functions being compared.

Univariate ANOVA tests were performed to determine which factors (see table 4.1 for a list of variables included in analysis) had statistically significant effects on the Akaike weights. For all statistical tests, the analyses were performed in SPSS (SPSS Inc., Chicago, Illinois) and $\alpha < 0.05$ as considered significant.

4.2.3 Gompertz Versus Exponential Functions

The goal of the first simulation was to determine when the Gompertz function would fit the data better than the exponential function. Data sets with fitting curves as shown in figure 4.2 were generated. Akaike weights were calculated for each data set to allow the fit of both functions to be compared. Given that the data sets all arose from an underlying generalized logistic function, which is more complex than the functions being fitted, the goal is to determine under what conditions it is possible to tell that the tumour growth curve is more complex in shape than an exponential function.

A univariate ANOVA was performed on the Akaike weights for the imaging interval, number of time points, minimum volume and range of volumes imaged. An $\alpha < 0.05$ was considered significant. Estimated marginal means were plotted for main effects and interaction to determine the trend in the Akaike weights as a function of the variables. Based on the results of this simulation, if it is possible for the Gompertz function to be the preferred model under a range of conditions, subsequent analyses will be performed with only Gompertz functions, otherwise they will be performed with exponential functions.

4.2.4 Comparison of Two Different Growth Curves

Subsequently, two generalized logistic functions with different coefficients were generated. Gompertz functions were fit to both data sets independently, and in addition, parallel Gompertz growth curve were fit to each pair of data sets. For the parallel Gompertz functions, the a and b parameters were constrained to be the same for the two curves and only V_0 could differ. The fit of the constrained parallel curves versus the unconstrained curves were compared to determine which method provides the better fit. If the constrained parallel curves fit better, a difference in growth between the two data sets cannot be determined, although it is known that the two

data sets arose from different growth curves.

A univariate ANOVA was performed to determine which parameters had significant effects on the Akaike weight for the two unconstrained fitted curves compared to the two parallel curves. The equivalent analysis to the previous section was performed, with the difference of comparing individual versus parallel Gompertz functions instead of two different functions.

4.2.5 Comparison of Early and Late Time Growth Curves

To determine the effect of fitting separate curves to two non-overlapping tumour volume time series, the simulated data sets were divided in half to give an early and late set of data points. The data will be analyzed with the Gompertz function as determined to be appropriate based on the results of section 4.2.3. Individual Gompertz functions were fit to each of the data sets without constraint initially. Additionally, Gompertz functions were fit to each data set with the constraint that the fitted functions must be parallel, with a and b the same between the two functions. This constraint allows only a variation in the y-axis crossing and consequently can approximate a growth delay that could potentially occur in the case of a treatment study. The fitted results were then analyzed to determine the conditions required for the parallel functions to fit better than the individual functions (i.e., Akaike weight less than 0.50, as this is selecting the less complex solution) given that it is known they arise from the same underlying growth curve.

Univariate ANOVAs were performed on the Akaike weights for the individual versus parallel Gompertz functions. Variables had to be analyzed in multiple ANOVAs due to the limitations in the number of levels and factors SPSS could handle. One ANOVA was performed on the imaging interval and the number of time points in each data set. A second ANOVA was performed on the imaging interval, minimum volume in the early and late growth curves and the volume range in the early and late growth curves.

4.3 Results

4.3.1 Gompertz Versus Exponential Functions

Figure 4.3 shows an example of a growth curve with both the exponential and Gompertz functions fitted to the data. In this case, at early time points both functions deviate from the simulated growth and data points. The univariate ANOVA

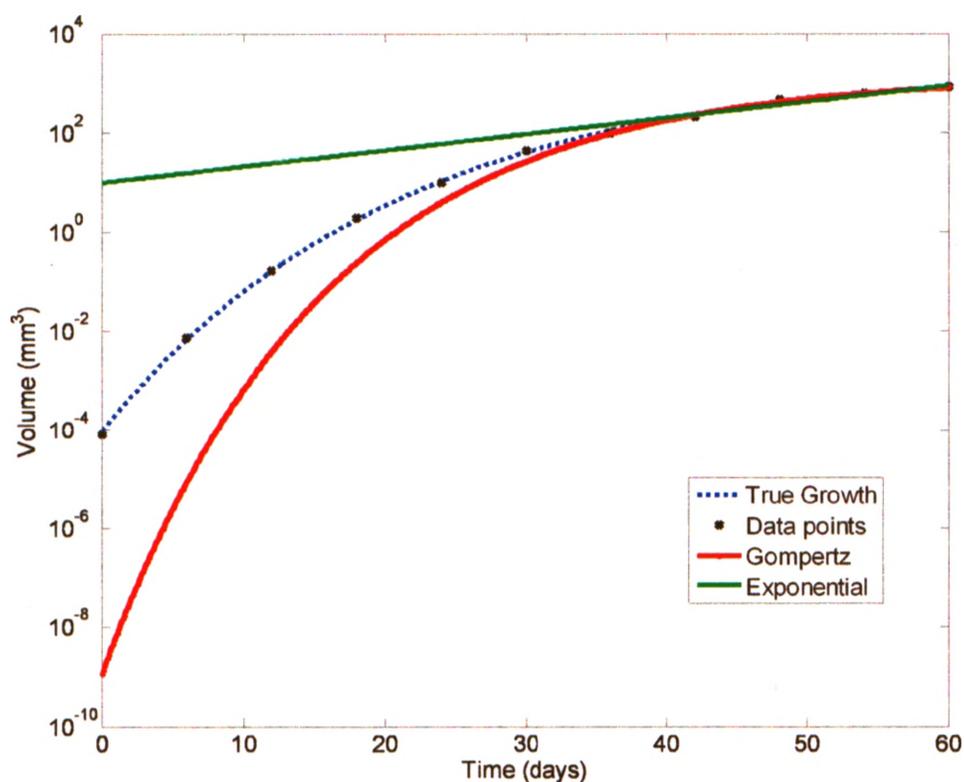


Fig. 4.3: An example of a simulated tumour growth function (blue) with a Gompertz (red) and an exponential (green) function fitted to the data points. Although there are data points distributed evenly over the entire time range, neither the Gompertz nor the exponential function is able to maintain a good fit with the data over the entire duration.

showed significant main effects and all interactions between for the imaging interval,

number of time points, minimum volume and range of volumes imaged. This means that the Akaike weights depend significantly on each variable, and that each variable has a relationship with all the other variables. This makes it difficult to draw simple conclusions regarding what ranges of each variable would be required to see a Gompertz fit as the preferred function, as that range would be depend on the values of other variables. Main effects indicate the value of the Akaike weight depends on the value of the different factors analyzed in the ANOVA.

Figure 4.4 shows a plot of the Akaike weights as a function of the length of time data is collected and the imaging interval in panel (A) and minimum volume imaged and volume range in panel (B). Figure 4.4A shows that as the number of data points increase, the Akaike weights increase, with the most frequent imaging schedule (1 day intervals) approaching the maximum Akaike weight for shorter lengths of time. An Akaike weight greater than 0.5 indicates that the Gompertz function has a higher chance of being the better function to choose. For the two most frequent imaging intervals (1 and 2 days) the Akaike weight is greater than 0.50 for any length of time longer than 15 days. For imaging every 6 days on the other hand, the experiment would need to last at least 40 days to produce sufficient data to choose the Gompertz function over the exponential function. Note that the Akaike weight increases rapidly at the low end of the length-of-time axis: in this region, only a small increase in number of data points leads to a noticeable increase in the Akaike weights.

The Akaike weight marginal means show a steady decrease for increasing minimum volumes when averaged over all volume ranges. When the volume ranges are examined separately, as shown in figure 4.4B, for small minimum volumes all curves show relatively high Akaike weights. However, as the minimum volume increases the smallest volume ranges show a very rapid decrease in Akaike weight. The results for the smallest volume ranges are highly dependent on the minimum volume that is imaged. If the minimum volume imaged was maintained less than approximately 0.09 mm^3 (about $550 \mu\text{m}$ diameter assuming spherical tumours), then for all other

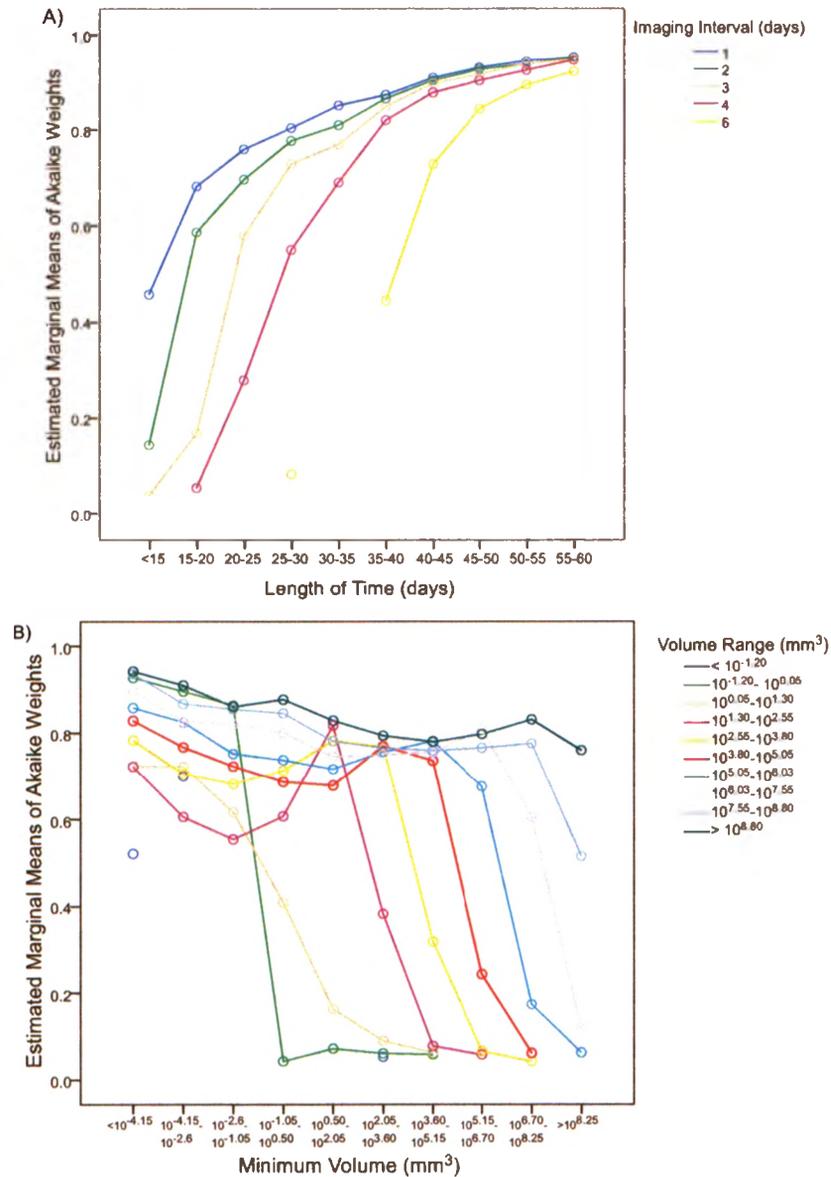


Fig. 4.4: Graphs show the marginal means for the Akaike weight comparing a Gompertz function to an exponential function. A) The Akaike weight increases with increased length of time the data spans and approaches close to 100% asymptotically. More frequent imaging results in a more rapid approach to the maximum value. B) For small minimum volumes there are relatively consistently high Akaike values, however there is a sudden drop as the minimum volume is increased which occurs at higher minimum volumes as the range of volumes is increased. This suggests that small volume ranges are particularly sensitive to the minimum size.

combinations of parameters the Akaike weight could be kept in the region where the Gompertz fits the curves well. Conversely, if the volume range was greater than 20 mm³ (above bin 3) the minimum volume could be increased up to 140,000 mm³ (about 64 mm diameter) before the Akaike weight fell below 0.5.

