American College of Radiology

Annual Progress Report: 2011 Formula Grant

Reporting Period

July 1, 2014 – June 30, 2015

Formula Grant Overview

The American College of Radiology received $1,777,126 in formula funds for the grant award period January 1, 2012 through December 31, 2015. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

Evaluation of Biomarker Focused Projects – The Radiation Therapy Oncology Group (RTOG), a National Cancer Institute funded multi-institutional clinical cooperative group has been collecting and banking biospecimens (biopsies, blood, urine, etc.) from patients enrolled on its clinical trials for decades. Often these specimens are collected without a pre-identified analysis – they are "banked" for future use. As technology and new biomarkers are developed, investigators request permission to use the specimens for research to identify new biomarkers or validate new procedures. These "secondary" analyses are not required by the original protocol, and may not be funded as part of that protocol. This project will allow for the investigation, including the statistical analysis, of five specified biomarker focused projects.

Anticipated Duration of Project

1/1/2012 – 12/31/2015

Project Overview

This project aims to use clinical data and biomarkers assessed from tissue specimens that have been collected in previous RTOG studies to advance current knowledge regarding the treatment and prognosis of cancer patients. The specific research objectives of this project relate to five project aims that will contribute to the overall project.

Aim 1: Evaluation of Candidate Pathways of Therapeutic Intervention in Anal Cancer: RTOG 9811 is a Phase III trial of patients with carcinoma of the anal canal. Using data and samples from this trial, this project will correlate expression of ERCC1, p53, p16 and PTEN via AQUA® technology, protein co-expression of EGFR and Ki-67 and amplification of EGFR with clinical outcome factors.

Note: This project is funded, in part, under a grant with the Pennsylvania Department of Health. The Department specifically disclaims responsibility for any analyses, interpretations or conclusions.

Aim 2: Validating cytoplasmic Hu antigen R (HuR) expression as a marker for gemcitabine (Gem) response in pancreatic cancer patients: RTOG 9704 is a Phase III trial of patients with resected pancreatic cancer. Using this trial's data and samples, this project will evaluate HuR cytoplasmic expression as an independent predictor of response to Gemcitabine treatment. This project will also evaluate correlations between deoxycytidine kinase (dCK) and HuR expression.

Aim 3: Correlation of Soft Tissue Sarcoma Tissue Biomarker Expression Patterns with Treatment Response and Outcomes: RTOG 9514 is a Phase II Study of Neoadjuvant Chemo and Radiation Therapy in the Management of High-Risk, High-Grade, Soft Tissue Sarcomas of the Extremities and Body Wall. Using this trial's data and samples, expression patterns for "candidate" tumor biomarkers will be defined and subsequently correlated with clinically relevant outcomes.

Aim 4: Markers and Potential Therapeutic Targets for Improving Tumor Response in Head & Neck (H&N) Cancer: This project will use data from two Phase III trials for locally advanced H&N cancer: RTOG 0129 and 0522. A separate grant, using specimens/data from non-RTOG trials, will develop candidate DNA methylation and mi-RNA biomarkers of response to treatment and prioritize them for clinical validation. This Aim 4 project will validate and refine the selected signatures of therapeutic response, using the above mentioned RTOG trials.

Aim 5: Correlating Pathologic Variables with Outcomes in Patients with Non-Urothelial Muscle Invasive Bladder Cancer After Bladder Preserving Trimodality Therapy: The final project utilizes multiple RTOG bladder sparing trials, focusing specifically on those patients with variant histologies. This project will focus on correlating central pathologic review data for these patients with long-term outcomes.

Principal Investigator

Kathryn Winter, MS
Director, RTOG Statistical and Data Management Center
American College of Radiology
Radiation Therapy Oncology Group
1818 Market Street, Suite 1600
Philadelphia, PA 19103

Other Participating Researchers

Qiang (Ed) Zhang, PhD; Jonathan Harris, MS; Jennifer Moughan, MS; Yan Yan, MD, MS; Asha George, MS – employed by the American College of Radiology

Expected Research Outcomes and Benefits

Aim 1: Evaluation of Candidate Pathways of Therapeutic Intervention in Anal Cancer: Identification of designated pathways and accurate measurement of designated proteins using

standardized and quantitative technologies will enable improved patient selection for treatment and provide possible targets for therapeutic interventions.

Aim 2: Validating cytoplasmic Hu antigen R (HuR) expression as a marker for gemcitabine (Gem) response in pancreatic cancer patients: A strong relationship between HuR and survival after Gem treatment may provide a foundation to develop strategies for tailoring or intensifying Gem-based therapeutic regimens, including the potential to direct clinical trial eligibility.

Aim 3: Correlation of Soft Tissue Sarcoma Tissue Biomarker Expression Patterns with Treatment Response and Outcomes: This project may aid in prognostication and guide trial design for this relatively rare tumor, where a limited number of potential patients makes every clinical trial extremely valuable as a way to improve the treatment of this disease.

Aim 4: Markers and Potential Therapeutic Targets for Improving Tumor Response in Head & Neck Cancer: This project may serve to streamline cancer therapy by logical selection of tumor-specific treatment. Such individualization of therapy may result in increased efficacy and reduced overall treatment toxicity.

Aim 5: Correlating Pathologic Variables with Outcomes in Patients with Non-Urothelial Muscle Invasive Bladder Cancer After Bladder Preserving Trimodality Therapy: This project may provide information for use in the design of future RTOG and other trials that utilize trimodality therapy in the treatment of muscle-invasive bladder cancer patients with variant histologies.

Summary of Research Completed

Aim 1: No progress to report for this period.

Aim 2: Completed.

Aim 3: Statistical analyses were completed for two more biomarkers of interest, XPF and HIF1a. Mortality from high risk (large, deep, high grade) soft tissue sarcomas (STS) remains high and adjuvant chemotherapy has shown mixed results. Biomarker predictors of treatment response and outcome could improve patient selection for neoadjuvant therapy. Other biomarkers have been previously reported. The markers that were analyzed in this report period were XFP and HIF1a. Tissue microarrays (TMA) for biomarker expression were created using pre and post treatment tumor from 2 prospective high risk STS trials (Massachusetts General Hospital pilot and RTOG 9514) of neoadjuvant mesna, adriamycin, ifosfamide, dacarbazine (MAID)/44 Gy radiation/adjuvant chemotherapy. Automated scoring of the biomarker expression was performed using the Automated Quantitative Analysis (AQUA™) System and for each marker, scores were determined for nuclear, cytoplasm, and tumor mask. Biomarker expression was correlated with pathologic complete response (PCR) using Chi-squared tests and for disease-free survival (DFS), distant disease-free survival (DDFS), and overall survival (OS) using Cox regression models. Markers with multiple values per time point were handled as follows: XPF – minimum; Hif1a – maximum. There were 36 cases evaluable for XPF and 32 for HIF1a. Analyses looking at distributions of the markers and correlations with outcomes for the pre and post values have been

completed. The five tables included in this progress report (Tables 1 – 5, selected due to space constraints) show the distribution of post-treatment marker values as well as correlation with OS, DFS, and DDFS. The results from these markers are being interpreted, in conjunction with the other markers that have already been analyzed, in preparation for a single manuscript reporting the results of all of the markers.

Aim 4: No progress to report for this period.

Aim 5: Some additional analyses breaking out invasive disease was done for the ASTRO 2014 presentation. A CONSORT diagram and a set of panel graphs for overall survival, disease-specific survival, and bladder recurrence-free survival were created. The manuscript was submitted to the International Journal of Radiology Oncology Biology and Physics.

Aim 3: Table 1
Marker Value Summary: Post-Treatment

Marker	Study	n	Mean	SD	Min	Q1	Median	Q3	Max
XPF nuclear AQUA norm	RTOG 9514	17	8016.57	1992.93	6095.47	7250.20	7567.05	8018.98	15215.35
	MGH	19	8455.28	1371.33	6569.48	7322.19	8260.02	9324.91	11625.12
	Total	36	8248.11	1682.90	6095.47	7286.19	7778.64	8891.30	15215.35
XPF cytoplasm AQUA norm	RTOG 9514	17	3818.47	1178.17	2233.00	3334.60	3771.62	4338.96	7225.68
	MGH	19	3598.13	720.17	2624.27	2914.04	3353.94	3998.37	5245.54
	Total	36	3702.18	955.89	2233.00	3047.95	3583.84	4221.85	7225.68
XPF tumor mask AQUA norm	RTOG 9514	17	4756.24	1483.84	2717.92	4057.78	4733.15	5223.34	9430.09
	MGH	19	4824.24	912.59	3588.97	4111.84	4587.36	5329.66	6806.78
	Total	36	4792.13	1198.34	2717.92	4084.81	4652.70	5242.49	9430.09
Hif1a nuclear AQUA norm	RTOG 9514	13	4362.21	1583.99	1618.79	3248.35	4248.39	4966.20	7602.49
	MGH	19	6601.11	2439.43	2274.81	4810.64	6498.57	8335.45	12844.03
	Total	32	5691.56	2382.16	1618.79	3842.67	5502.91	7332.17	12844.03
Hif1a cytoplasm AQUA norm	RTOG 9514	13	1893.30	779.93	966.06	1414.18	1668.49	2292.86	3969.33
	MGH	19	3403.24	1395.52	1269.41	2394.39	3051.16	4094.26	6565.47
	Total	32	2789.83	1390.66	966.06	1702.60	2438.41	3784.14	6565.47
Hif1a tumor mask AQUA norm	RTOG 9514	13	2368.68	932.79	1098.09	1841.48	2355.22	2895.31	4754.04
	MGH	19	4421.83	1754.38	1552.10	3119.17	4519.56	5358.44	7781.92
	Total	32	3587.74	1781.45	1098.09	2190.53	3075.78	4688.94	7781.92

SD: standard deviation; Q1: first quartile; Q3: third quartile.