The strong dependence on the minimum imaged volume for the different volume ranges suggests that improving the ability to image smaller tumours could have a substantial effect on the ability to fit curves to data that span a small volume range. The minimum volume suggested at 0.09 mm³ would correspond to a spherical tumour with approximately 0.55 mm diameter, which has been detected by ultrasound with the B16F1 liver metastases. This minimum volume would allow Gompertz function to be used for all volume ranges. The lowest two volume-range bins correspond to very narrow ranges of growth, ≤ 1 mm³, but even a 20-mm³ volume range, which is included in bin 3, is a reasonable range of volumes to measure.

4.3.2 Comparison of Two Different Growth Curves

Subsequent results are only presented for the Gompertz function as the previous section demonstrates that it is possible to select variables such that the complex shape of the tumour growth can be fitted better with a Gompertz function than an exponential function. Figure 4.5 shows an example of two curves fitted with individual and then parallel Gompertz functions. The example shows a case where the parallel Gompertz functions are not a good match to fit both sets of data as their growth rates are substantially different.

A high value for the Akaike weight indicates that the more complex model - where the two data sets are fitted separately - is a better choice for fitting the data than the constrained parallel curves. The ANOVA shows a significant four-way interaction between the imaging interval, number of time points, minimum volume and volume range of the data set, indicating that all the experimental design variables combine to influence the value of the Akaike weight. All main effects are significant as well as

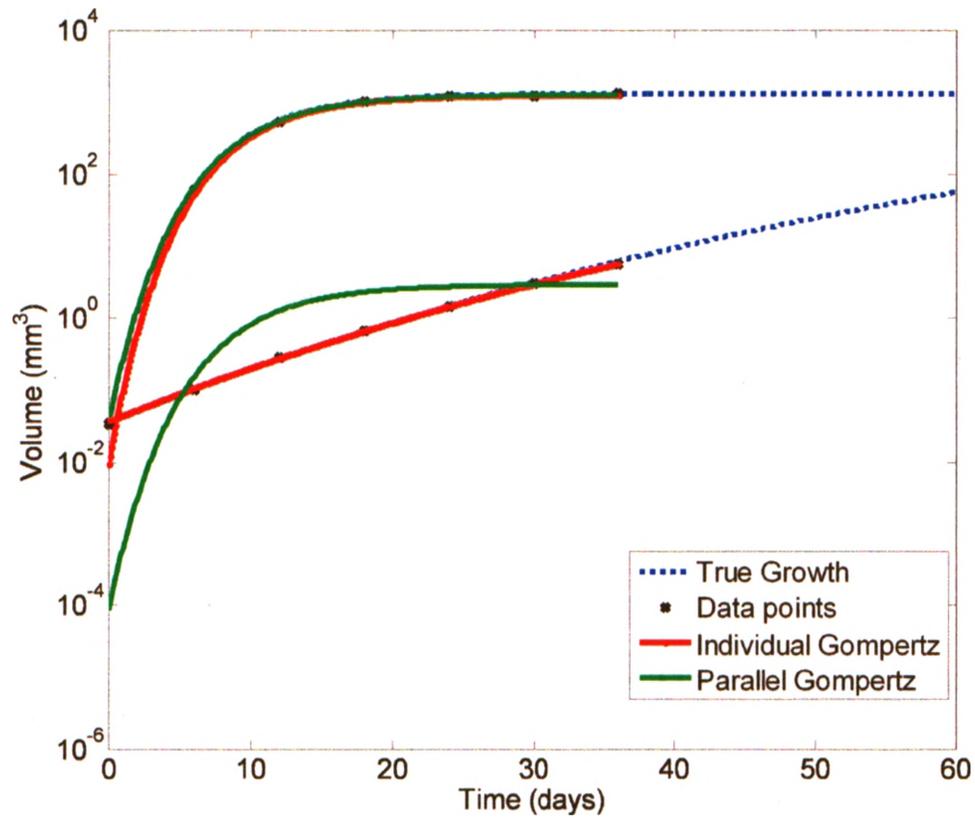


Fig. 4.5: Two different simulated tumour growth curves are shown with a Gompertz function fit to each data set (red) and with parallel Gompertz function fit to each data set (green). These two growth curves are clearly different enough that the parallel Gompertz functions are not able to follow the shape of the lower valued curve at all.

almost all the interactions. From plots of the marginal means for each of the variables, Akaike weights over 0.5 (indicating data sets should be fitted with separate functions) can be achieved for imaging intervals every four days or less, for lengths of time in bin 3 (25 days) or higher, for volume ranges above 20 mm^3 and for all minimum volumes. Figure 4.6A shows the values of the Akaike weight gradually decrease for increased imaging interval. The multiple lines for the individual binned volume ranges are approximately parallel, as is suggested by the non-significant interaction between these parameters. Figure 4.6B shows the same trend as in part A of the figure, with very high minimum volumes or long imaging intervals having Akaike weights below 0.5.

The requirements to be able to decide that two functions are in fact different based on requiring different Gompertz functions to fit the data correspond well with the requirements for the previous section to be able to fit the Gompertz functions to the data. The minimum volume (0.09 mm^3) and volume range (20 mm^3) are the same as the previous section with the length of time being slightly higher at 25 days with a maximum imaging interval of 4 days.

4.3.3 Comparison of Early and Late Time Growth Curves

Results for this section are only presented with Gompertz functions. Figure 4.7 shows an example simulation with the data divided into an early and late segments. In this example, using parallel Gompertz functions yields a poorer fit for the data than is achieved with individually fitted Gompertz functions.

For the first ANOVA, the imaging interval and number of time points were both statistically significant main effects as was their interaction. Marginal means remained below 0.50 (meaning constrained parallel Gompertz functions are acceptable) for all imaging intervals and were approximately 0.50 or below for imaging at less than or equal to 19 time points. In the second ANOVA, all the factors had a significant effect on the Akaike weights and many of the combinations of factors varied significantly

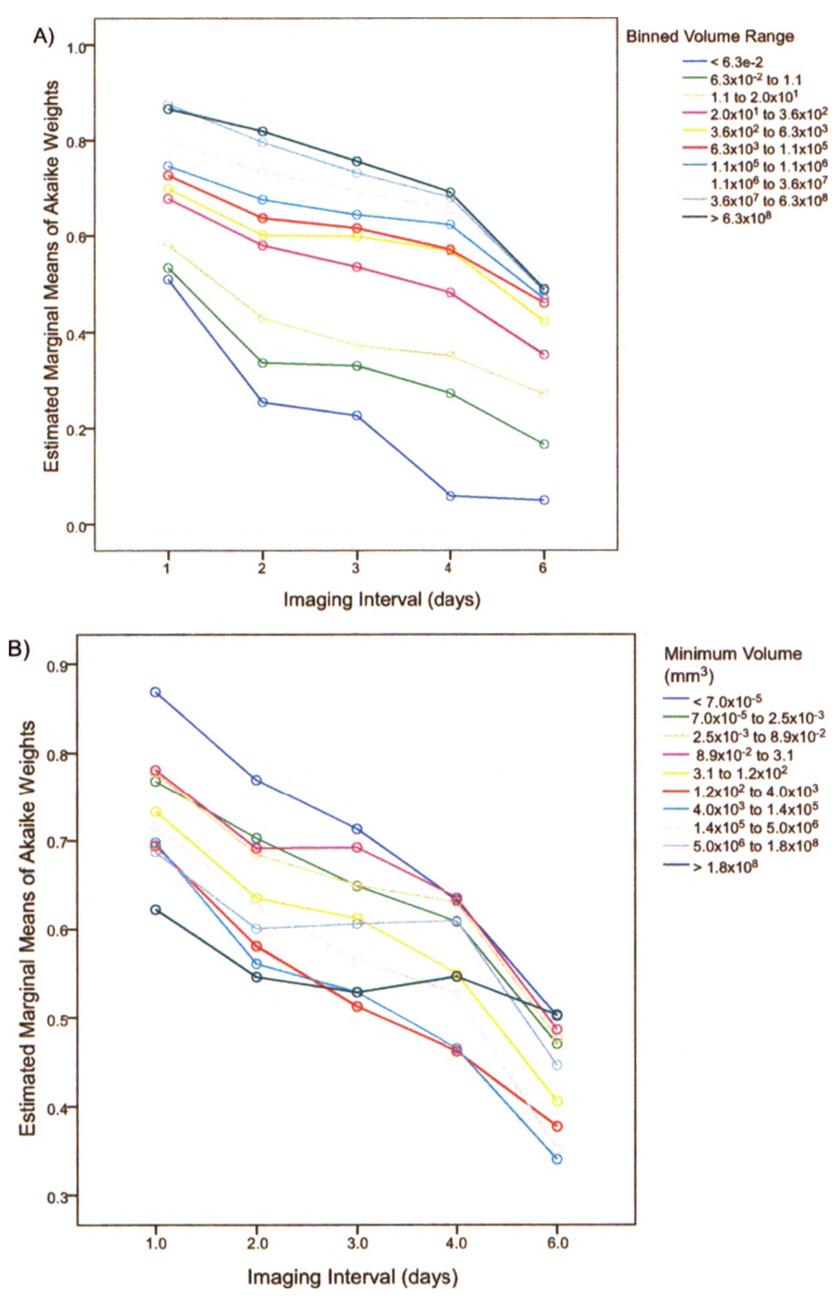


Fig. 4.6: Both graphs show the marginal means of the Akaike weight for comparing the fit individual Gompertz functions to two separate sets of growth data compared to fitting parallel functions. Akaike weights decrease for increased imaging interval and volume range, however, they decrease for increase minimum volume. A large range of parameters allow for the Akaike weight to be maintained above 0.5.

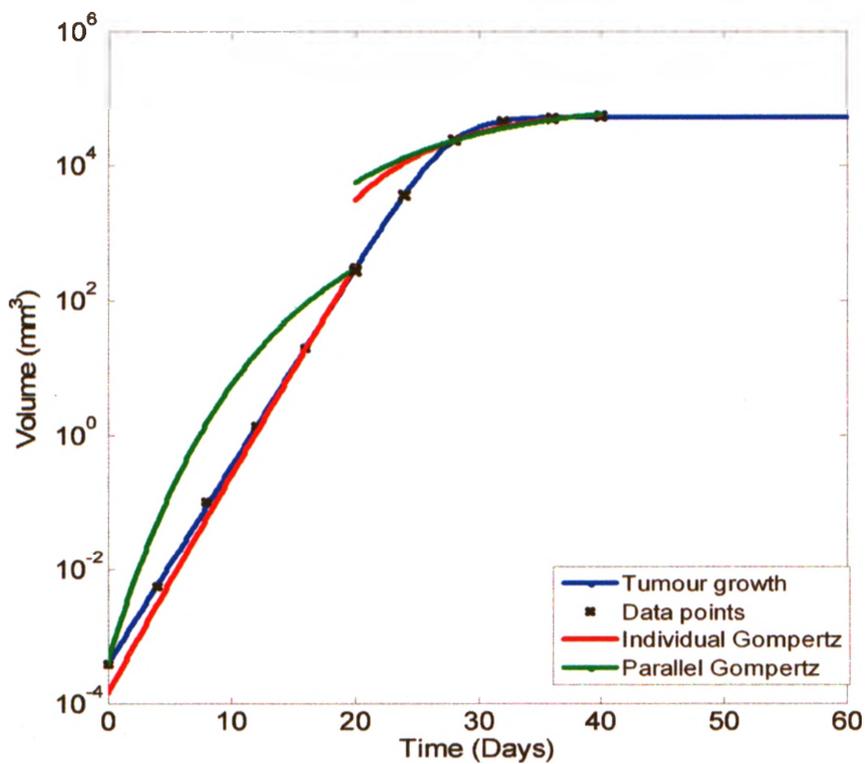


Fig. 4.7: The graph shows simulated tumour growth (blue) with individual Gompertz function fitted to the first half (up to day 20) and second half (beyond day 25) shown in red. Gompertz functions constrained to be parallel were fitted to the two data sets (green). For the early data set the constrained Gompertz curve fitting did not fit the data points as well as the unconstrained Gompertz function. For the later time points, there is very little difference between the two curves.

with each other. Based on the marginal means of the individual factors, to avoid an Akaike weight over 0.50 it would be necessary to maintain an imaging interval of at least 2 days and a minimum volume in the late curve above 0.5 mm^3 . Figure 4.8 shows the marginal means of the Akaike weights and the influence of the imaging interval with the minimum volume of each part of the curve. In figure 4.8A it is seen that for imaging intervals of 1 day, the lower minimum volumes imaged could have Akaike weights above 0.5. In figure 4.8B the results are more consistently below 0.5; however, for frequent imaging intervals and low minimum volume for the late data, the Akaike weight can be above 0.5.

Although there are ranges of all the plots where Akaike weights below 0.5 can be maintained, this would put limits on the experimental design factors in the opposite direction compared to the previous analyses (i.e., setting a maximum number of time points, whereas previous analyses have found a minimum number of time points). Determining if two segments of a data set arise from the same growth function is likely to be difficult to achieve consistently. Potential reasons for this outcome are described in the discussion.

4.4 Discussion

In order to fit a more complex function, such as a Gompertz function, sufficient data are required. It can be seen from the simulations that the majority of the experimental designs result in a better fit for the Gompertz function over the exponential function. The goodness-of-fit of the Gompertz function is more favourable when imaging begins with smaller tumour volumes rather than larger volumes, even across the same range of volumes being imaged. This suggests an opportunity to improve the ability to fit functions to the data. As described in section 4.3.1, the range of minimum volumes that show an improvement include tumours that would be approximately 0.55 mm in diameter (assuming a sphere), which have been imaged *in vivo*

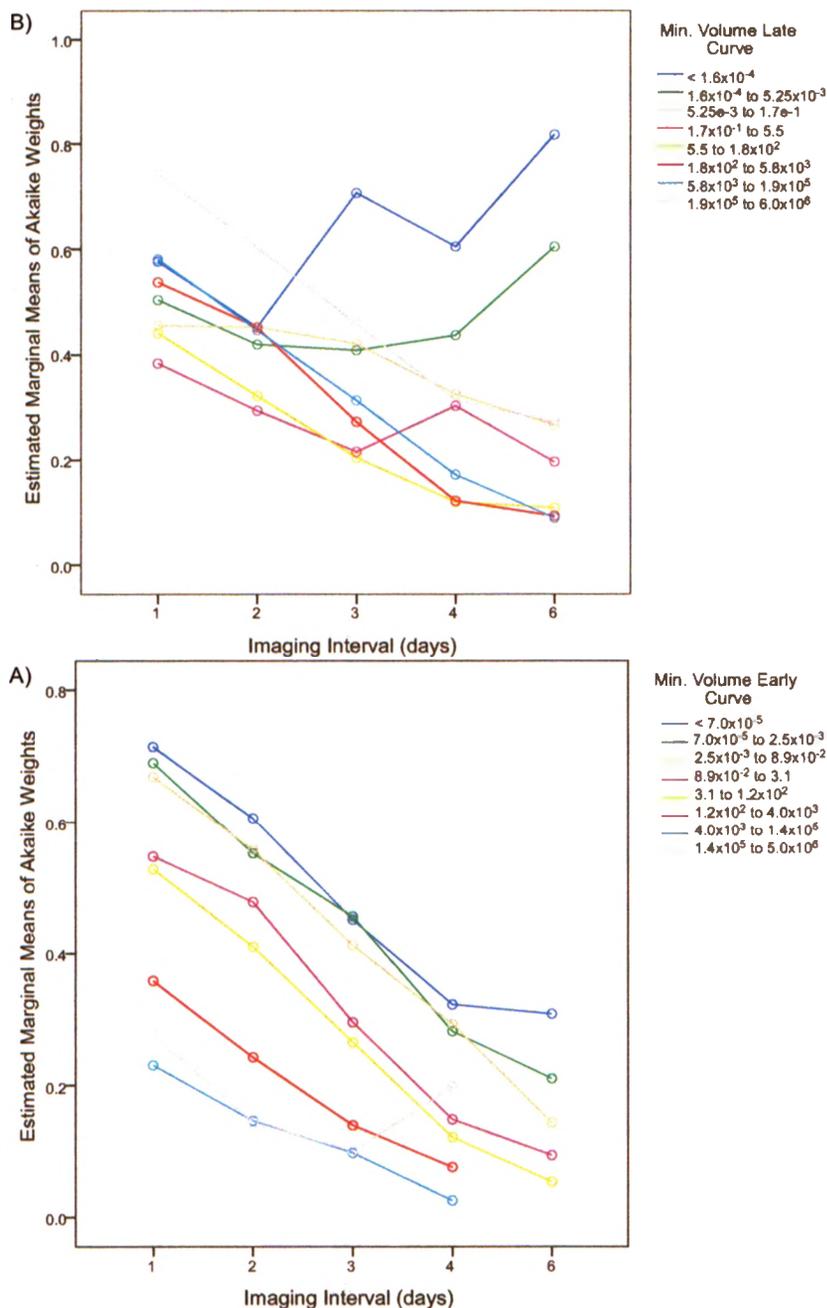


Fig. 4.8: Both graphs show the trends for the minimum volumes for each segment of the curve over different imaging intervals. A) The effects of minimum volume for the early part of the data set and imaging interval are plotted and show a range where the values of the Akaike weights are above 0.5, when it is desired to have values below 0.5 to indicate the parallel curves fitting better than the individual curves. B) The Akaike weights are more consistently below 0.5 across the imaging interval for the minimum volume of the late curve; however, the lowest minimum volumes still maintain values above 0.5.

in the B16F1 liver metastases model [2]. However, not every tumour was detected when it was that small. There are several reasons why this might be the case. There is a chance that some tumours are more easily visualized at this size than others due to biological variations between tumours. It is also possible that there is a minimum detectable tumour size and tumours just below this threshold grow from below the detection limit to well above it between imaging sessions.

Being able to correctly determine that two tumours are growing differently is reasonable as the constraints on the variables to achieve an Akaike weight over 0.5 in section 4.3.2 are not very much more rigorous than those required to have a better fit for the Gompertz function over the exponential function in the previous section. The imaging interval needs to be four days or less, whereas to determine a Gompertzian shape any imaging interval could be employed. Also, the length of time needs to be 25 days or longer compared to 15 days or longer for the previous section.

The measurement variability, which had a fixed dependence on tumour volume in these simulations, will have an impact on how well curve fitting works. Reduced variability will allow more precise measurements of tumour volume and consequently a better estimate of the growth parameters when curve fitting.

The feasibility of determining whether two data sets have the same growth rate parameters may be more limited as more data may result in over fitting, suggesting these data sets are in fact different from each other. Although the paired data sets in section 4.3.3 were both based on the same generalized logistic curve, the measurement variability introduces errors that appear to increase the difficulty of determining that the early and later data arise from the same growth curve, especially since the variability is a function of volume. Constraining the experimental design with both a maximum and minimum volume to image is likely not practical and could lead to the erroneous interpretation of results.

4.5 Conclusion

Simulations of tumour growth and curve fitting can be used to determine when there is sufficient data to draw conclusions about data sets. Simulations presented indicate that the Gompertz function can be fitted to volume data resembling B16F1 liver metastasis growth with parameters that are realistic. This information could be used to plan preclinical cancer studies to better enable imaging measurements to quantitatively characterize tumour growth and treatment responses.

References

- [1] ,” .
- [2] K. C. Graham, L. A. Wirtzfeld, L. T. MacKenzie, C. O. Postenka, A. C. Groom, I. C. MacDonald, A. Fenster, J. C. Lacefield, and A. F. Chambers, “Three-dimensional high-frequency ultrasound imaging for longitudinal evaluation of liver metastases in preclinical models,” *Cancer Res*, vol. 65, no. 12, pp. 5231–7, 2005.
- [3] A. M. Cheung, A. S. Brown, L. A. Hastie, V. Cucevic, M. Roy, J. C. Lacefield, A. Fenster, and F. S. Foster, “Three-dimensional ultrasound biomicroscopy for xenograft growth analysis,” *Ultrasound Med Biol*, vol. 31, no. 6, pp. 865–70, 2005.
- [4] Michael W. Retsky, Romano Demicheli, Douglas E. Swartzendruber, Paul D. Bame, Robert H. Wardwell, Gianni Bonadonna, John F. Speer, and Pinuccia Valagussa, “Computer simulations of a breast cancer metastasis model,” *Breast Cancer Research and Treatment*, vol. 45, pp. 193–202, 1997.
- [5] L. Norton, “Theoretical concepts and the emerging role of taxanes in adjuvant therapy,” *Oncologist*, vol. 6 Suppl 3, pp. 30–5, 2001.
- [6] Susan E. Clare, Faina Nakhlis, and John Carl Panetta, “Molecular biology of breast cancer metastasis the use of mathematical models to determine relapse and to predict response to chemotherapy in breast cancer,” *Breast Cancer Res*, vol. 2, no. 6, pp. 430–435, 2000.
- [7] *Cancer Medicine*, BC Decker Inc, Hamilton, Canada, 2003.
- [8] H. E. Skipper, Jr. F.M. Schabel, and W. S. Wilcox, “Experimental evaluation of potential anticancer agents: Xiii. on the criteria and kinetics associated with ‘curability’ of experimental leukemia,” *Cancer Chemother. Rep.*, vol. 35

- [9] L. Norton and R. Simon, "Growth curve of an experimental solid tumor following radiotherapy," *J Natl Cancer Inst*, vol. 58, no. 6, pp. 1735-41, 1977.
- [10] R. Chignola and R. I. Foroni, "Estimating the growth kinetics of experimental tumors from as few as two determinations of tumor size: implications for clinical oncology," *IEEE Trans Biomed Eng*, vol. 52, no. 5, pp. 808-15, 2005.
- [11] K. Rygaard and M. Spang-Thomsen, "Quantitation and gompertzian analysis of tumor growth," *Breast Cancer Res Treat*, vol. 46, no. 2-3, pp. 303-12, 1997.
- [12] L. Norton, "A gompertzian model of human breast cancer growth," *Cancer Res*, vol. 48, no. 24 Pt 1, pp. 7067-71, 1988.
- [13] L. Norton, "Evolving concepts in the systemic drug therapy of breast cancer," *Semin Oncol*, vol. 24, no. 4 Suppl 10, pp. S10-3-S10-10, 1997.
- [14] L. Norton and R. Simon, "Tumor size, sensitivity to therapy, and design of treatment schedules," *Cancer Treat Rep*, vol. 61, no. 7, pp. 1307-17, 1977.
- [15] Brian D. Ross, Yong-Jie Zhao, Eric R. Neal, Lauren D. Stegman, Matthew Ercolani, Oded Ben-Yoseph, and Thomas L. Chenevert, "Contributions of cell kill and posttreatment tumor growth rates to the repopulation of intracerebral 9l tumors after chemotherapy: An MRI study," *Proc. Natl. Acad. Sci.*, vol. 95, pp. 7012-7017, 1998.
- [16] L. Preziosi, "Cancer modelling and simulation," pp. 77-90. Chapman & Hall/CRC, New York, 2003.
- [17] L. A. Wirtzfeld, K. C. Graham, A. C. Groom, I. C. Macdonald, A. F. Chambers, A. Fenster, and J. C. Lacefield, "Volume measurement variability in three-dimensional high-frequency ultrasound images of murine liver metastases," *Phys Med Biol*, vol. 51, no. 10, pp. 2367-81, 2006.