Aim 3: Table 2
Overall Survival: Univariate Analysis
Marker Expression as a Log-Transformed Continuous Variable: Post-Treatment

Marker	Events/total	HR (95%CI)	p-value
XPF nuclear AQUA norm	15/36	0.07 (0.00-3.57)	0.1858
XPF cytoplasm AQUA norm	15/36	0.29 (0.03-2.97)	0.2970
XPF tumor mask AQUA norm	15/36	0.17 (0.01-2.01)	0.1593
Hif1a nuclear AQUA norm	15/32	0.56 (0.16-1.97)	0.3677
Hif1a cytoplasm AQUA norm	15/32	0.42 (0.13-1.35)	0.1445
Hif1a tumor mask AQUA norm	15/32	0.45 (0.15-1.37)	0.1602

HR: hazard ratio; CI: confidence interval.

Aim 3: Table 3
Disease-Free Survival: Univariate Analysis
Marker Expression as a Log-Transformed Continuous Variable: Post-Treatment

Marker	Events/total	HR (95%CI)	p-value
XPF nuclear AQUA norm	17/36	0.01 (0.00-0.92)	0.0457
XPF cytoplasm AQUA norm	17/36	0.25 (0.03-2.35)	0.2271
XPF tumor mask AQUA norm	17/36	0.16 (0.02-1.62)	0.1207
Hif1a nuclear AQUA norm	16/32	0.59 (0.18-1.89)	0.3712
Hif1a cytoplasm AQUA norm	16/32	0.37 (0.12-1.16)	0.0885
Hif1a tumor mask AQUA norm	16/32	0.40 (0.14-1.17)	0.0945

HR: hazard ratio; CI: confidence interval.

Aim 3: Table 4
Distant Disease-Free Survival: Univariate Analysis
Marker Expression as a Log-Transformed Continuous Variable: Post-Treatment

Marker	Events/total	HR (95%CI)	p-value
XPF nuclear AQUA norm	17/36	0.01 (0.00-0.93)	0.0462
XPF cytoplasm AQUA norm	17/36	0.27 (0.03-2.43)	0.2415
XPF tumor mask AQUA norm	17/36	0.17 (0.02-1.69)	0.1304
Hif1a nuclear AQUA norm	16/32	0.59 (0.18-1.90)	0.3717
Hif1a cytoplasm AQUA norm	16/32	0.37 (0.12-1.17)	0.0913
Hif1a tumor mask AQUA norm	16/32	0.40 (0.14-1.17)	0.0951

HR: hazard ratio; CI: confidence interval.

Aim 3: Table 5

Marker Value Summary: Pre- and Post-Treatment Pairs

Marker	Timepoint	n	Mean	SD	Min	Q1	Median	Q3	Max
XPF nuclear AQUA norm	Pre-Tx	10	8772.27	1076.08	7070.68	7677.74	8901.68	9489.52	10503.53
	Post-Tx	10	8880.11	2407.45	6723.11	7442.14	8441.52	9324.91	15215.35
	Change (p=0.9291[1]) (p=0.8457[2])	10	107.84	2449.25	-2896.66	-1599.30	-291.79	1159.90	5458.47
XPF cytoplasm AQUA norm	Pre-Tx	10	4694.75	1556.58	2527.43	3972.63	4153.96	5806.80	7712.99
	Post-Tx	10	4240.83	1254.75	2853.02	3334.60	4179.81	4734.10	7225.68
	Change (p=0.5502[1]) (p=0.3750[2])	10	-453.93	2200.15	-4450.56	-1733.18	-427.68	553.59	4019.62
XPF tumor mask AQUA norm	Pre-Tx	10	6004.00	1598.71	3747.40	4964.83	5472.87	7412.93	8998.24
	Post-Tx	10	5511.37	1619.80	4057.78	4302.90	5284.91	5799.35	9430.09
	Change (p=0.5185[1]) (p=0.4316[2])	10	-492.64	2585.74	-4886.41	-2172.76	-519.42	839.96	4554.48
Hif1a nuclear AQUA norm	Pre-Tx	8	6267.08	1603.95	4593.31	4949.78	5993.26	7179.49	9298.29
	Post-Tx	8	5787.91	2241.58	2978.37	3842.67	5498.83	7877.98	8885.95
	Change (p=0.3892[1]) (p=0.4609[2])	8	-479.17	1921.65	-3662.78	-1590.75	-624.72	829.40	2601.54
Hif1a cytoplasm AQUA norm	Pre-Tx	8	3857.75	1221.78	2608.94	2959.48	3676.52	4301.47	6378.13
	Post-Tx	8	2978.52	1433.30	1294.85	1927.90	2501.12	4134.17	5406.95
	Change (p=0.1453[1]) (p=0.1484[2])	8	-879.23	1832.05	-2892.86	-2123.84	-1671.06	631.09	2186.66
Hif1a tumor mask AQUA norm	Pre-Tx	8	4582.47	1435.49	3103.41	3368.95	4509.57	5156.28	7486.74
	Post-Tx	8	4009.57	2108.14	1841.48	2435.49	3123.87	5738.09	7640.18
	Change (p=0.3219[1]) (p=0.3828[2])	8	-572.90	2380.72	-3409.88	-2080.00	-1397.47	1383.66	3014.32

SD: standard deviation; Q1: first quartile; Q3: third quartile.

[1] p-values are from paired t-test on log-transformed values.

[2] p-values are from Wilcoxon signed rank test on log-transformed values.

Research Project 2: Project Title and Purpose

Development and Evaluation of Novel Methods for Cancer Clinical Trial Interim Monitoring – Clinical trials provide first line scientific evidence necessary to advance treatment for cancer. With the increasing number of new treatment options being tested, there is a need for improvements in trial design and monitoring in order to a) terminate a trial in a timely manner when the therapy is ineffective, b) plan activities that take place during the trial efficiently (for example, interim safety and efficacy analyses), and c) derive and apply trial stopping rules and statistical power estimates that realistically reflect the interim data structure. To address these needs, we propose a series of methodological projects aimed at addressing current questions in clinical trial monitoring. These projects encompass a range of challenges in clinical trial conduct that apply broadly to cancer research as well as clinical research in general.

Anticipated Duration of Project

7/1/2012 – 12/31/2015

Project Overview

Three specific investigations are proposed as follows:

Aim 1: Comparison of futility monitoring methods using oncology trials: Futility monitoring is an important component in the conduct of clinical trials. An optimal rule would allow timely stopping if the new therapy is harmful or is unlikely to ultimately prove effective. Commonly used methods for futility monitoring include conditional power (CP) boundaries, repeated confidence intervals (RCI, adjusted for multiple looks), testing to reject the alternative hypothesis, and the recently proposed linear inefficacy boundary (LIB20). We will evaluate and compare the performances of these methods, using event histories from completed clinical trials of the Radiation Therapy Oncology Group (RTOG) and other cooperative groups

Aim 2: Prediction of landmark event times in oncology trials: In clinical trials with planned interim analysis, it can be valuable for logistical reasons to predict the times of landmark events such as the 50th and 100th event. Parametric (for example, Exponential, Weibull) models and nonparametric methods have been proposed for this purpose and these work well in simulation studies. However, the performance of these approaches has not been fully evaluated in real clinical trials. For this analysis, we will apply these prediction models to data from RTOG oncology trials with time to event as an outcome. These methods when applied to ongoing clinical trials will be useful tools for planning interim analyses and Data Safety Monitoring Board (DSMB) meetings.

Aim 3: Repeated confidence intervals and prediction intervals under fractional Brownian motion for stochastically curtailed tests: The repeated confidence interval (RCI) approach is an important method for sequential monitoring of clinical trials. Stochastically curtailed tests (SCT) also known as conditional power is another common approach. These methods are based on Brownian motion (BM) assumption, which is a special case of fractional Brownian motion

(FBM). However, it is possible that the interim statistic is an aggregated process of many different processes. Therefore, the future path will depend on both the past and current interim statistics. For these cases FBM is a good sensitivity measure to see if the observed processes deviate from BM and adjust the design accordingly. For this project we will derive RCIs based on SCTs under FBM and investigate the impact on sample size and design characteristics.

Principal Investigator

Qiang (Ed) Zhang, MD, PhD
Senior Statistician
American College of Radiology
1818 Market St. Suite 1600
Philadelphia, PA, 19103

Other Participating Researchers

James J. Dignam, PhD – employed by the American College of Radiology
Ed Korn, PhD; Boris Freidlin, PhD – employed by the National Cancer Institute
Dan Heitjan, PhD; Gui-Shuang Ying, PhD – employed by the University of Pennsylvania
Barry Davis, MD, PhD; Dejian Lai, PhD – employed by the University of Texas

Expected Research Outcomes and Benefits

Clinical trials are a critical step in the search for effective therapies of cancer as all reliable treatment options arise through this process. However, the process can be slower than desired if, for example, a futility monitoring method is unable to stop a trial when interim results are showing inefficacy of the new regimen; or, due to lack of accurate and precise method of predicting event times, DSMB meetings and trial closeout cannot be planned and scheduled efficiently. Clinical trials are process intensive, and most importantly require the greatly valued contribution of patient participants, who are seeking the best possible option for their personal situation, while at the same time contributing to research. A more efficient treatment evaluation strategy and logistic planning could improve both knowledge acquisition and patient care. Lastly, common assumptions for interim analysis tools need to be verified during actual trial conduct and adjusted if the conditions are not satisfied to guarantee enough statistical power to detect the expected treatment effect of the new therapy. We propose three areas of research that have immediate practical implications for cancer clinical trials. Comparing and selecting the optimal futility boundary will lessen patients' exposure to inactive treatment, improve resource utilization, and accelerate dissemination of important clinical information. Decisions regarding whether to continue or terminate a trial will be made more timely and efficiently if interim analyses and DSMB meetings take place at the accurately predicted event times. And finally, results on impact of deviations from design assumptions will inform investigators the importance of assessing these conditions and provide necessary tools for realistic clinical trial design and monitoring.

Summary of Research Completed

Aim 1: During the past year, we have received interim data from three other cooperative groups funded by National Cancer Institute (NCI) for trials to be included in this analysis for aim 1. From the National Surgical Adjuvant Breast and Bowel Project (NSABP), we examined eight large trials for breast cancer (N=5) and colon cancer (N=3), also there are 20 trials from the Mayo Clinic and 14 studies from Cancer and Leukemia Group B (CALGB). We applied each of the stopping rules to the 70 trials. Results of this analysis are shared with the co-PIs and each of the groups provided the corresponding datasets. Figures 1 and 2 show two examples of how a trial is monitored by different futility stopping rules, particularly when there are negative or positive results. We also started collecting time and sample sizes saved for each trial from the three groups.

Aim 2: From 07/01/2014 to 06/30/2015, we further analyzed results from the three prediction models (exponential, nonparametric, and Weibull) to predict event times in the head and neck cancer trial (RTOG 0129), which compared the survival from accelerated-fractionation radiotherapy (AFX) vs. standard-fractionation radiotherapy (SFX). From the preliminary results of 3 prediction approaches, it is clear that the exponential and Weibull models predicted poorly for the landmark dates, particularly for the prediction of time to reach the 206^{th} and 303^{rd} death. This poor performance in prediction could be due to the fact that some cancer patients may be cured from the disease by the treatments, and our current parametric prediction models did not consider the cure probability and the long term survivorship after chemo-radiation therapy.

Then we primarily focused on developing a Weibull cure-mixture model that could accommodate the cure probability from treatments. We did theoretical derivations regarding the improved model and developed programs for analyzing this type of data. We then applied this Weibull cure-mixture model to RTOG 0129 after justifying prior information for the model parameters. The prediction results are summarized and discussed in the draft paper; these show that the Weibull cure-mixture model performs better than the other three models. The manuscript has gone through several iterations of revisions and phone call discussions among Drs. Zhang, Heitjan and Ying. The manuscript will be ready for submission to the Journal of Clinical Oncology later this summer.

Aim 3: Based on the results of Davis and Hardy (1992), Zhang (2011) and Zhang, Lai, Davis (2012), we continued the analysis with previous established R programs. The impact of the different number of interim analyses and type I and type II error rates are also studied in table 1 with alpha=0.05. We can see these RCIs (Repeated Confidence Intervals) are more conservative than the Pocock and OBF designs and the Hurst parameter impact RCI width when it deviates from 0.5. We have chosen to use the Beta Blocker Heart Attack Trial (BHAT) as an example to illustrate how to apply these design methods. The H value is estimated to be 0.54, SD=0.0349. The repeated confidence intervals exclude hazard ratio of 1 when t=0.791 under both the Brownian motion and fractional Brownian motion models. With H=0.5, 0.54, the adjusted type I errors are 0.024893/0.0470 and 0.024890/0.04971 for the one-sided derived test. The ratios of RCI width to a fixed design at each time point is 3.31, 2.29, 1.69, 1.37 and 0.92 under Brownian motion and 3.31, 2.18, 1.64, 1.34 and 0.92 under fractional Brownian motion. These results

indicate that these one-sided derived tests based on SCTs have narrower final confidence intervals and require smaller sample sizes than those using classical group sequential designs. The Hurst parameter has more impact on the RCI width than on the sample size requirements for the proposed designs. This paper is now accepted for publication by *Communications in Statistics – Theory and methods*.

Table 1. Ratios of sample sizes for different number of interim analyses, alpha=0.05

	H=0.1	*H*=0.3	*H*=0.5	*H*=0.7	*H*=0.9	Pocock	OBF
K=3	0.66 0.83	0.68 0.85	0.69 0.85	0.69 0.85	0.68 0.85	0.99 1.19	0.73 0.89
K=4	0.70 0.87	0.70 0.86	0.69 0.85	0.69 0.85	0.69 0.85	1.05 1.25	0.75 0.91
K=5	0.73 0.90	0.71 0.87	0.70 0.86	0.69 0.85	0.69 0.85	1.09 1.29	0.77 0.93
K=10	0.83 0.99	0.74 0.90	0.71 0.87	0.70 0.86	0.79 0.97	1.18 1.39	0.82 0.98

*The ratios are between sample sizes of designs using RCIs based on SCTs, Pocock and OBF design types and a fixed sample size design with type II errors of 0.2 and 0.1.

Figure 1**:** Example of when futility rules are applied to a negative study

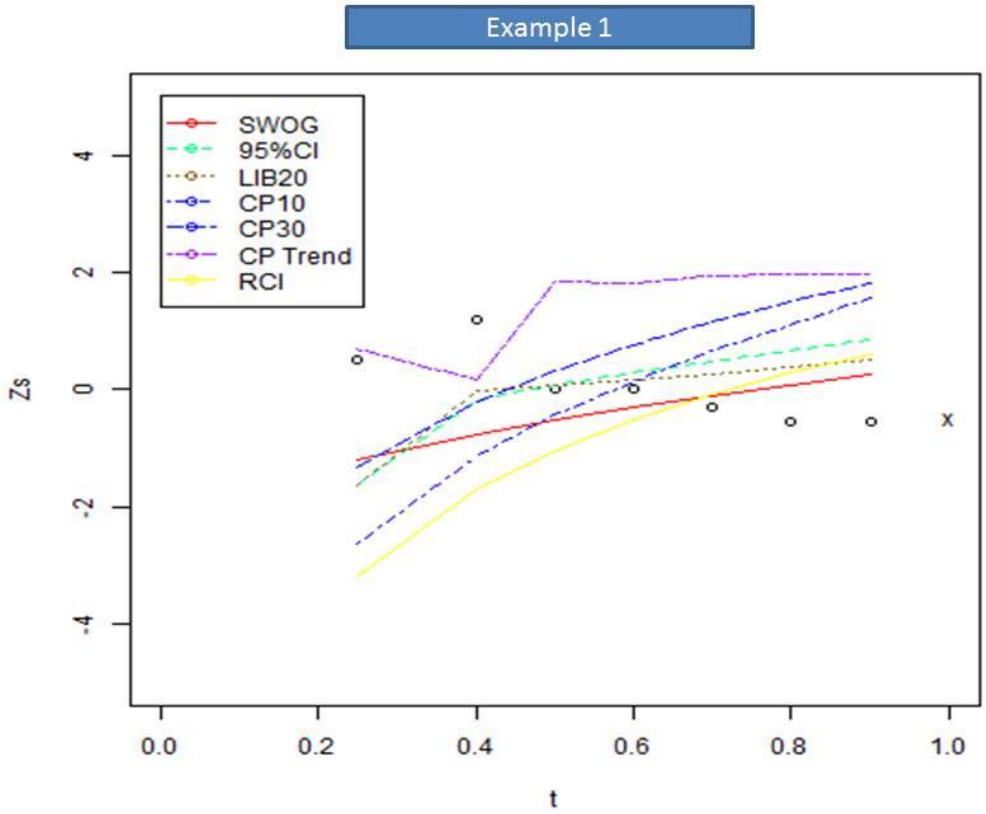
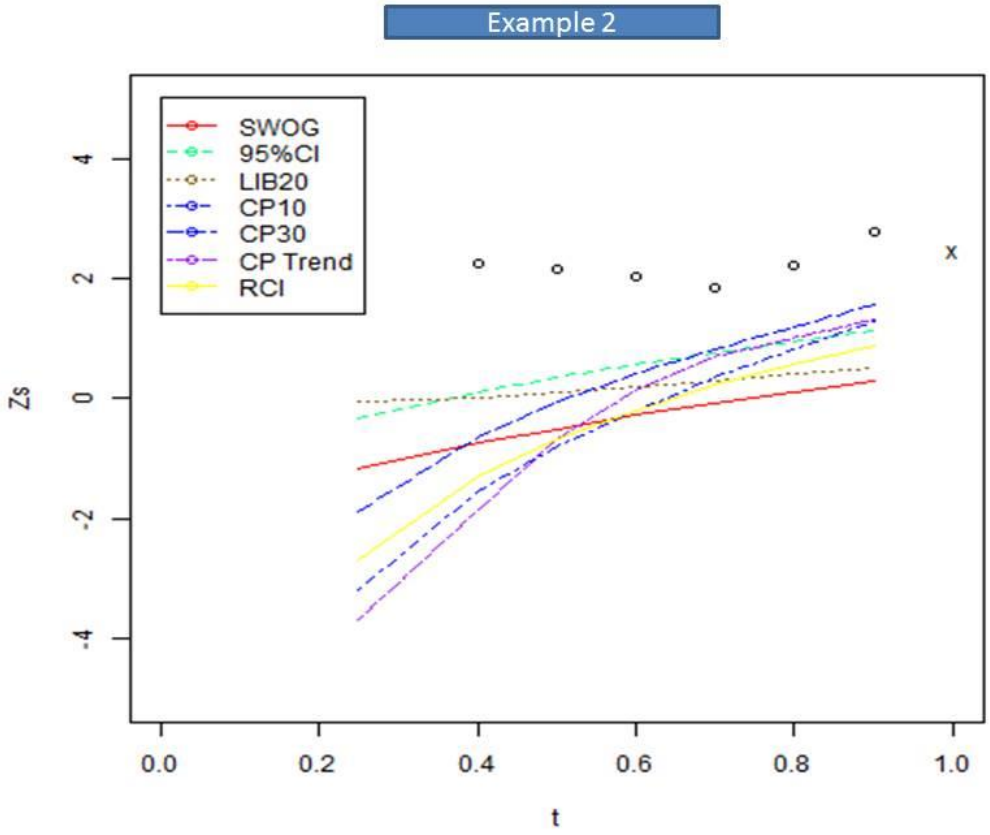