- [18] Harvey Motulsky and Arthur Christopoulos, *Fitting Models to Biological Data using Linear and Nonlinear Regression A practical guide to curve fitting*, Oxford University Press, New York, 2004.
- [19] Kenneth P. Burnham, *Model selection and multimodel inference : a practical information-theoretic approach* / Kenneth P. Burnham, David R. Anderson., Springer, New York, 2002.

Chapter 5

Synthetic Aperture Focusing of a Single Element 40 MHz System

The contents of this chapter are in preparation to be submitted as a paper with the author list: L. A. Wirtzfeld, A. F. Chambers, A. Fenster and J. C. Lacefield

5.1 Introduction:

Three-dimensional high-frequency (30-40 MHz) ultrasound imaging is used to monitor tumour growth and treatment responses in mouse cancer models [1-3]. To perform these experiments, it is important to quantify and minimize the measurement variability to detect small changes in volume and determine the functional form of the growth for curving fitting. The tumour measurement variability depends on the tumour size and location within the field of view [4]. This poses a challenge determining values for the measurement variability as it will vary based on the object location within the field of view and will also change as the tumour grows due to the size dependence resulting in changing variability over the longitudinal experiments.

Commercially available high-frequency ultrasound systems use mechanically scanned, single element transducers. Low f-numbers are used, resulting in tightly focused

transducers with high lateral resolution at the fixed focal depth that rapidly degrades towards the near and far fields.

To achieve consistent high resolution through the entire image, the two options are to use an array to acquire the images or to use synthetic aperture focusing to retrospectively focus the images. For the first case, array transducers allow for the ultrasound beam to be focused at multiple depths within the image. There are a number of research groups working on developing both annular [5-7] and linear [8, 9] high-frequency arrays. Annular arrays allow for focusing along the beam but must still be mechanically steered to form a B-mode image, whereas linear arrays can focus the beam and do not require any mechanical scanning to produce an image. High-frequency arrays are technically challenging to produce due to the small element size required and the need to bond a wire to each element. In the case of linear arrays, an element spacing of the wavelength, λ , is required to avoid grating lobes within the images and would require the elements to be spaced approximately 40 μm apart for a 40 MHz probe.

Although high-frequency arrays are in development and experimental use, it will likely be some time before all high-frequency systems are using arrays as there are over 450 VisualSonics commercial systems* and many other custom single-element systems in use that will continue to be used for imaging. In these cases, using a synthetic aperture technique to retrospectively improve the image focusing would offer benefits.

Synthetic aperture focusing techniques (SAFT) have been employed with clinical array systems [10] and can be adapted for use with tightly focused single element transducers using a virtual source approach [11]. In this technique, the fixed focus is treated as a virtual source with sound waves propagating towards both the near and far fields as approximately spherical waves over a limited range of angles corresponding to the shape of the original beam. Each acquired scan line then produces a virtual

*Personal conversation with VisualSonics, October 2008

source at its focus. All these virtual sources can be combined in a fashion similar to an array to reconstruct an image focused at multiple depths. As the individual scan lines are acquired, the beam produced along a given scan line will overlap the beam produced for adjacent scan lines. Where the beams overlap at a given point, those scan lines can be used to retrospectively focus at that point by applying appropriate time delays as shown in figure 5.1.

The time delays can be calculated by determining the additional time it takes for the sound to travel from the virtual source to the point of interest compared to the time for sound to travel from the virtual source to the same depth as the point but along the centre of the beam. This can be seen in figure 5.1 as the difference between the length of time to travel the distance r compared to r' and it can be seen that this calculation will depend on the specific scan geometry. These time delays can be applied to reconstruct scan lines in a manner equivalent to applying delays to a linear or curvilinear array.

Synthetic aperture focusing can be performed by delay-and-sum beamforming with or without applying additional weighting. Apodization is commonly used to reduce sidelobe levels when beamforming arrays by giving higher weighting to the central elements and lower weighting to the outer elements of the subaperture. In addition, there are a number of adaptive weighting techniques that use the data to determine the appropriate weighting. Two adaptive techniques that have been applied to ultrasound are minimum variance [12, 13] and the generalized coherence factor [14, 15].

The generalized coherence factor (GAFF) is based on an earlier weighting known as the coherence factor (CF) which is the ratio of the energy of the coherent sum to the total incoherent energy for the subaperture [15]. The GAFF extends the CF by taking a Fourier transform of the data in the lateral direction and calculating the ratio of energy around zero frequency (DC) to the energy in the entire spectrum. The CF suffered from large variations due to the fact that it only takes the value

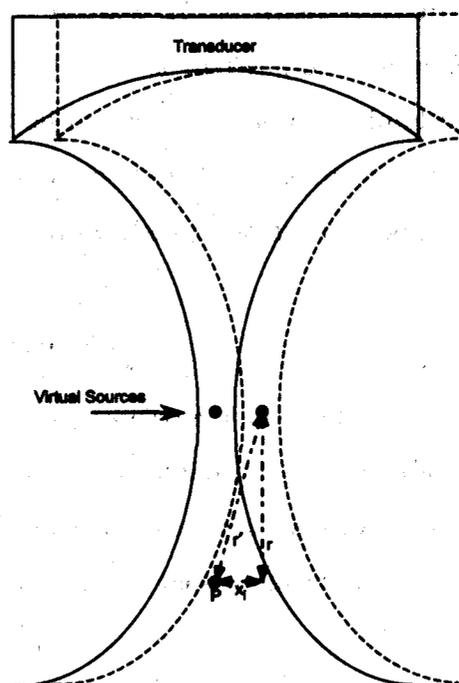


Fig. 5.1: The width of the ultrasound beam converges to the focus, which can be treated as a virtual source for each scan line. The set of virtual sources which are formed as the transducer is mechanically scanned can be treated in a manner analogous to array elements. Above and below the virtual source, the beam rapidly widens which means the beam from several scan lines will pass through each point, such as point P . The multiple scan-lines that pass through point P can be used to perform synthetic aperture focusing. Two beams from two different scan lines are shown that will overlap on point P . Delays can be calculated by determining the distance in flight time between r and r' , where x_i is the distance between scan lines.

at DC squared which is susceptible to noise. The goal of the GAFF is to assign lower weightings to strong off-axis signals compared to on-axis signals. The minimum variance beamformer aims to minimize the power for the beamformed signal and maintain unity gain at the focal point to reduce the main lobe width and suppress the sidelobes. The weights can be computed from the autocovariance matrix of the subaperture data, $W_{mv} = R^{-1}a/a^H R^{-1}a$, where W_{mv} are the weights, R is the autocovariance and a is a steering vector that can be maintained as a vector of ones. Although this calculation is fairly straight forward to implement, the coherence of the ultrasound signals means an additional step is required to smooth the data and give a good estimate of the autocovariance. Details can be found in [13].

5.1.1 Outline of chapter

This chapter describes synthetic aperture focusing techniques using several weighting schemes, both conventional and adaptive, to improve the focusing in ultrasound images acquired with a 40-MHz, single-element transducer and evaluate the effect of synthetic aperture focusing on the size measurement variability of point-like targets and small lesions.

To determine the impact on point-like targets, the cross-sectional area of small air channels was compared for the original and SAFT images. The area measurements were compared to determine which weighting technique provided the greatest improvement in point resolution. Gelatin lesion phantoms were employed to evaluate the impact of SAFT on 2D lesion-size measurement variability and determine how the various weighting techniques change the variability. Lesion images were acquired and measured with the centre of the lesion within, above and below the depth of field. To evaluate the technique *in vivo*, an experimental liver metastasis model was imaged and the SAFT were compared to the original.

5.2 Materials and Methods:

5.2.1 Synthetic Aperture Focusing Technique

In order to perform the SAFT, it is necessary to know the number of scan lines that pass through the beam width at each depth, z . In order to do this, the beamplot of the transducer was simulated from 2 mm to 10 mm depth with Field II [16] and the lateral resolution (LR) was measured as the -6 dB beam width at the centre of each focusing depth band. The LR varied from 0.1 to 2 mm. The number of scan lines used at a depth z in each sub-aperture was calculated,

$$N_z = \frac{2}{\text{line spacing}} \arctan \frac{LR}{2(SL + z)} \quad (5.1)$$

where SL is the shaft length from the pivot point to the transducer surface, z is the image depth from the transducer surface and the line spacing is defined above. Since circular sector images with azimuth angles ranging from $\pm 3.76^\circ$ were acquired in our experiments, the line spacing was 0.02° with 377 scan lines acquired over the sector. Values of N_z range from 6 lines at the 6 mm focal depth to 102 lines at 2 mm deep and 94 mm at 10 mm deep. Focusing delays were recalculated for each 0.5-mm depth band.

For the unapodized case, after the focusing delays were applied to the subaperture, the coherent sum was calculated. In order to shift the signal to baseband, the coherent sum was quadrature demodulated and low-pass filtered with a Hamming window centred around the DC component of the signal. For the apodized case, the coherent sum of the Hamming window weighted RF signals was calculated and the signals shifted to baseband as in the unapodized case.

For the minimum variance and generalized coherence factor weighting, focusing delays were applied, and delayed subaperture data were demodulated as described above, and then the weightings were computed. For a detailed explanation of GAFF, refer to [15]. In brief, one-dimensional Fourier transforms were calculated in the lateral

direction and the maximum projection of the Fourier transforms was calculated in the axial direction. From the axial projection, the ratio of energy around DC ($f < |M_0|$) of the spectrum to the total energy in the spectrum was calculated and multiplied by the coherent sum to obtain the weighted image. The Fourier transform is in the angular spatial frequency domain with DC being the component straight ahead of the centre of the subaperture.

The minimum variance technique is based on the methods presented in [12, 13]. The weighting for each scan line in the subaperture is calculated by minimizing the power of the beamformed signal, $E[|b(t)|^2]$, where $b(t)$ is the output of the beamformer, with the constraint of unity gain being maintained at the focal point. The weightings can be calculated analytically from the sample auto-covariance matrices of the subaperture data [13].

5.2.2 Experimental data - RF acquisition

All images were acquired with a Vevo 770 (VisualSonics Inc., Toronto, Canada) high-frequency ultrasound scanner using a 40 MHz centre frequency, single-element transducer with an f -number of 2. The resolution was $40 \times 80 \times 80 \mu\text{m}^3$ with a 1.5 mm depth of field (manufacturer's specifications) at the 6 mm geometric focal distance. The transducer is mechanically swept, acquiring 337 scan lines over a sector to create an 8 mm by 8 mm image. The radiofrequency (RF) data, required for the retrospective focusing, were digitized using an oscilloscope (Waverunner LT345, Lecroy Corp., Chestnut Ridge, NY, USA) and saved with software written in LabVIEW 6.1 (National Instruments Corp., Austin, TX, USA). All data were sampled at 250 MHz and digitized at 8 bit. In all phantom experiments, each scan line was acquired ten times to allow averaging to reduce the electronic noise. For the mouse imaging experiments each scan line was acquired three times to increase the speed of acquisition. The image acquisition time for one 2D image was approximately 5 minutes to acquire 10 repeated lines or 2 minutes to acquire 3 repeated lines. The software was set to

save a depth range of 2 mm to 10 mm, corresponding to the B-mode field of view. Two-dimensional images were reconstructed from the individual RF scan lines using Matlab (Mathworks, Natick, Massachusetts).