Figure 2**:** Example of when futility rules are applied to a positive study



Research Project 3: Project Title and Purpose

Biological Modeling of Tumor Control and Normal Tissue Complication for NSCLC Treated with SABR – Hypo-fractionated stereotactic ablative radiation therapy (SABR) is currently being used to treat early stage non-small cell lung cancer patients**.** The responses of tumor and nearby critical structures to SABR may be quite different from that of the conventional radiation therapy (RT) for which dose and radiobiological parameters for tumor control and toxicities of critical organs have been accumulated over the past two decades. Such parameters are still sparse and far from consensus for SABR treatment for lung cancer. The purpose of this study is to establish clinically useful nomogram for dose tolerance parameters and model the biological parameters for tumor control and normal tissue complication based on institutional data for hypo-fractionated lung SABR.

Anticipated Duration of Project

1/1/2012 – 12/31/2015

Project Overview

Three specific sub-projects are proposed as follows:

Aim 1: Establish a web-based software system to manage and evaluate patient clinical data, dose delivery and treatment outcomes for NSCLC patients treated with hypofractionated stereotactic ablative radiation therapy across institution: The core of the system software used in this project is a database that will store patient data. It will allow for a systemic integration of data for each patient as well as for all patients. As for the former, data of various provenances as Digital Imaging and Communications in Medicine-Radiation Therapy (DICOM RT) files or treatment outcomes will be available for each patient. This software system provides data-mining capability since it can seek emergent properties as derived from the entire content of the database. Each patient is quantum of information in a space of all treatment outcomes. Treatment outcomes of all accrued patients will be mapped back to their treatment plans.

Aim 2: Establishment of correlation between the dosimetric characteristics of treatment plans obtained with various dose computation algorithms and treatment outcome: Various dose engines exhibit different accuracy of calculation. The database will store all dose distributions resulted from these different dose engines and identify the most accurate dose distributions, which in turn will be equated with dose delivered during treatment. Software system will allow for turn-on key calculation of these dosimetric parameters that can be correlated to individual patient treatment outcomes. Moreover, the dosimetric parameters of all accrued patients with their treatment outcomes will be analyzed as a population-based data. This analysis will be implemented as one of the functionalities of the software system.

Aim 3: Establishment of correlation between biological characteristics inherent to treatment plans and treatment outcomes: Biologically-based characteristics of each treatment plan as is the case for their dosimetric characteristics are fixed after the plan is approved and the course of radiation treatment completed. Based upon the dosimetric data and treatment outcomes, biological parameters for tumor control and normal tissue complication will be derived. From these parameters and treatment outcomes, a clinically useful nomogram for dose tolerance parameters will be developed.

Principal Investigator

M. Saiful Huq, PhD
Professor
UPMC Cancer Pavilion
5150 Centre Ave., Room 542
Pittsburgh, PA 15232

Other Participating Researchers

Yongqian Zhang, Ph.D., Xiang Li, PhD; Ron Lalonde, Ph.D., Darek Michalski, PhD; Dwight E. Heron, MD – employed by the University of Pittsburgh. John Kang, MD, PhD – resident physician in internal medicine. Zachary Horne, MD – resident physician in radiation oncology at UPMC

Expected Research Outcomes and Benefits

The proposed methodology utilizes the current state-of-the-art medical informatics approach to investigate the combination and consolidation of patient data and outcome results. This type of study explores emergent features, which can only be derived from a consolidated approach to data, i.e., the sought results and conclusions would not be obtained looking and evaluating all data components separately. It combines data from various sources and stores data under unique patient key in database. Clinically driven data mining will expose patterns of dose distributions and resulting treatment outcome and will connect it with biological modeling of the treatment parameters. This will help establish a clinically useful nomogram for dose tolerance parameters and model biological parameters for tumor control and normal tissue complications. This knowledge cannot be acquired otherwise since the discovery of the correlation calls for a consolidated approach as described in this project. Data as saved in an electronic patient record (ePR) for this SABR-based database is always accessible and can be retrieved and processed in the future if new developments warrant. The quality of the results clearly will depend on the number of patients accrued in the system. We have treated over 700 non-small cell lung cancer (NSCLC) patients with SABR during the past 10 years. This along with a busy SABR program in our clinic guarantees dependable influx of patients for such processing. The nomogram to be developed in this project is expected to provide personalized adaptive management of NSCLC patients. The database will also allow for quantification of efficacy of SABR.

Summary of Research Completed

Here are the milestones completed during this past reporting year, 7/1/2014-6/30/2015:

Aim 1:

- Patient enrollment for treatment with different prescription dose. (Also for Aim 2)
    - Status – completed (we have 12Gy x4 and 30Gy x3 data)
    - 30 Gy x3 data will be focused on as it was meticulously curated by a single researcher

Aim 2:

- Conduct analyses of LQ vs. non-LQ for BED calculations and submit abstract for presentation and national meeting
    - Status –abstract was submitted and accepted at ASTRO 2014
- Determine the effect of various dose calculation algorithms, such as pencil beam convolution algorithm and anisotropic analytical algorithm, different dose scheme, and from different treatment modalities.
    - Status – in progress

Aim 3:

- Compare linear quadratic (LQ) and non-LQ models to calculate tumor control probability (TCP) for NSCLC and correlate with treatment outcomes
    - 12 Gy x4 schedule status – completed, presented at ASTRO 2014
    - 30 Gy x3 schedule status – in progress
- Compare classical survival analysis methods (Cox regression, logistic regression, etc.) and machine learning methods (support vector machine, neural network, etc.) to predict survival/TCP outcomes (local control, regional control, progression free survival, distant mets)
    - Status – in progress

We have completed collecting our dataset of 165 patients and are working on determining significant clinical/biological predictors of tumor control probability (TCP) by comparing classical survival models such as Cox proportional hazards regression with promising machine learning methods, including regularized logistic regression, support vector machine, and artificial neural networks. Cox regression is a commonly used survival model that is adept at handling censoring and assumes multiplicative, proportional hazards at any arbitrary point in time. Broadly defined, machine learning is a branch of computer science that deals with making predictions from complex data through statistical models. We decided to use machine learning methods in this project as it is a promising avenue that clinical translational research is moving towards that one of the researchers (John Kang) has significant experience in. To consolidate and showcase our knowledge, we have, during this past reporting year, written a review article on logistic regression, support vector machine, and neural networks titled "Machine Learning Approaches for Predicting Radiotherapy Outcomes: A Clinician's Perspective" with the participating researcher John Kang as the first author. It has finished peer review and we have been told it is about to be accepted by the International Journal of Radiation Oncology Biology Physics (*IJROBP*).

We decided to focus on TCP survival models for local control data since (1) there have been previous papers published making similar comparisons in lung SABR (Klement et al., *IJROBP* 2014), (2) the data is more objective as complication data may vary based on who collected it, (3) there is not yet an established prognostic tool for SABR TCP. From our dataset, normal tissue complication probability (NTCP) turned out to be too variable to be reliably modeled.

As the Klement et al. 2014 paper mentioned in the previous paragraph looked at logistic regression vs. support vector machine for predicting local control (with a minimum of 6 month follow up), we began our analysis by looking for the same outcome. As an initial screen for predictors that may be important, we performed univariate analysis with logistic regression. The results are shown in **Table 1** below.

Potentially interesting predictors on univariate analysis for multivariate analysis (based on a p-value cutoff of 0.2) include pack-years, lobe (right/left, middle/upper/lower), max pretreatment standardized uptake value (SUV), max pretreatment SUV >5, SUV/PTV ratio (PTV standing for "planning target volume"), clinical stage 1b vs. 1a, treatment gating, percent isodose, and max tumor dose mean. While not all these predictors may be clinically relevant, they form a starting

point. We are currently also running a similar univariate analysis for Cox proportional hazards regression as well and the preliminary results for significant predictors look similar to logistic regression.