5.2.3 Phantom Experiments

In order to evaluate the synthetic aperture focusing techniques, phantoms were constructed from gelatin containing either thin air channels, to act as pseudo point targets, or tissue-mimicking spheres, to mimic the appearance of tumours on ultrasound images. Gelatin phantoms were made according to the procedure outlined by Ryan *et al.* [17], with 30% by weight gelatin powder, and a background of 2% by weight of amorphous silica. All phantoms were imaged at an angle of approximately 90°.

5.2.3.1 Air channels

A phantom mould was created that was comprised of a $15 \times 15 \times 15 \text{ mm}^3$ box with eight holes drilled in the sides to allow capillary tubes to be threaded parallel to each other and arranged on a diagonal with a 1 mm spacing in depth and 0.75 mm spacing horizontally. Glass capillary tubes with a 150 μm outer diameter (part number TSP030150, Polymicro Technologies, LLC, Phoenix, Arizona) were positioned through the guide holes and the tissue-mimicking material was poured into the mould. Once solidified, the tubes were carefully removed, leaving small channels filled with air. Ultrasound images of the cross-section of the channels were acquired with up to 7 channel cross-sections in an image.

The channel cross-sectional area was measured using a semi-automated region-growing algorithm that started from a user selected point to find the boundaries above the user selected threshold for the values surrounding the channel [18]. Once the entire region was found, the centroid of the area was calculated and used in

the analysis as the measured depth of the channel. To assess the measured area as a function of depth in the original and SAFT images, four 2D images of different planes of the phantom were acquired. Each channel in each image was measured three times and the averages of the measured cross-sectional areas and depths were computed. As each image was acquired at a separate time, the air channels are not at the exact same depth relative to the transducer in each case showing the measured areas over a range of around 0.3 mm.

5.2.3.2 Gelatin lesion phantoms and depth of field

Emulated lesions were created by embedding hypoechoic gelatin spheres with 0.5 % by weight amorphous silica into surrounding tissue-mimicking material with 2 % amorphous silica. A spherical mold was used to make the gelatin spheres. For each lesion phantom, 2D RF data were acquired with the lesion centred at the geometric focus, 1 mm above the focus and 1 mm below the focus. Each image was acquired at the varying depths to enable a comparison of the measurements as a function of location within the image field of view.

Sixty images were acquired of 21 phantom lesions, including 18 sets with an image of the lesion at all three depths. All images were reconstructed as described in section 5.2.2 and synthetic aperture focusing with each of the four weighting methods was applied. The magnitude of the images were taken followed by logarithmic compression. The images' window, level and gain for each 1 mm depth band were adjusted manually. Images were subsequently scan converted to create B-mode images. File names for each image were randomized and the 2D lesions were manually segmented three times each by a blinded observer using Matlab.

The coefficient of variation (COV; standard deviation divided by the mean) of the measured lesion area was calculated for each image. The mean and standard deviation of the COV was used to evaluate the measurement variability for single images. To compare the size-measurement variability as a function of location within

the field of view, for each set of three images, the maximum measured area minus the minimum measured area was used to give the range of measured areas. The mean and standard deviation of the measurement ranges were used to compare the variability.

A repeated measures ANOVA was performed on the data to determine the effects of repeated measurements, weighting method and location within the field of view. The data was analyzed considering the location within the image to be a within-subjects factor. A value of $\alpha < 0.05$ was considered significant for all tests.

5.2.4 Liver metastasis imaging

The syngeneic experimental liver metastasis model was employed (described in [1]). B16F1 murine melanoma cells were injected into the mesenteric vein to produce liver metastases. Five C57BL/6 mice with B16F1 liver metastases were imaged. Due to the rapid motion of the liver from respiratory and intestinal motion combined with the long acquisition times for the RF data, images were acquired immediately after sacrifice before obvious changes in the ultrasonic appearance of the tumour could be observed. To minimize the scan time, only three repeated scan lines were acquired at each location and averaged.

5.3 Results:

5.3.1 Phantom validation

5.3.1.1 Air channels

The diagonal pattern allowed for a visual comparison of the size and shape of the channel cross section as a function of depth. Figure 5.2 shows the original image and the identical image processed with SAFT and the different weighting techniques. Improvements in both the shape as well as cross sectional area can be seen. The cross-section of the air channels demonstrate the tight focusing at 6 mm and the

rapid decrease in lateral resolution away from this fixed focus, as can be seen in figure 5.2. After the application of the SAFT a clear improvement in focusing consistency can be seen, most noticeably in the cross-sections 1 mm above and below the fixed focus in figure 5.2.

Figure 5.3 displays the measured cross-sectional areas of the air channels for original and SAFT images as a function of image depth. All data sets show the same trend of increasing area away from the focal zone, with the original image showing much more rapid increases compared to the SAFT images. The smallest cross-sectional areas and corresponding depth at which they were measured are summarized in table 5.1, compared to an actual cross-sectional area of 0.018 mm^2 . It would be expected that the original image had the smallest cross-sectional area closest to 6 mm; however, the SAFT images with minimum variance and apodized weighting show improved air channel measurement even at the fixed focus. The SAFT images show improvement in the air channel cross-sectional area beyond the minimum in the original images.

Weighting method	Minimum cross-sectional area (mm^2)	Image depth (mm)
Actual cross-sectional area	1.77×10^{-2}	
Original image	4.51×10^{-2}	6.36
Unapodized	4.31×10^{-2}	5.39
Hamming	3.90×10^{-2}	6.21
Minimum variance	2.67×10^{-2}	6.32
Generalized coherence factor	2.69×10^{-2}	5.38

Table 5.1: Minimum measured cross-sectional area of $150\text{-}\mu\text{m}$ diameter air channels for each SAFT weighting method.

All SAFT weighting techniques showed an improvement in the cross-sectional area measured over the original images. The unapodized image showed the smallest improvement, followed by the apodized results. The GAFF and MV images showed greater improvement in most cases; however, the focusing was less consistent as a function of depth than the non-adaptive weighting techniques, as can be seen in figure 5.3.

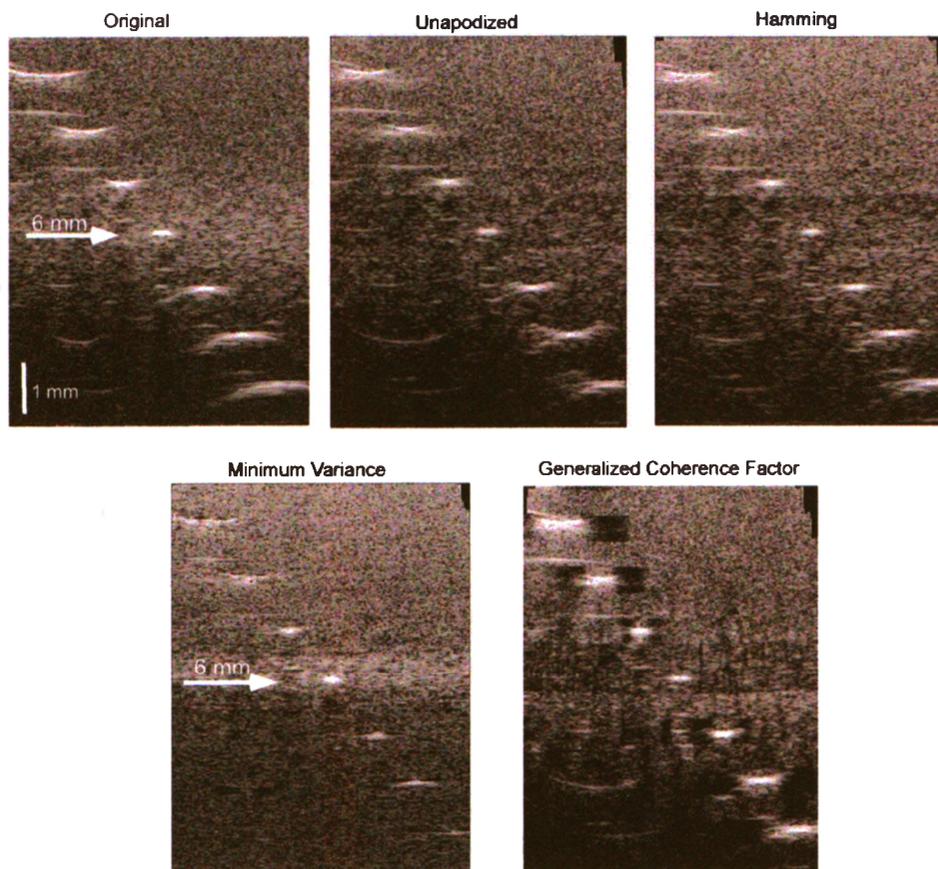


Fig. 5.2: An image showing the $150\ \mu\text{m}$ air channels in cross-section with the original data and with synthetic aperture focusing with each of the four weighting methods. The arrow indicates the 6 mm geometric focus and the vertical scale bar is 1 mm. Near the geometric focus, all weighting methods appear to reduce the width of the air channel. The minimum variance weighting improves the cross-sectional area in this region but is lower contrast than the other methods whereas the generalized coherence factor weighting results in higher contrast but introduces additional artifacts into the image. A 40 dB dynamic range is displayed.

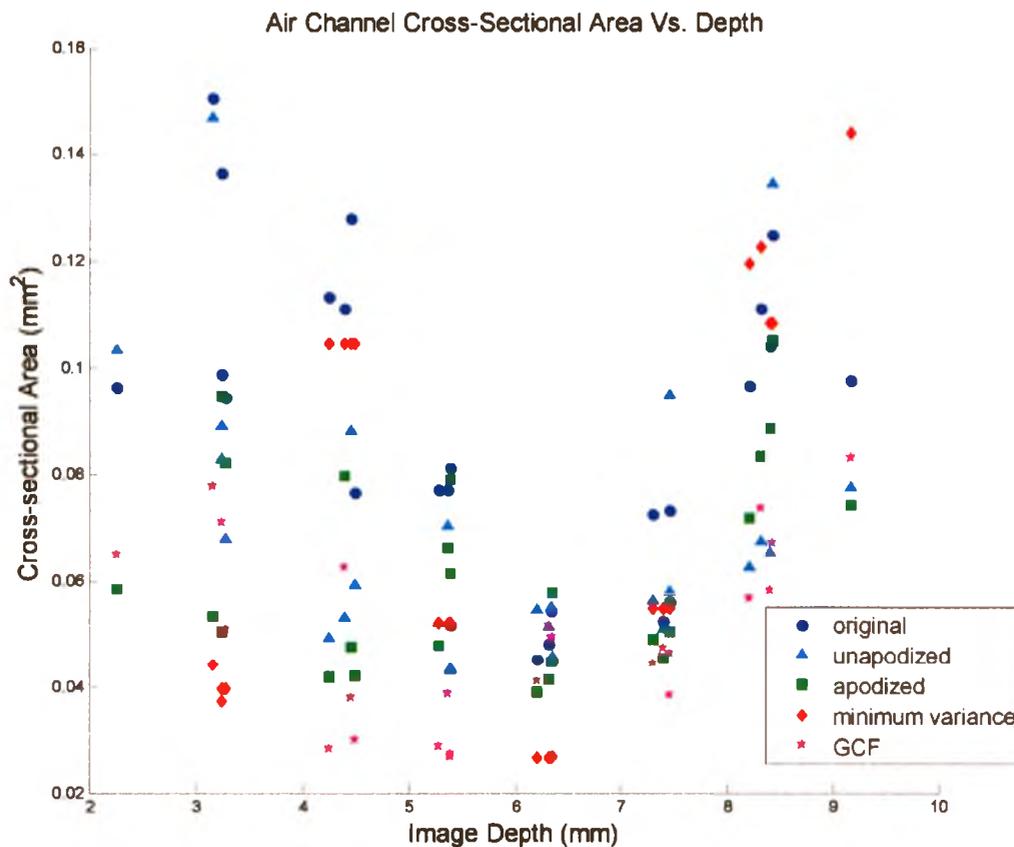


Fig. 5.3: Average measured cross-sectional area of the air channels is plotted against image depth. Data from four images of the phantom are presented and have been measured after each of the four weighting techniques have been applied. The GAFF (generalized coherence factor) shows the most consistent cross-sectional area with a large improvement over the original images over the full range of depth. Minimum variance shows improvement around the geometric focus, but begins to increase in cross-sectional area within about 2 mm of the focus. The unapodized and apodized weighting shows some improvement, particularly near the geometric focus, but does not maintain a consistently low measurement across all depths.

5.3.1.2 Gelatin lesion phantoms and depth of field

The average lesion area measured on the original image was 4.88 mm^2 , with maximum diameters (measured in 3D ultrasound) ranging from 1.94 mm to 3.45 mm with an average of 2.52 mm. Visual inspection shows some improvement in margin definition for the unapodized and Hamming window weighting (figure 5.4). The GAFF weighting shows substantial improvement in boundary delineation in some areas, with other areas being degraded due to artifacts introduced from bright objects near the boundary (figure 5.5). The MV weighting also shows improvement in the boundary delineation, most noticeable near and beyond the fixed focus, but lesions situated between the transducer and the fixed focus do not show improved margins. Overall, the MV weighting suffers from fewer image artifacts, even in the presence of extraneous bright areas, compared to the GAFF (figure 5.5).

Table 5.2 shows the measurement variability for individual lesion measurements and the range of measurements for the lesion at multiple depths. The average COV ranges from 2.60 % for the apodized weighted up to 3.19 % for the minimum variance weighting, with the original images having an average COV of 3.00 %. Although there is a range of values, there are no significant differences between weighting methods for the COV ($\alpha > 0.05$). The same pattern is also seen for the range of measured areas for each lesion, where the Hamming and unapodized weighting provided modest improvement over the original images but the two adaptive techniques did not.

The repeated-measures ANOVA showed a significant interaction between the repeats and the weighting technique and a significant main effect for the weighting technique, indicating that the measured volume does change based on the different SAFT weighting techniques. Depending on how well defined the margins are, the observer will vary where they place the segmented boundary. No significant effects were seen due to lesion location within the image, indicating that there is not significant variation in the measured size as a function of image depth.

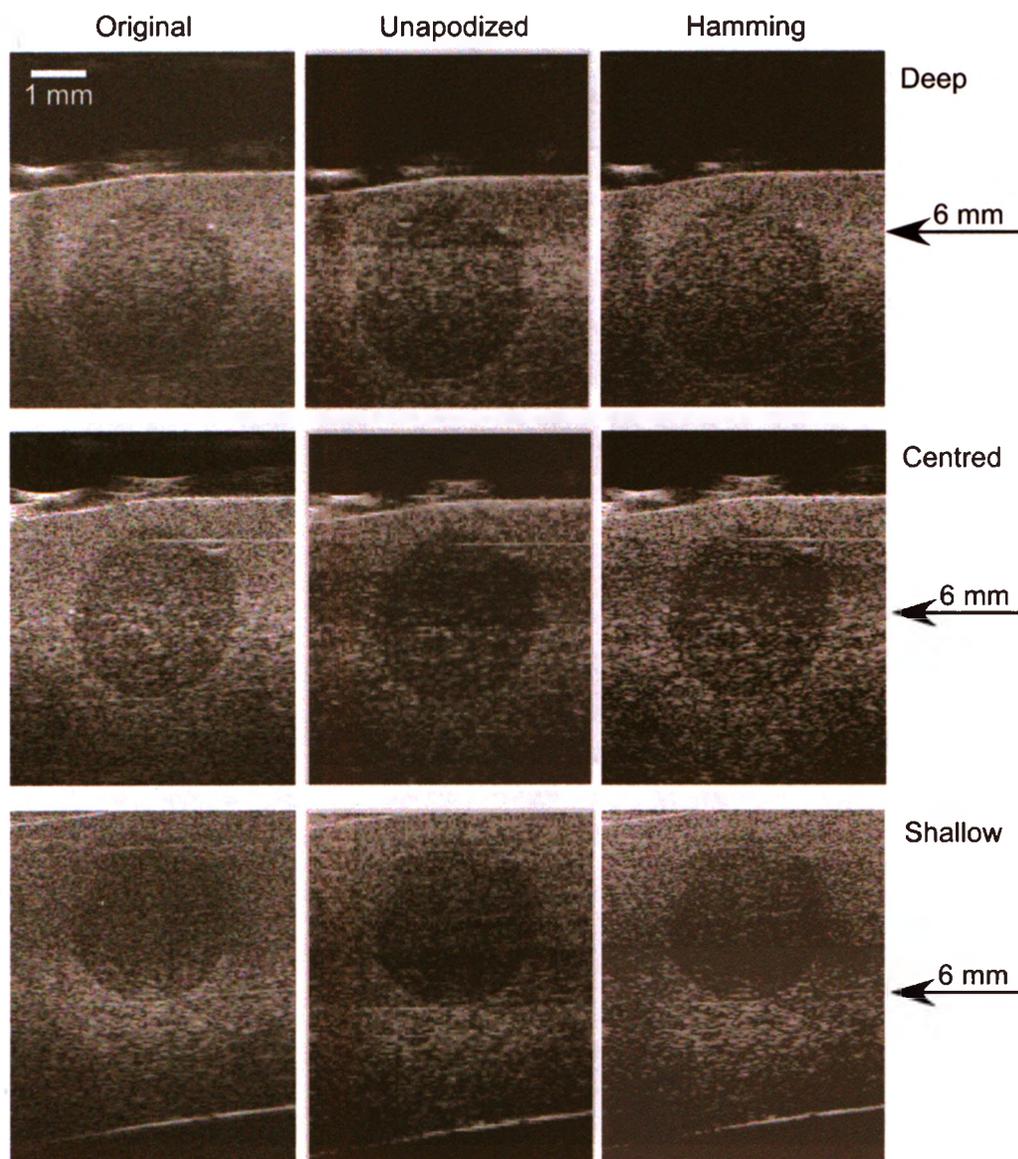


Fig. 5.4: Original images of a lesion phantom are shown compared to SAFT with unapodized and Hamming window weighting for three depths of the region of interest. Increased contrast can be seen at the lesion border for the SAFT images, particularly for the Hamming window weighting where there is a good contrast at the bottom of the deep lesion compared to the original image. A 40 dB dynamic range is displayed.

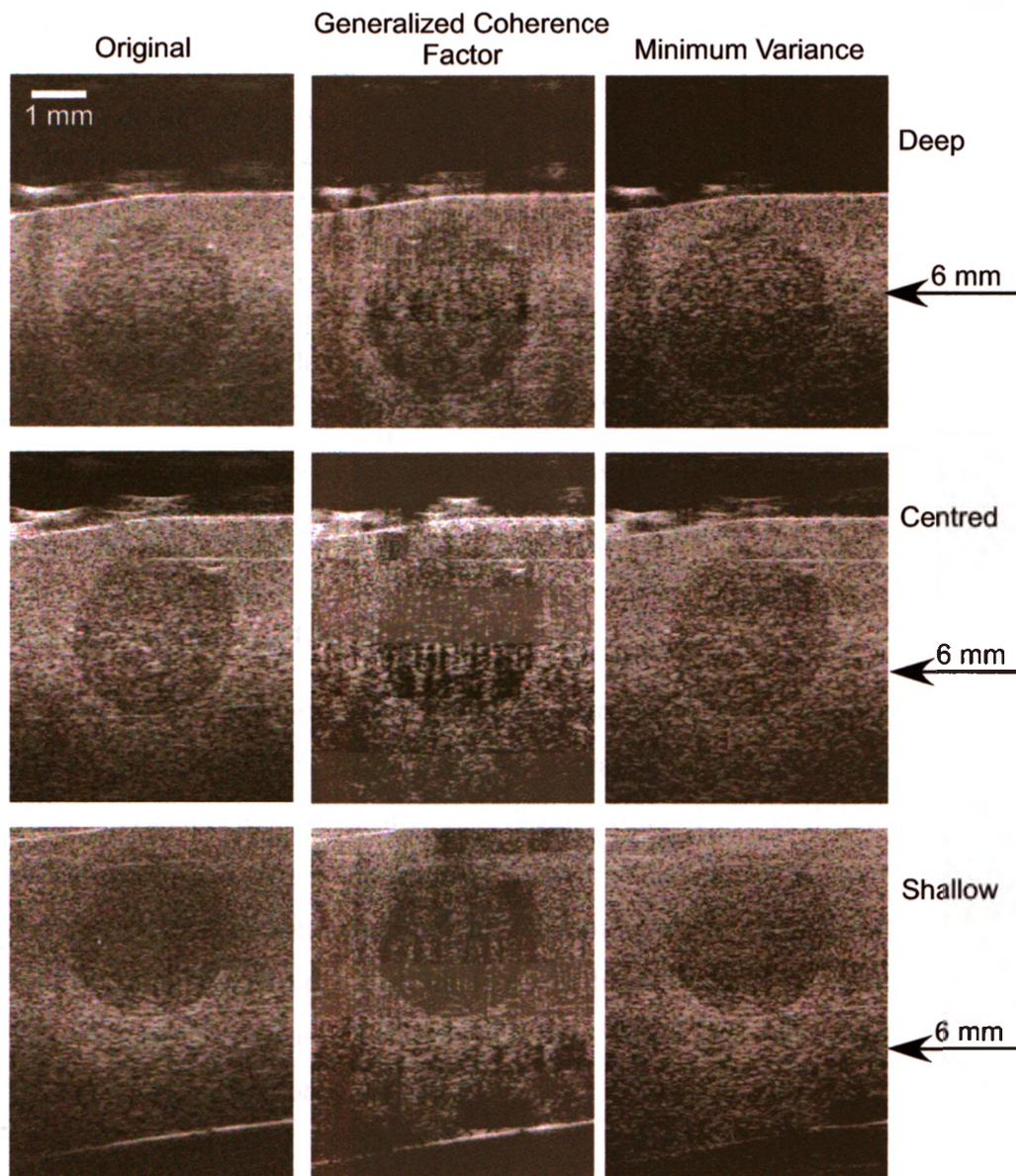


Fig. 5.5: Original images of a lesion phantom are shown compared to SAFT with generalized coherence factor and minimum variance adaptive weighting for three depths of the region of interest. The generalized coherence factor images show improved contrast, especially in the shallow image, however, some artifacts that are introduced can be seen in the centred image which can make the image interpretation difficult. The minimum variance method shows good contrast improvements for the deep and centred lesion; however, for the shallow case no improvement is seen. A 40 dB dynamic range is displayed.

Weighting method	Mean COV (%)	STD COV (%)	Mean range (mm ²)	STD range (mm ²)
Original	3.00	1.81	0.54	0.43
Unapodized	3.02	2.19	0.46	0.33
Hamming	2.60	1.53	0.50	0.27
MV	3.19	2.68	0.57	0.33
GAFF	3.14	2.49	0.56	0.39

Table 5.2: The table shows the results from the repeated measurements of the phantom lesion areas. The mean and standard deviation (STD) of the coefficient of variation (COV) is recorded as an indication of measurement variability. In addition, the mean and standard deviation of the range of area measurements for a lesion at multiple image depths is given. The values for the Hamming weighting show improvements over the original, uncorrected images. However, the minimum variance (MV) and generalized coherence factor (GAFF) images show degraded measurement variability. The unapodized shows almost equivalent COV and improved mean range.