For our machine learning methods, we began by using support vector machine (SVM) to attempt to predict which patients will have local control. SVM is a method that is resistant to multicollinearity effects and handles high dimensional data very well as it serves to non-parametrically separate data using a "n minus 1"-dimensional hypersurface for n-dimensional space. A brief and clinically relevant tutorial on SVM can be seen in Noble, *Nature Biotechnology* 2006. Our preliminary methods are using the WEKA (Waikato Environment for Knowledge Analysis) machine learning software and the LibSVM library with the radial basis function (RBF) kernel. We apply 10-fold cross validation for coarse grid search on the cost and the gamma hyperparameter (part of the RBF kernel) and are obtaining AUC of 0.591 (where 0.50 is a coin flip). We are finding that just slight alterations in the hyperparameters produce large variations in the AUC and thus we plan on doing a fine grid search, which will take longer computational time. Additionally, another major barrier is the fact that the data is unbalanced (an 8:1 ratio of local control to local failure). While this is good for our patients, it is not good for many models, which often ran better with information for both sides. We are mitigating this by looking at other important outcomes in our dataset which are more balanced, such as progression free survival (PFS) that has closer to a 2:1 ratio.

Currently, we are focusing only on 60 Gy (20 Gy x3) peripheral tumor patients as the data collection for these patients is robust and meticulously performed by Zachary Horne (a participating researcher). As such, the hypofractionated TCP MATLAB model (Aim 2) we developed and used in the July 1, 2013-June 30, 2014 reporting period is of limited use as it depends on variations in dosing in order to create variation in biological effective dose (BED). Once we refine our 60 Gy model, we plan on using our 48 Gy data (12 Gy x4) for central tumors to compare our results. At this point, our BED calculator will be relevant again and we plan on revisiting this comparison.

Table 1: logistic regression univariate analysis with potentially useful predictors bolded

Variable	No local failure (min 6 m f/u) (n=118)	Local failure (min 6 m f/u) (n=17)	p-value with the logistic regression
Cohort=1 N (%)	29/118 (25%)	3/17 (18%)	0.5323
Age at therapy Mean (SD)	76.0 (8.4)	76.0 (8.8)	0.9875
Female n (%)	67/118 (57%)	8/17 (47%)	0.4527
Black race vs white n (%)	11/96 (11%)	1/16 (6%)	0.5396
Smoking history N (%) Current Past	 18/109 (17%) 81/109 (74%)	 1/15 (7%) 12/15 (80%)	0.7589 Exact score test p=0.6264 0.9913 0.9871 0.9853

Never	9/109 (8%)	2/15 (13%)	reference
Unknown	1/109 (1%)	0/15 (0%)	
Smoking Packyears			**0.1020**
Mean (SD)	51.3 (29.9)	35.8 (22.2)	
Smoking			0.5433
N (%)			0.3204
None	7/84 (8%)	2/10 (20%)	0.6384
0.5 ppd	10/84 (12%)	2/10 (20%)	0.2375
0.5-1 ppd	39/84 (46%)	3/10 (30%)	reference
>1 ppd	28/84 (33%)	3/10 (30%)	
Years Since Quitting			0.6891
Mean (SD)	13.8 (13.3)	15.5 (12.9)	
Charlson Comorbidity Index			0.5964
Mean (SD)	3.8 (1.4)	3.6 (1.2)	
Age Adjusted CCI			0.7310
Mean (SD)	6.9 (1.4)	6.8 (1.0)	
Pre-Therapy HTN=1			0.4185
N (%)	48/88 (43%)	6/14 (43%)	
Pre_Therapy COPD=1			0.4230
N (%)	54/88 (61%)	7/14 (50%)	
Pre-Therapy DM=1			0.5336
N (%)	19/88 (22%)	2/14 (14%)	
Pre-Therapy CAD=1			0.4230
N (%)	34/88 (39%)	7/14 (50%)	
Pre-Therapy KPS			0.5943
Mean (SD)	79.6 (10.4)	78.2 (8.8)	
Laterality=L			0.8166
N (%)	52/118 (44%)	8/17 (47%)	
Up mid or lower			0.4000 Exact score test p=0.2519
N (%)			0.9707
Up	74/118 (63%)	8/17 (47%)	0.9674
Mid	3/118 (3%)	0 (0%)	reference
Lower	41/118 (35%)	9 (53%)	
Lobe			**0.1280 Exact score test p=0.0688**
N (%)			0.9516
RLL	19/118 (16%)	7/17 (41%)	0.9679
RML	3/118 (3%)	0/17 (0%)	0.9848
RUL	44/118 (37%)	2/17 (12%)	0.9738
LLL	22/118 (19%)	2/17 (17%)	reference
LUL	30/118 (25%)	6/17 (35%)	
Tumor Size			**0.1414**
Mean (SD)	2.3 (1.0)	2.7 (1.3)	
Max Pretreatment SUV			**0.1045**
Mean (SD)	6.8 (5.0)	9.6 (6.2)	
SUVMax_5=1			**0.1946**
N (%)	49/95 (52%)	8/11 (73%)	
SUV_PTV			**0.0127**
Mean (SD)	0.3 (0.2)	0.5 (0.3)	
Clinical T			0.2569
N (%)			0.9860
1a	61/118 (52%)	9/17 (53%)	

1b 2a	39/118 (33%) 18/118 (15%)	3/17 (18%) 5/17 (29%)	0.1513 reference
Clin_Stage=1b vs 1a N (%)	18/118 (15%)	5/17 (29%)	**0.1553**
Biopsy_Proven all one			
Histology N (%) Squamous Adeno Adenosquam large cell carcinoid NSCLC NOS	 37/118 (31%) 54/118 (46%) 1/118 (1%) 1/118 (1%) 1/118 (1%) 24/118 (20%)	 6/17 (35%) 10/17 (59%) 0/17 (0%) 1/17 (6%) 0/17 (0%) 0/17 (0%)	0.9079 Exact score test p=0.3011 0.9450 0.9433 0.9823 0.9220 0.9823 reference
Histology simplified N (%) Squamous Adeno NSCLC NOS/others	 37/118 (31%) 54/118 (46%) 27/118 (23%)	 6/17 (35%) 10/17 (59%) 1/17 (6%)	0.3255 0.3334 0.1821 reference
Squamous_vs_Adeno N (%)	54/91 (59%)	10/16 (63%)	0.8122
Type_of_Therapy N (%) Segment Cyberknife Trilogy Trubeam	 0/118 (0%) 42/118 (36%) 72/118 (61%) 4/118 (3%)	 0/17 (0%) 9/17 (53%) 8/17 (47%) 0/17 (0%)	0.4548 Exact score test p=0.3020 - 0.9821 0.9848 reference
PET_CT_planning N (%) 0 1 2 3 n/a yes	 5/108 (5%) 28/108 (26%) 33/108 (31%) 40/108 (37%) 1/108 (1%) 1/108 (1%)	 0/16 (0%) 7/16 (44%) 4/16 (25%) 5/16 (31%) 0/16 (0%) 0/16 (0%)	0.8951 Exact score test p=0.5938 0.9856 0.9722 0.9755 0.9754 0.9926 reference
Treatment Gated N (%)	39/69 (57%)	3/10 (30%)	**0.1293**
GTV volume Mean (SD)	12.5 (11.7)	15.7 (16.8)	0.5430
PTV volume Mean (SD)	34.2 (28.0)	36.1 (22.4)	0.8124
SBRT_prescribed_dose N (%) 48 60 6000	 1/118 (1%) 88/118 (75%) 29/118 (25%)	 0/17 (0%) 14/17 (82%) 3/17 (18%)	0.8137 Exact score test p=0.6232 0.9725 0.9706 reference
SBRT_fractions=4 vs 3 N (%)	1/118 (1%)	0/17 (0%)	0.9918

Fraction_size=20 vs 12 N (%)	88/89 (99%)	14/14 (100%)	0.9915 Exact score test p=1.0000
Time from Diagnosis (days) Mean (SD)	68.9 (45.0)	59.0 (33.8)	0.4008
Pct_isodose Mean (SD)	85.0 (5.1)	81.9 (3.0)	**0.0442**
Fields or Beams Mean (SD)	74.9 (81.8)	87.3 (85.0)	0.5973
Max Tumor Dose Mean (SD)	69.6 (7.3)	72.8 (3.1)	**0.0692**
Min Tumor Dose Mean (SD)	55.5 (8.8)	54.4 (13.0)	0.7074

Research Project 4: Project Title and Purpose

Quantitative Uncertainty Investigations for Clinical Trial Protocols – There are many factors that can confound the interpretation of results from cancer clinical trials that use radiation for therapy. Focus on such factors increases when a study gives an unexpected result that is counterintuitive. This was the case for a recent Radiation Therapy Oncology Group (RTOG) protocol, #0617 for Non-Small Cell Lung Cancer (NSCLC) where a lower dose arm gave significantly improved survival. The research proposed here examines radiation dose uncertainties that are either intentionally included in the protocol design process to improve accrual, or are unanticipated. Uncertainties for the RTOG 0617 protocol will be carefully analyzed to identify and quantify potential uncertainties.

Anticipated Duration of Project

7/1/2012 – 12/31/2015

Project Overview

Three specific investigations comprise this project:

Specific Aim 1.0 Quantitative uncertainty investigation: Evaluate uncertainties for the radiotherapy processes
Specific Aim 1.1 Structure delineation
Target definition is a major source of errors in radiation treatment. The variability in delineation of targets and critical normal structures has been shown to be highly variable. These variations, in turn, have been shown to have significant impact on dosimetric and radiobiological outcome. These variations will be quantified and their impact on treatment outcome will be simulated. Special attention will be paid to the process of image fusion, and the impact of this technique on accurate target delineation.

Specific Aim 1.2 Radiotherapy treatment planning
At least two components of treatment planning can produce major uncertainties for the radiotherapy process. First, variations can be introduced when the dose prescription is adapted for use in a clinical trial that is multi-institutional. Second, dose calculation, and plan optimization processes can introduce uncertainties. This sub aim will identify and model uncertainties associated with the treatment planning process.

Specific Aim 1.3 Radiotherapy delivery
Significant variations in the dose delivery step of the process have been reported to have discernable impact on treatment outcome. Examples are uncertainties in the equipment dose calibration and the possibility of delayed or deleted treatments. These variations will be quantified and simulated for inclusion in outcome predictions.

Specific Aim 1.4 TCP/NTCP and outcome model uncertainty
Techniques for Modeling of Tumor Control Probability (TCP) and Normal Tissue Complication Probability (NTCP) remain crude. Thus, differences between the outcome modeling methods used to analyze the data is another major variation. The various models emphasize different aspects of the input data. These uncertainties will be quantified as part of this research.

Specific Aim 2.0 Propagation of uncertainties to outcome
Strategies will be developed to combine the uncertainties derived from the investigations of specific aim 1 to the eventual outcome.

Specific Aim 3.0 Application of modeling to the RTOG 0617 protocol
The model for uncertainty propagation will be applied to the example of the unexpected results for the RTOG 0617 protocol.

Principal Investigator

Ying Xiao, PhD
Radiation Oncology Core Lab Physicist
American College of Radiology and Jefferson Medical College
1818 Market St. Suite 1600
Philadelphia, PA, 19103

Other Participating Researchers

Rebecca Paulus, BS; Elizabeth O'Meara, BS, RT (R) (T), CCRP – employed by the American College of Radiology
Adam Dicker, MD, PhD; Yan Yu, PhD – employed by Jefferson Medical College
James Galvin, PhD - Consultant

Expected Research Outcomes and Benefits

Evidence-based medicine has become the foundation of radiation oncology development. There are different levels of evidence upon which radiation therapy practice is based. The quality of the evidence depends significantly upon any variance for the individual chain of steps for the different techniques upon which the evidence is based. Radiation therapy consists of a large number of such stepwise processes, each of which poses variance/uncertainty that impact upon the final outcome, such as survival or quality of life.

In this project, we will identify and investigate the uncertainties associated with radiotherapy processes. These uncertainties will be propagated to affect the outcome predictions, using appropriate mathematical algorithms. We will collaborate with experts in the fields of computational/mathematical modeling.

The preliminary results from RTOG #0617 protocol comparing high-dose (74 Gy) with standard-dose (60 Gy) radiotherapy for treatment of non-small cell lung cancer were unexpected in that survival for the low-dose arm was significantly improved. This difference was not accompanied with any identifiable difference in radiation toxicity. This unexpected result might be explained by the propagation of uncertainties that could, in effect, decrease the stated prescribed dose differences. Investigation of this unexpected result will serve as an example of the use of the modeling techniques developed in this research project. This investigation is critically important for informing the design of subsequent similar trials. We plan to study uncertainties associated with all the processes of the trial, and quantify these uncertainties and their impact upon the outcome.

Summary of Research Completed

Specific Aim 1.1: While investigating the uncertainties in structure delineation, the team also explored the possibility of validating a quality assurance tool for cardiac structure delineation, and studied the optimization of technical parameters for using the tool for the evaluation of contour accuracy. The efforts yield the following publication level abstract:

Abstract 1: Optimizing Technical Parameters for Using Atlas Based Automatic Segmentation for Evaluation of Contour Accuracy – Experience with Cardiac Structures from NRG Oncology/RTOG 0617
Purpose: Accurate contour delineation is crucial for radiotherapy. Atlas based automatic segmentation tools can be used to increase the efficiency of contour accuracy evaluation. This study aims to optimize technical parameters utilized in the tool by exploring the impact of library size and atlas number on the accuracy of cardiac contour evaluation.
Methods and materials: Patient Computed Tomography Digital Imaging and Communications in Medicine (CT DICOMs) from RTOG 0617 were used for this study. Five experienced physicians delineated the cardiac structures including pericardium, atria and ventricles following an atlas guideline. The consistency of cardiac structured delineation using the atlas guideline was verified by a study with four observers and seventeen patients. The CT and cardiac structure DICOM files were then used for the Atlas-based Autosegmentation (ABAS) technique.

To study the impact of library size (LS) and atlas number (AN) on automatic contour accuracy, automatic contours were generated with varied technique parameters for five randomly selected patients.
Three LS (20, 60, and 100) were studied using commercially available software. The AN was four, recommended by the manufacturer. Using the manual contour as the gold standard, the Dice Similarity Coefficient (DSC) was calculated between the manual and automatic contours. Five-patient averaged DSCs were calculated for comparison for each cardiac structure.
In order to study the impact of AN, the LS was set at 100, and AN was tested from one to five. The five-patient averaged DSCs were also calculated for each cardiac structure.
Results: The variation of DSC against the LS was shown in Table 1 (A), which suggests the best LS to be 100. The variation of DSC against the AN was shown in Table1 (B), indicating the four-atlas gave the best DSC performance when LS was 100.
Conclusions: By comparing DSC values, the combination AN=4 and LS=100 gives the best performance (See Table 1).

Specific Aim 1.2: The team investigated the quality of treatment planning for two RTOG trials, and the feasibility of dosimetric compliance criteria for one trial. The efforts yielded the following three publication level abstracts:

Abstract 2: Evaluation of Lung Intensity Modulated Radiation Therapy (IMRT) Plans Quality using a Knowledge Based Engineering Tool
Purpose: A knowledge based engineering tool is used to generate a dose volume histogram for the organs at risk based on the anatomical information of the patient and a model trained using a database of high quality plans. The purpose of this work is to use this tool to 1) guide the establishment of dosimetric evaluation criteria for the RTOG 1308 lung clinical trial and 2) assist in the quality assessment and improvement of treatment plans that fails to meet the dosimetric compliance criteria.
Materials and Methods: Twenty six lung IMRT treatment plans developed at two institutions were included in the study. These plans were also used to guide the development of the dosimetric compliance criteria of the RTOG 1308 lung trial. DICOM CT, RP, RD and RS data were imported to the tool. The tool requires the matching of planning target volume (PTV) and organs at risks (OARs) and the prescription dose of a candidate review to select one or more plans among those in the database of original plans. The tool then generates a modeled dose volume histogram (DVH) for organs at risk for the candidates' plan. The predicted DVHs were compared with the planned DVHs to assess the plan quality.
Results: DVHs for organs at risk generated by the tool were compared to those of the respective treatment plans. The comparison of DVHs for organs included in the dosimetric compliance criteria of the trial (total lungs, heart, esophagus and spinal cord) indicated some variations between the modeled and calculated DVHs. However, most of the variations were within the confidence limits of the predicted DVHs. The results also indicated that IMRT plans used during the design study of the RTOG 1308 trial were of good quality.
Conclusions: The quality of lung intensity modulated radiation therapy (IMRT) plans used during the development of the RTOG 1308 clinical trial was investigated using a knowledge based engineering tool. The results indicated that the plans are of good quality. We aim to

eventually use this tool to provide real time feedback for plan optimization, the process for which will be tested in future studies.