5.3.2 Liver metastasis imaging

Due to the small sample size and limitations in image quality due to the low number of repeated line acquisitions for averaging, liver metastases are shown as a proof of principle for tumour imaging. A small tumour within the centre of the liver is presented. Visual inspection showed improved tumour to background contrast for the conventional weighting. The minimum variance shows a slight increase in contrast, whereas the generalized coherence factor reduces contrast for most of the tumour. The changes in appearance for the different weightings follow visually the results from the lesion phantoms.

5.4 Discussion:

5.4.1 Air channels

Measurements of the cross-sectional area of the small air channels gives an approximation to a point target and allows the approximate shape of the point spread function to be seen within the image. The minimum cross-sectional area due to SAFT

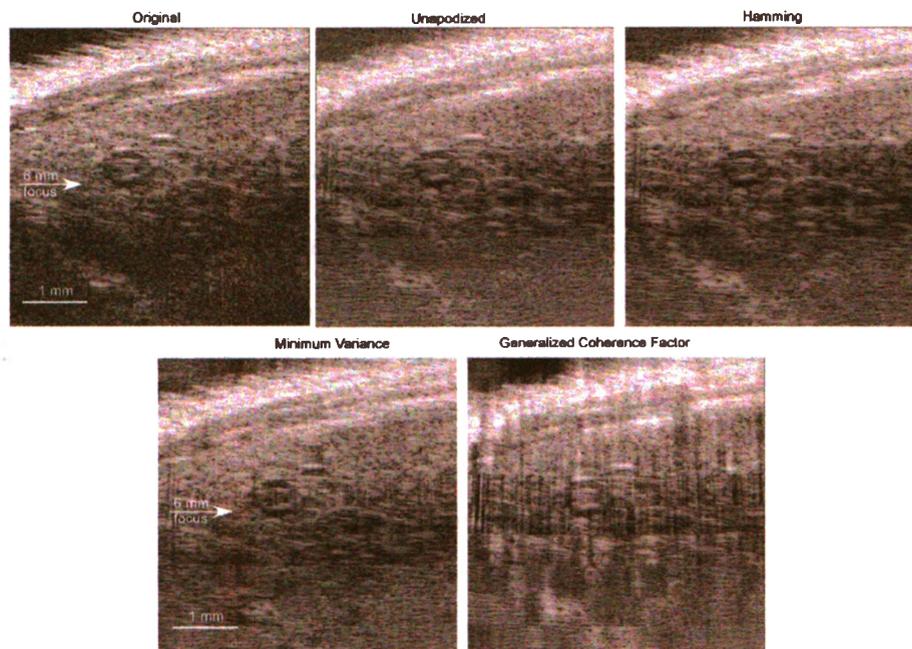


Fig. 5.6: An example liver metastasis is shown in the original image and for the different weighting techniques. The unapodized and Hamming weighting show an improvement in contrast between the tumour boundary and the liver. The minimum variance shows high contrast for part of the tumour boundary. The generalized coherence factor makes the tumour very difficult to see within the image. A 40 dB dynamic range is displayed.

was shallower than the 6 mm geometric focus for the GAFF weighted and unapodized images, the limited number of scan lines used to reconstruct the image at the geometric focus likely contributes to greater improvement in mainlobe width. Difficulty in measuring the cross-sectional areas for the MV weighted images likely contributed to the high measured values at shallow and deep points. The contrast between the air channel and surrounding material was poor and the region growing algorithm tended to encompass a relatively large region surrounding the centre of the air channel.

5.4.2 Gelatin lesion phantoms

The improved measurements from the unapodized to Hamming to adaptive weighting techniques is not surprising. However, due to the complex nature of the adaptive weighting algorithms and the unique artifacts introduced into the images, they offer less consistent improvements than expected. Both techniques aim to minimize the contribution of off-axis scatterers to the SAFT image; however, within the context of the lesion images this seemed to create regions with very low weighting next to regions with bright speckle. While the variability, expressed as the COV, is very similar between all groups it is very interesting to note that the adaptive techniques performed worse based this parameter compared to the non-adaptive techniques.

5.4.3 Liver metastasis imaging

Due to limitations in the number of acquisitions, the original images are not as high of a quality as the lesion phantoms images. When the SAFT is applied, the unapodized and Hamming weights show improved contrast which could help delineate the borders for segmenting. The generalized coherence factor degrades some of the image quality making it difficult to define borders.

5.5 Conclusion:

For point targets, there are definite benefits to using any of the weighting techniques. The GAFF weighting consistently maintained a reduced area measurement compared to the original image. The MV weighting, in contrast, performed very inconsistently.

For the phantom and *in vivo* lesion images, the results are harder to interpret. The adaptive techniques did not show the improvement in measurement variability that was expected to result from improved resolution. For lesion images, it is likely best to use an unapodized or Hamming weighting for synthetic aperture focusing. These methods gave the greatest benefits in terms of measurement variability and are less computationally demanding than the adaptive techniques.

References

- [1] K. C. Graham, L. A. Wirtzfeld, L. T. MacKenzie, C. O. Postenka, A. C. Groom, I. C. MacDonald, A. Fenster, J. C. Lacefield, and A. F. Chambers, "Three-dimensional high-frequency ultrasound imaging for longitudinal evaluation of liver metastases in preclinical models," *Cancer Res*, vol. 65, no. 12, pp. 5231-7, 2005.
- [2] Ai-Ho Liao, Chen-Han Li, Weng-Fang Cheng, and Pai-Chi Li, "Non-invasive imaging of small-animal tumors: High-frequency ultrasound vs. microPET," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2005, pp. 5695-8.
- [3] A. M. Cheung, A. S. Brown, L. A. Hastie, V. Cucevic, M. Roy, J. C. Lacefield, A. Fenster, and F. S. Foster, "Three-dimensional ultrasound biomicroscopy for xenograft growth analysis," *Ultrasound Med Biol*, vol. 31, no. 6, pp. 865-70, 2005.
- [4] L. A. Wirtzfeld, K. C. Graham, A. C. Groom, I. C. Macdonald, A. F. Chambers, A. Fenster, and J. C. Lacefield, "Volume measurement variability in three-dimensional high-frequency ultrasound images of murine liver metastases," *Phys Med Biol*, vol. 51, no. 10, pp. 2367-81, 2006.
- [5] J. A. Brown and G. R. Lockwood, "A digital beamformer for high-frequency annular arrays," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 52, no. 8, pp. 1262-1269, 2005.
- [6] J. A. Ketterling, S. Ramachandran, and O. Aristizabal, "Operational verification of a 40-MHz annular array transducer," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 53, no. 3, pp. 623-630, 2006.

- [7] E. J. Gottlieb, J. M. Cannata, Hu Chang-Hong, and K. K. Shung, "Development of a high-frequency (> 50 MHz) copolymer annular-array, ultrasound transducer," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 53, no. 5, pp. 1037–1045, 2006.
- [8] J. M. Cannata, J. A. Williams, Zhou Qifa, T. A. Ritter, and K. K. Shung, "Development of a 35-MHz piezo-composite ultrasound array for medical imaging," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 53, no. 1, pp. 224–236, 2006.
- [9] M. Lukacs, Yin Jianhua, Pang Guofeng, R. C. Garcia, E. Cherin, R. Williams, J. Mehi, and F. S. Foster, "Performance and characterization of new micromachined high-frequency linear arrays," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 53, no. 10, pp. 1719–29, 2006.
- [10] Mustafa Karaman, Pai-Chi Li, and Matthew O'Donnell, "Synthetic aperture imaging for small scale systems," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 42, no. 3, pp. 419–442, 1995.
- [11] C. H. Frazier and Jr. O'Brien, W. D., "Synthetic aperture techniques with a virtual source element," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 45, no. 1, pp. 196–207, 1998.
- [12] Suhyun Park, Andrei B. Karpiouk, Salavat R. Aglyamov, and Stanislav Y. Emelianov, "Adaptive beamforming for photoacoustic imaging," *Optics Letters*, vol. 33, no. 12, pp. 1291–1293, 2008.
- [13] Johan-Fredrik Synnevag, Andreas Austeng, and Sverre Holm, "Adaptive beamforming applied to medical ultrasound imaging," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 54, no. 8, pp. 1606–1613, 2007.

- [14] P. C. Li and M. L. Li, "Adaptive imaging using the generalized coherence factor," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 50, no. 2, pp. 128-41, 2003.
- [15] M. L. Li, W. J. Guan, and P. C. Li, "Improved synthetic aperture focusing technique with applications in high-frequency ultrasound imaging," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 51, no. 1, pp. 63-70, 2004.
- [16] J. A. Jensen, "Field: A program for simulating ultrasound systems," *Med. Biol. Eng. Comp.*, vol. 4, Supplement 1, no. Part 1, pp. 351-353, 1996.
- [17] Linda K. Ryan and F. Stuart Foster, "Tissue equivalent vessel phantoms for intravascular ultrasound," *Ultrasound in Medicine & Biology*, vol. 23, no. 2, pp. 261-273, 1997.
- [18] M. Y. Wang, Jr. Maurer, C. R., J. M. Fitzpatrick, and R. J. Maciunas, "An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head," *IEEE Trans Biomed Eng.*, vol. 43, no. 6, pp. 627-37, 1996.

Chapter 6

Summary and Future work

6.1 Summary

Three-dimensional high-frequency ultrasound can be a valuable tool for providing information on the growth of soft-tissue tumours in mouse models to complement conventional endpoint analysis. The lack of agreement between conventional caliper measurements and 3D ultrasound volume measurements suggests that the calipers may only provide an approximate measure of size and lack the precision to draw conclusions regarding the growth of the tumours or allow for comparisons between tumours and treatment groups.

6.1.1 Chapter 2: Three-Dimensional High-Frequency Ultrasound Imaging For Longitudinal Evaluation Of Liver Metastases In Preclinical Models

Chapter 2 demonstrates that liver metastases can be longitudinally imaged without exogenous contrast agents to construct growth curves for individual tumours. In addition to the work presented in chapter 2, a similar study examining prostate tumour growth in a transgenic mouse model was performed [1]. Prostate tumours

spontaneously developed over time and were detected through regular ultrasound screening and subsequently monitored over time to measure tumour growth. To further show the usefulness of high-frequency ultrasound for imaging soft-tissue tumours using endogenous contrast, a study of the growth of mouse melanoma [2] and unpublished studies of mouse mammary fat pad tumours have been conducted. All cell lines have shown sufficient endogenous contrast to allow tumours to be monitored over time to construct volume growth curves.

The ability to track individual tumours growth over time in a precise manner will allow the tumour to act as its own control as any changes in growth due to treatment response can be quantified for each tumour. Tracking individual tumour's opens the possibility of being able to evaluate different tumours response per animal and even per tumour. The longitudinal experiments with the different cell lines in chapter 2 demonstrate that individual tumours can grow at varying rates from each other. The single HT-29 tumour that showed no growth over the course of the experiment opens the possibility of using ultrasound to identify small but dormant tumours. Whether this tumour would respond to a therapeutic agent or at some point begin to grow is unknown. The variability in growth makes it difficult to interpret endpoint analysis, as a tumour that did not grow for the duration of the experiment cannot be differentiated from a tumour that grew rapidly to a larger size and then regressed due to an effective treatment.