Abstract 3: Treatment Plan Quality Assurance for XXX* With Knowledge Engineering
Purpose/Objective(s): The purpose is to report on the treatment plan quality of cases submitted for XXX, and to use the knowledge-guided treatment planning tool to predict the best achievable plans for organs-at-risk (OARs) sparing. We hypothesized that this tool would aid in improving radiation plan quality for cooperative group trials
Materials/Methods: The dosimetric data from 80 cases submitted for XXX, a phase 2 post prostatectomy adjuvant radiation trial, were evaluated against the criteria specified in the protocol. The knowledge-guided treatment planning tool was used to predict best achievable plans and was used to evaluate whether we can improve the plan quality indices. This tool was developed from the evidence-based approach by learning from a database of 88 high-quality prior prostate IMRT plans. The anatomical features of the PTVs and OARs and their spatial relationships on OARs dose sparing variation were modeled by the multiple regression method. The DVHs were analyzed and tested against the protocol criteria. For cases failing the criteria, we implemented the knowledge-guided treatment planning tool on their DICOM data and obtained the best achievable DVH ranges, which were checked against the criteria.
Results: Of the 80 submitted cases (10 cases used 3DCRT technique and 70 used IMRT), 72 cases (90%) met per protocol/variation acceptable criteria for target coverage, 66 cases (83%) for OAR sparing. We used the knowledge-guided treatment planning tool to predict the best achievable OARs dosimetric values. For the 14 cases that failed OAR criteria, 72% improved to per protocol, 14% improved to variation acceptable, and 14% still deviation unacceptable. All dosimetric values improved except in one 3D and one IMRT case for bladder V50Gy criterion. In summary, of the 14 cases having deviation unacceptable, 12 of them were improved to variation acceptable or better.
Conclusion: We reported good dosimetric plan quality for cases submitted for XXX. The knowledge-guided treatment planning tool can be used to guide improvement of treatment plans in future clinical trials.
**The actual trial number has been de-identified due to the discussion of quality assurance issues.*

Abstract 4: On the Feasibility of the Dosimetric Compliance Criteria of NRG-HN002: A Randomized Phase II Trial for Patients with p16 Positive, Non-Smoking Associated, Locoregionally Advanced Oropharyngeal Cancer
Purpose: NRG-HN002 is a randomized phase II trial for patients with good-risk human papillomavirus (HPV)-associated oropharyngeal cancer (OPC). The trial will evaluate two new treatment options, one consisting of chemoradiotherapy and one of radiotherapy alone. The prescription dose is 60 Gy (2.0 Gy/fraction) over 6 weeks for arm 1 and over five weeks for arm 2. The trial will select the arm(s) achieving a 2 year progression-free survival rate of $\geq$ 85% without unacceptable swallowing toxicity. As part of the trial planning process, we assessed the feasibility of the required dosimetric compliance criteria using different treatment planning systems and techniques. We offer technical guidance in regards to the expected outcomes of these dosimetric planning goals.
Materials and Methods: Five benchmark patients meeting the trial inclusion criteria were used for dosimetric testing at different institutions. Contours for target volumes and organs at risk

were drawn according to study guidelines. Plans using IMRT with nine beams at 40 degree separation from AP and VMAT with two full arcs were developed using Varian Eclipse version 11.0.30 and Philips Pinnacle3 version 9.10. Planning priorities, according to the protocol, were: (1) meet spinal cord and brain stem constraints; (2) meet PTV coverage and homogeneity objectives; (3) meet remaining normal tissue. The dosimetric compliance criteria included target volumes (CTVs and PTVs); SpinalCord; SpinalCord_05; and BrainStem_03. Recommended dose levels were assigned to other structures. Dose volume histograms of the plans were analyzed and tested against the compliance criteria of the protocol.
Results: All dosimetric compliance criteria were met using the two planning techniques in Eclipse and Pinnacle with the exception of PTV_4800 maximum dose criteria, when the volume overlaps with PTV_6000. Coverage (D95% (Gy) and D99% (Gy) for the three PTVs (PTV_6000, PTV_5400 and PTV_4800) was met in all cases. The average maximum dose in all the plans for SpinalCord_05, SpinalCord and BrainStem_03 were 45 ± 4 Gy, 40 ± 5 Gy and 30 ± 9. Gy. Recommended dose values for other OARs (Parotid, Larynx, Pharynx, Submandibula_R or Submandibula_L (contralateral), OralCavity, Esophagus_Up, Mandible and NonPTV) were also achieved with minor deviations in a very few occasions.
Conclusions: All dosimetric compliance criteria for the newly developed head and neck NRG HN002 clinical trial were readily achieved using different treatment planning systems and techniques, with the exception of PTV_4800 maximum dose. This information will be used to adjust protocol parameters as necessary. This represents a practical improvement in trial planning processes. By establishing in advance that specified dosimetric criteria are feasible across a variety of centers with varied techniques, the efficiency of the clinical trial launch may be enhanced. Technical guidance to achieve the dosimetric planning goals is available prior to trial launch.

Specific Aim 1.3: While investigating the variations in treatment plan delivery, we explored quantitatively evaluating the accuracy of two dimensional image guidance radiation therapy for clinical trial credentialing. This effort yielded the following publication level abstract:
Abstract 5: Quantitative Evaluation of 2D-2D and 2D-3D Image Guided Radiation Therapy (IGRT) for Clinical Trial Credentialing
Purpose: 2D-2D kV IGRT credentialing evaluation for clinical trial qualification was historically qualitative and was conducted by submitting screen captures of the fusion process. However, as quantitative DICOM 2D-2D and 2D-3D image registration tools are implemented in clinical practice for better precision, especially in centers that treat patients with protons, better image guided radiation therapy credentialing techniques are needed. The aim of this work is to establish methodologies for quantitatively reviewing (IGRT) submissions based on DICOM 2D-2D and 2D-3D image registration, and to test the methodologies in reviewing 2D-2D and 2D-3D IGRT submissions for NRG Oncology clinical trials qualifications.
Methods: DICOM 2D-2D and 2D-3D automated and manual image registration has been tested using the Harmony tool in MIM software. 2D kV orthogonal portal images are fused with the reference digital reconstructed radiographs (DRR) in the 2D-2D registration while the 2D portal images are fused with DICOM planning CT image in the 2D-3D registration. The Harmony tool allows alignment of the two images used in the registration process and also calculates the required shifts. Shifts calculated using MIM are compared with those submitted by institutions

for IGRT credentialing. Reported shifts are considered to be acceptable if differences are less than 3mm.
Results: Several tests have been performed on the 2D-2D and 2D-3D registration using Harmony tool in MIM. The results indicated good agreement between submitted and calculated shifts. A workflow for reviewing these IGRT submissions has been developed and will eventually be used by medical physics co -chairs of NRG Oncology clinical trials to review IGRT submissions.
Conclusion: The IROC Philadelphia RTQA center has developed and tested a new workflow for reviewing DICOM 2D-2D and 2D-3D IGRT credentialing submissions made by different cancer clinical centers, especially proton centers. The NRG Center for Innovation in Radiation Oncology (CIRO) and the IROC RTQA center continue their collaborative efforts to enhance quality assurance services and to be consistently adaptive to the new advances in radiation therapy.

Specific Aim 1.4: One factor that contributes to the uncertainties of patient TCPs is the definition of minimum and maximum dose for target volume. This effort yielded the following publication level abstract:

Abstract 6: On Definition of Minimum and Maximum Dose for Target Volume
Purpose: This study aims to investigate the impact of different minimum and maximum dose definitions in radiotherapy treatment plan quality evaluation criteria by using tumor control probability (TCP) models.
Methods: Dosimetric criteria used in the RTOG 1308 protocol are used in this investigation. RTOG 1308 is a phase III randomized trial comparing overall survival after photon versus proton chemoradiotherapy for inoperable stage II-IIIB NSCLC (non small cell lung cancer). The prescription dose for planning target volume (PTV) is 70Gy. Maximum dose (Dmax) should not exceed 84Gy and minimum dose (Dmin) should not go below 59.5Gy in order for the plan to be "per protocol" (satisfactory).
A mathematical model that simulates the characteristics of the PTV dose volume histogram (DVH) curve with normalized volume is built. The Dmax and Dmin are noted as percentage volumes Dη% and D(100-δ)%, with η and δ ranging from 0 to 3.5. The model includes three straight line sections and goes through four points: D95%= 70Gy, Dη%= 84Gy, D(100-δ)%= 59.5 Gy, and D100%= 0Gy. For each set of η and δ, the TCP value is calculated using the inhomogeneously irradiated tumor logistic model with D50= 74.5Gy and γ50=3.52.
Results: TCP varies within 0.9% with η and δ values between 0 and 1. With η and δ varies between 0 and 2, TCP change was up to 2.4%. With η and δ variations from 0 to 3.5, a maximum of 8.3% TCP difference is seen.
Conclusion: When the defined maximum and minimum volume varied more than 2%, significant TCP variations were seen. It is recommended that less than 2% volume be used in the definition of Dmax or Dmin for target dosimetric evaluation criteria.

Specific Aim 2.0: No progress on this aim.

Specific Aim 3.0: No progress on this aim.

Table 1

DSC	(A)			(B)				
	4_20	4_60	4_100	1_100	2_100	3_100	4_100	5_100
Pericardium	0.89±0.04	0.90±0.02	0.90±0.02	0.90±0.02	0.89±0.03	0.91±0.03	0.90±0.02	0.91±0.03
Atria	0.77±0.06	0.72±0.05	0.75±0.06	0.75±0.05	0.73±0.06	0.73±0.08	0.75±0.06	0.69±0.08
Ventricles	0.85±0.07	0.84±0.05	0.86±0.02	0.83±0.05	0.81±0.04	0.85±0.06	0.86±0.02	0.86±0.04

Research Project 5: Project Title and Purpose

Arterial Stiffness and Wave Reflections as Determinants of Regression of Left Ventricular Hypertrophy and Fibrosis Assessed with Cardiac MRI After Aortic Valve Replacement for Severe Aortic Stenosis – This project will evaluate the importance of arterial stiffness and wave reflections as determinants of persistent left ventricular (LV) hypertrophy and fibrosis (assessed using cardiac magnetic resonance imaging [MRI]) after correction of severe stenosis (tightness) of the aortic valve. We aim to test the hypothesis that stiff arteries and increased wave reflections impede pumping of blood by the LV after aortic valve replacement and prevent adequate regression (improvement) of hypertrophy and fibrosis of the myocardium despite correction of aortic valve stenosis. Proof of hypothesis would identify potentially treatable abnormalities identifiable on imaging for future targeted therapy. This project also will assess the value of a novel cardiac MRI sequence to characterize myocardial fibrosis without the use of gadolinium.