6.1.2 Chapter 3: Volume Measurement Variability In Three-Dimensional High-Frequency Ultrasound Images Of Murine Liver Metastases

Tumour measurement variability in three-dimensional high-frequency ultrasound images due to segmentation variability by observers and experimental factors has been quantified. Experimental B16F1 liver metastases were analyzed in different size

ranges including less than 1 mm³, 1 to 4 mm³, 4 to 8 mm³ and 8 to 70 mm³. The intra- and inter-observer repeatability was high over a large range of tumour volumes, but the coefficient of variation (COV) varied over the volume ranges. The minimum and maximum intra-observer COV were 4% and 14% for the 1 to 4 mm³ and < 1 mm³ tumours, respectively. For tumour volumes measured by segmenting parallel planes, the maximum inter-slice distance that maintained acceptable measurement variability increased from 100 to 600 μm as tumour volume increased. Comparison of free breathing versus ventilated animals demonstrated that respiratory motion did not significantly change the measured volume. These results enable design of more efficient imaging studies by using the measured variability to estimate the time required to observe a significant change in tumour volume.

6.1.3 Chapter 4: Monte Carlo Growth Curve Simulations for Planning Longitudinal Imaging Experiments with Mouse Cancer Models

Growth curve simulations, based on the growth of the B16F1 liver metastases and their measurement variability, were performed to identify longitudinal experiment designs for which it should be more appropriate to analyse tumour volume data using a Gompertzian growth curve than an exponential growth curve. Imaging tumours beginning when they are small (< 0.09 mm³) and imaging over a large volume range support the use of Gompertzian growth analysis. Volume data needs to be acquired over a minimum of 15 days and over as many as 40 days for infrequent imaging or when it is desired to determine if the two data sets are growing with different parameters. In most cases, an imaging interval of 6 days does not provide sufficient data to perform the analyses with Gompertz functions and decreasing the interval to 4 days shows an improvement in the fitting results.

6.1.4 Chapter 5: Synthetic Aperture Focusing of a Single Element 40 MHz System

Synthetic aperture focusing techniques (SAFT) with a variety of weighting methods, including both conventional and adaptive, were applied to images of both point-like targets and lesions. Reductions in the measured cross-sectional areas of the point-like targets, indicating improved spatial resolution, was seen for all weighting methods, although the generalized coherence factor offered the most consistently improved resolution over all depths. Synthetic aperture focusing was less effective than expected for reducing the area measurement variability, whereas the conventional apodization showed a small improvement. Although measurement variability on an individual lesion basis did not show the expected improvements using SAFT, the variation in the measured lesion size as a function of depth was reduced with SAFT, suggesting a reduction in the spatial variance of the point-spread function.

6.1.5 Current Progress

Current developments include the use of clinical scale micro-bubble (1 to 5 μm) and newer liquid-core nanoparticle (anywhere from 100 to 1000 nm) contrast agents to enhance the B-mode imaging. The growth curve analysis indicates that it is beneficial to be able to detect and measure smaller tumour volumes when fitting Gompertz functions to the growth data and comparing curves. The use of contrast agents may allow for small tumours to be detected prior to them developing sufficient contrast to be detected and measured with only endogenous contrast. The lack of perfusion or limited perfusion in small tumours would provide negative contrast against the contrast-enhanced liver parenchyma, making the identification of even small tumours easier than the unenhanced tumours.

Contrast agents are frequently also used to assess blood flow within organs and tumours. Changes in blood flow within the tumours could provide information that

would complement the volume measurements and perhaps point to mechanisms of treatment action. Changes in the surrounding blood flow to the tumour may also be of interest as the feeder vessels may offer a therapeutic target.

Assessing blood flow can also be performed using either colour flow or power Doppler ultrasound techniques that can be used with or without contrast agents. Several groups have demonstrated the use of power Doppler to quantify tumour blood flow, including using power Doppler to monitor the tumour response to antiangiogenic drugs [3], antivascular drugs [4] and surgical interventions [5]. Work has been done to quantify Doppler techniques and compare them to other gold standards including histological microscopic quantification of tumour blood vessels [6] and microfil-enhanced CT [5].

6.2 Future Work

VisualSonics has recently released their new Vevo 2100 scanner including linear array transducers. The use of arrays could have benefits in performing longitudinal experiments imaging tumour models, as they would allow a much larger range of depths to be imaged in focus and thus reduce measurement variability. To determine the effect of the array on measurement variability, the study presented in chapter 3 would need to be repeated using images acquired with this new system. Any changes in measurement variability could then be used in the growth simulations, as presented in chapter 4, to allow for comparison of how the experiments should be designed to take into consideration the tumour-volume measurement variability differences between single-element and linear array scanners.

Using the results from the growth curve simulations in chapter 4, an *in vivo* experiment should be set up to obtain sufficient data for reliable growth curve fitting at the beginning of the tumour growth and have additional measurements acquired beyond the minimum time points required. The first set of time points could be

used to fit a Gompertz function and then extrapolate the curve to the additional data to determine how well the fitted curve can predict how the tumour will continue to grow. The ability to extrapolate late stage growth from early stage volume data would support treatment-response experiments in which each tumour is used as its own control, i.e., measured growth after treatment is compared to predicted growth extrapolated from early stage, pre-treatment volume data.

Synthetic aperture focusing techniques, in addition to influencing the tumour-size measurement variability, will likely affect the detectability of smaller lesions. The minimum lesion size that can be observed in an image is a function of the lateral resolution and side-lobe levels, therefore improvements in focusing through SAFT could allow smaller tumours to be observed or the smallest tumours to be observed regardless of their location within the image. Due to the tightly focused transducer on the current system, it is possible that small tumours that can be detected when they are within the focal zone may not be detectable at other image depths. A phantom based contrast-detail experiment, similar to the one described by Smith *et al.* [7], could be employed to determine the smallest cross-sectional area that can be observed as a function of the lesion contrast. Two-dimensional assessment would be appropriate as initial scanning to find the tumours is performed in 2D and only once a tumour is found is a 3D image acquired.

6.3 Conclusions

High-frequency ultrasound is able to provide precise longitudinal analysis of liver metastasis growth in mouse models over a large range of tumour volumes. Retrospective focusing is able to improve these measurements through a reduction in the dependence of the volume measurement on the location within the field of view. Experiments can be designed to ensure sufficient data points are acquired to perform the desired curve fitting.

The ability to track these soft-tissue tumours through time and measure them with low variability allows for the construction of growth curves. These growth curves can be analyzed using curve fitting to allow multiple curves to be compared from within the same or different treatment groups in a study, thereby providing information that could be valuable in assessing therapeutic effects.

References

- [1] L. A. Wirtzfeld, G. Wu, M. Bygrave, Y. Yamasaki, H. Sakai, M. Moussa, J. I. Izawa, D. B. Downey, N. M. Greenberg, A. Fenster, J. W. Xuan, and J. C. Lacefield, "A new three-dimensional ultrasound microimaging technology for preclinical studies using a transgenic prostate cancer mouse model," *Cancer Res*, vol. 65, no. 14, pp. 6337-45, 2005.
- [2] A. M. Cheung, A. S. Brown, L. A. Hastie, V. Cucevic, M. Roy, J. C. Lacefield, A. Fenster, and F. S. Foster, "Three-dimensional ultrasound biomicroscopy for xenograft growth analysis," *Ultrasound Med Biol*, vol. 31, no. 6, pp. 865-70, 2005.
- [3] M. Jugold, M. Palmowski, J. Huppert, E. Woenne, M. Mueller, W. Semmler, and F. Kiessling, "Volumetric high-frequency doppler ultrasound enables the assessment of early antiangiogenic therapy effects on tumor xenografts in nude mice," *European Radiology*, vol. 18, no. 4, pp. 753-8, 2008.
- [4] D. E. Goertz, J. L. Yu, R. S. Kerbel, P. N. Burns, and F. S. Foster, "High-frequency doppler ultrasound monitors the effects of antivascular therapy on tumor blood flow," *Cancer Res*, vol. 62, no. 22, pp. 6371-5, 2002.
- [5] J. W. Xuan, M. Bygrave, H. Jiang, F. Valiyeva, J. Dunmore-Buyze, D. W. Holdsworth, J. I. Izawa, G. Bauman, M. Moussa, S. F. Winter, N. M. Greenberg, J. L. Chin, M. Drangova, A. Fenster, and J. C. Lacefield, "Functional neoangiogenesis imaging of genetically engineered mouse prostate cancer using three-dimensional power doppler ultrasound," *Cancer Res*, vol. 67, no. 6, pp. 2830-9, 2007.
- [6] E. F. Donnelly, L. Geng, W. E. Wojcicki, A. C. Fleischer, and D. E. Hallahan, "Quantified power doppler us of tumor blood flow correlates with microscopic quantification of tumor blood vessels," *Radiology*, vol. 219, no. 1, pp. 166-70, 2001.

- [7] S. W. Smith, R. F. Wagner, J. M. Sandrik, and H. Lopez, "Low contrast detectability and contrast/detail analysis in medical ultrasound," *Sonics and Ultrasonics, IEEE Transactions on*, vol. 30, no. 3, pp. 164-173, 1983.

Appendix A

Copyright agreements

This thesis contains material from previously published manuscripts in *Cancer Research** (chapter 2) and *Physics in Medicine and Biology*† (chapter 3). *Cancer Research* and *Physics in Medicine and Biology* do not require that the author of a publication receive copyright permission for duplication of the published article for the purpose of a thesis or dissertation, as long as the source of the material is properly referenced. Appropriate reference to the published sources is given in the co-authorship section.

*Details available at: <http://cancerres.aacrjournals.org/misc/ifora.shtml#1>

†Details available at: <http://authors.iop.org/atom/help.nsf/aF20EC7D4A1A670AA8025-6F1C0053EEFF?OpenDocument>

Appendix B

Ethics Approval for Animal Subjects



10.30.06

This is the 1st Renewal of this protocol***A Full Protocol submission will be required in 09.30.09***

Dear Dr. Chambers & Dr. MacDonald:

Your Animal Use Protocol form entitled:

Non-Invasive imaging of Metastasis: Detection, Monitoring and Intervention

has been approved by the Animal Use Subcommittee.

This approval is valid from **10.01.06 to 09.30.09**.The protocol number for this project remains as **2005-052-09**.

1. This number must be indicated when ordering animals for this project.
2. Animals for other projects may not be ordered under this number.
3. If no number appears please contact this office when grant approval is received.
If the application for funding is not successful and you wish to proceed with the project, request that an internal scientific peer review be performed by the Animal Use Subcommittee office.
4. Purchases of animals other than through this system must be cleared through the ACVS office. Health certificates will be required.

ANIMALS APPROVED FOR 1 YR. - Highest Pain Level: D

Species	Other Detail	Housing/Use Locations	Animal # Total for 1 Year
Mouse	Nudes 5-6 wks on arrival	HSACF	140
Mouse	NIH 111 5-6 wks on arrival	HSACF	80
Mouse	BalbC 5-6 wks on arrival	HSACF	80
Mouse	CS7 Blik 5-6 wks on arrival	HSACF	80
Mouse	SCID 5-6 wks on arrival	HSACF	80

REQUIREMENTS/COMMENTS

Please ensure that individual(s) performing procedures on live animals, as described in this protocol, are familiar with the contents of this document.

c.c. Approved Protocol - A. Chambers, I. MacDonald, L. MacKenzie, T. Kirkpatrick.
Approval Letter - L. MacKenzie.

The University of Western Ontario
Animal Use Subcommittee/University Council on Animal Care
Health Sciences Centre • London, Ontario • CANADA • N6A 5C1
Phone: 519-661-2111 ext. 86770 • Fax: 519-661-2028 • www.uwo.ca/animal