Anticipated Duration of Project

1/1/2012 – 12/31/2015

Project Overview

Previous C.U.R.E. funding established a network of academic medical centers in Pennsylvania (ACRIN PA) with the broad goal of advancing the role of imaging in the detection and/or treatment of disease by conducting early stage imaging clinical trials. This project seeks to continue the work of that network. This multi-institutional project will prospectively evaluate potential determinants of the regression (improvement) of left ventricular hypertrophy and fibrosis assessed by cardiac MRI before and after aortic valve replacement (AVR) for severe aortic stenosis. We will enroll 80 eligible participants with severe aortic stenosis who are scheduled to undergo AVR. A gadolinium-enhanced cardiac MRI scan, along with arterial pulse wave recordings and novel non-contrast myocardial tissue characterization sequences (T1rho

mapping), will be performed immediately prior to AVR and repeated 6 months after AVR. These data will be used to assess left ventricular mass (LVM), left ventricular myocardial fibrosis, arterial stiffness, and wave reflections. We will test the hypotheses that arterial stiffness and arterial wave reflections are associated with a less pronounced reduction of LVM and fibrosis and with a greater degree of residual fibrosis and hypertrophy despite correction of aortic stenosis via AVR. Importantly, we also will assess the value of T1rho imaging in detecting the degree of myocardial fibrosis at baseline and degree of reduction after AVR, using post-gadolinium T1-mapping as a reference method.

Principal Investigator

Mitchell D. Schnall, MD, PhD
Professor of Radiology
University of Pennsylvania
Hospital of the University of Pennsylvania
3400 Spruce St.
Philadelphia, PA 19104

Other Participating Researchers

Scott R. Akers, MD, PhD; Julio A. Chirinos, MD, Walter Witschey, PhD – employed by the Hospital of the University of Pennsylvania

Expected Research Outcomes and Benefits

We expect to demonstrate that arterial stiffness and wave reflections are important determinants of residual left ventricular hypertrophy and fibrosis (adverse prognostic markers assessed with cardiac MRI) after aortic valve replacement. This would identify a novel, potentially treatable mechanism that could be targeted with therapy in future trials and can be assessed by cardiac imaging studies.

We expect to validate T1rho, a novel MRI imaging method that does not require gadolinium contrast, as a technique for the assessment of myocardial fibrosis. This would allow for myocardial fibrosis (an important abnormality that needs to be assessed in several cardiac conditions) to be imaged without the use of gadolinium contrast, which is contraindicated in many patients who have advanced kidney impairment.

Summary of Research Completed

Specific Aim 1: We made progress in enrolling patients to this trial; Specific Aim 1 relates to the final data analysis which has not yet completed.

Specific Aim 2: We made progress in enrolling patients to this trial; Specific Aim 2 relates to the final data analysis which has not yet completed.

Specific Aim 3: We made progress in enrolling patients to this trial; Specific Aim 3 relates to the final data analysis which has not yet completed.

Specific Aim 4: We made progress in enrolling patients to this trial; Specific Aim 4 relates to the final data analysis which has not yet completed (planned for the 4th quarter of 2015).

The following is an update for the reporting period (July, 2014-June, 2015) for the research study and beginning steps for analysis that support the 4 aims noted above:

Study Accrual

During the reporting period (July, 2014-June, 2015), subject enrollment almost tripled with a total of 39 cases enrolled (Table 1). A fifth site was added, Lancaster General Hospital, to increase enrollment. The study closed to accrual on 5/15/2015. The baseline MRIs were completed by May 31 and it is anticipated that follow up will continue into October. Study demographics are included in Table 2.

Core Lab Quantification

All cardiac MRI images have been transferred to the cardiac MRI quantification core lab at the University of Pennsylvania. Similarly, all digitized central pressure and carotid-femoral pulse wave velocity tracings have been measured.

- We have finalized the modifications to our dedicated software programmed in Matlab (Segment Software as well as several custom-designed software interfaces and routines) for quantification purposes.
- We have quantified all MRI data for the baseline visit (all 39 subjects enrolled). The finalized quantification includes
    - LV segmentation (only 2 cases have not been quantified due to issues with image data transfer; we are in the process of troubleshooting the image data transfer to finalize the LV mass quantification)
    - Myocardial fibrosis (as described below): all cases have been quantified, except for 1 case in which Look Locker sequences had not been transferred appropriately. Once we receive these files, we will finalize all myocardial fibrosis quantification for the baseline visit.
    - Carotid-femoral pulse wave velocity: we have finalized the quality control and quantification for baseline visit data.
    - Wave reflections: we have finalized the quality control and quantification for baseline visit data.

Summary of methods (specific methodology for each quantification noted above)

Measurement of LV mass: Epicardial and endocardial borders of short-axis SSFP cines were contoured on diastolic and systolic frames. Total volumes were obtained by adding together volumes from the contiguous slices (slice summation method) allowing for calculation of diastolic and systolic volumes. Myocardial mass was calculated from the volume difference derived from epicardial and endocardial contours, assuming a specific myocardial density of 1.05 g/cm^3.

Quantification of myocardial fibrosis: We computed the extracellular volume fraction (Ve) in human myocardium by exploiting the linear relationship between change in relaxivity (i.e., 1/T1) and gadolinium (Gd) concentration. Ve is a direct quantitative metric of the interstitial compartment and has been validated against the collagen volume fraction in humans. This robust noninvasive measure employs Gd contrast as an extracellular marker and requires measurement of longitudinal relaxation times (T1) and the hematocrit. Myocardial T1 measures were derived from a series of images with variable inversion times from which the collagen volume fraction, Ve, is ultimately computed. To obtain T1 values from cardiac MRI data, we used a 3 parameter model to describe signal intensity (SI) as a function of inversion time (TI): $SI=| A - B \cdot e^{(-TI/T1^*)} |$, where T1=T1*((B/A)-1). Least square estimates of model parameters continue to be obtained using the Levenberg-Marquardt algorithm implemented in the Matlab software. Ve is measured as outlined by Jerosch-Herold et al. Specifically:

$$Ve = [\lambda \cdot \rho \cdot (1\text{-hematocrit})] - Vp$$

where Ve is the myocardial extracellular volume fraction, Vp is the myocardial plasma volume fraction (assumed to be 0.045, reflecting capillary density), and $\lambda=[\Delta R1_{myocardium}] / [\Delta R1_{bloodpool}]$ pre and post Gd contrast administration (where R1=1/T1). Hematocrit is measured from blood samples drawn at the time of each cardiac MRI session (usually through a freshly placed intravenous line). Ve measurement requires steady state equilibrium between plasma and interstitial Gd contrast. Given the relatively slow clearance of Gd, the bolus technique accurately measures Ve compared to a constant infusion technique (to achieve steady state equilibrium), thus greatly facilitating integration of LV myocardial fibrosis quantification into the workflow of the study, employing a single contrast bolus. A single contrast bolus technique has also been validated against histologic measures of fibrosis in humans.

Quantification of aortic stiffness: We quantified aortic stiffness as carotid-femoral pulse wave velocity, using arterial tonometry (Sphygmocor device).

Quantification of wave reflections: We measured aortic flow using through-plane proximal aortic phase-contrast MRI. The flow waveform along with the pressure waveform obtained via arterial tonometry is combined and modeled to assess aortic input impedance and arterial load as recently reviewed by Chirinos and Segers. Using the central pressure and flow waveforms, linear wave separation analysis was performed. First, characteristic impedance of the proximal aorta (Zc), which describes the relationship between pulsatile pressure and flow in the absence of wave reflections, is calculated by averaging the moduli of the 3rd to 10th harmonics of input impedance (i.e., pressure moduli divided by corresponding flow moduli). Pressure and flow waveforms are then used for wave separation as follows:

$$\text{Forward pressure} = (P+Zc*Q)/2$$

$$\text{Backward pressure} = (P-Zc*Q)/2$$

Where P and Q denote harmonics are derived from the measured pressure and flow waveform, using Fourier decomposition. To obtain the total forward and backward wave, forward and backward harmonics are summated. Reflection magnitude is calculated as backward wave amplitude/forward wave amplitude.

Table 1: Accrual by Institution

Institution	Date Site Opened	Total Accrual
Pennsylvania State/Hershey Med Ctr	10/24/2013	1
Hosp of the U Pennsylvania	07/31/2013	5
U Pittsburgh Med Ctr	11/04/2013	21
Veteran Affairs Med Ctr– Phila.	08/05/2013	6
Lancaster Gen Hospital	07/28/2014	6
Total (5 institutions):		**39**

Table 2: Demographics

AGE	
Median	74
Minimum	49
Maximum	91
RACE	
American Indian or Alaskan Native	0 (0%)
Asian	0 (0%)
Black or African American	3 (8%)
Native Hawaiian or Other Pacific Islander	0 (0%)
White	36 (92%)
Reported As Unknown	0 (0%)
ETHNICITY	
Hispanic or Latino	0 (0%)
Not Hispanic or Latino	39 (100%)
Reported as Not Reported	0 (0%)
Reported as Unknown	0 (0%)
GENDER	
Male	26 (67%)
Female	13 (33%)