American College of Radiology

Annual Progress Report: 2009 Formula Grant

Reporting Period

July 1, 2010 – June 30, 2011

Formula Grant Overview

The American College of Radiology received $2,043,960 in formula funds for the grant award period January 1, 2009 through December 31, 2013. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

Novel Methods for Cancer Clinical Trial Design and Analysis - Clinical trials provide the critical evidence necessary to advance treatment for cancer. With the ever growing number of promising interventions, there is a need for improvements in trial design in order to a) obtain answers more quickly, b) conserve and optimize resources, and c) make better choices of what treatments to pursue in further evaluation. In addition, as treatment regimens become more complex and multimodal, the ability to accurately characterize whether anticipated benefits with respect to specific disease event reduction have occurred requires extensions of standard analytic methods. To address these needs, we propose a series of methodological projects aimed at addressing current questions in clinical trial design and analysis. These projects encompass a range of needs that apply broadly to cancer clinical trials and research in general.

Anticipated Duration of Project

1/1/2010 - 12/31/2013

Project Overview

Three specific investigations are proposed as follows:

Aim 1: Development and Use of an Efficient Phase II/III Transition Design – The traditional paradigm for therapy development involves a pilot safety and efficacy trial (phase II) followed by a definitive Phase III comparative trial if warranted. This development model is intensive with respect to the time and logistical overhead involved in conducting sequential studies, and too often leads to failure in Phase III despite promising Phase II data on seemingly similar targeted populations. We propose to evaluate and implement a novel Phase II/III transition design that has thus far been little used in the oncology setting.
Aim 2: Alternative Metrics for Time to Event Endpoints in Phase II and III Trials – Phase II trials have traditionally been formulated as one-sample designs where all patients receive the treatment of interest. While statistical power for comparison to fixed benchmark values can be adequate

within a feasible sample size, the design suffers from dependence on historical comparisons that may not prove reliable. An alternative is a randomized Phase II design, where either a) treatment arms are not formally compared, but rather the arm that prevails to any degree is taken as more favorable with respect to further development, or b) adequately powered comparisons for simple endpoints such as fixed-time proportions failure-free are feasible. We propose analytic development of an approach using quantile (median, etc) estimation and comparison in the randomized Phase II setting. The approach is equally applicable to Phase III trials, and may have particular advantages in the presence of non-proportionality.
Aim 3: Estimating Treatment and Covariate Effects Under Competing Risks – Competing risks, whereby patients are subject to multiple potential failure types, with only one of these occurring as the primary first failure, are ubiquitous in cancer. In addition to multiple cancer-specific events (i.e., local, regional, distant recurrence), patients may experience second primary cancers or deaths from other causes that preclude any cancer event. While correct estimation of event-specific probabilities of occurrence for competing risks is straightforward, inference in the presence of competing risks remains more challenging. We propose to investigate and compare different recently developed competing risks modeling methods.

Principal Investigator

Meihua Wang, PhD
Senior Statistician
American College of Radiology
1818 Market Street, Suite 1600
Philadelphia, PA 19103

Other Participating Researchers

Ed Zhang, PhD, Thomas Pajak, PhD - employed by American College of Radiology
James J. Dignam, PhD, Maria Kocherginsky, PhD - employed by University of Chicago

Expected Research Outcomes and Benefits

Cancer clinical trials are a critical component of cancer care and as all reliable treatment options arise through this process, it is only through systemic and comprehensive evaluation in a trial setting that the risks and benefits of any option can be assessed. However, the process can be slower than desired, is always process intensive, and requires 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. In addition, an increasing array of agents in which to try in some disease settings further strains the development system. A more efficient treatment evaluation strategy could improve both knowledge acquisition and patient care. Previous proposals to accelerate the development process have often been too ambitious, and remain unused. We propose three areas of research that have immediate practical implications for cancer clinical trials. The integrated Phase II/III design fits well into the current framework while at the same time offering the opportunity to improve it, and thus is more likely to become widely used. More informative and reliable endpoints for Phase II trials are needed, as typical response endpoints often fail to correlate adequately with survival and also do not fully apply to many modern therapeutic agents. Analytic methods that can more directly assess risks, benefits, and

effects on intended targets will increase the efficiency of trials and produce more informative reports. This investigation will provide a concrete demonstration of the worth of these innovative concepts, furthering knowledge in cancer research and treatment.

Summary of Research Completed

Aim 1: Previously, we evaluated the efficiency of the integrated phase II/III design in the setting of a multi-center cooperative group, compared to the traditional approach. During the past reporting period, we documented the findings in a manuscript which was submitted to Cancer Investigation. Unfortunately, the revised manuscript, which incorporated four reviewers' comments, was not accepted for publication due to not meeting one reviewer's expectation. During this current reporting period we revised the manuscript and resubmitted it to Clinical Trials, currently in review process.

A summary of results extracted from the manuscript are provided here. The Radiation Therapy Oncology Group (RTOG) is interested in testing whether bevacizumab added to lomustine is more beneficial than lomustine alone in the setting of recurrent temozolomide-resistant anaplastic glioma (AG). We describe in detail what a phase II/III study for the brain tumor example would look like. The primary endpoint would be progression-free survival (PFS) for the phase II portion and overall survival (OS) for the phase III portion. Based on the published data in this patient population, the median PFS and OS in this patient population with standard treatment are estimated as 5.1 and 14.6 months, respectively. Hunsberger's integrated II/III design has been proposed for this study. The calculations are based on a conservative sample size approach that assumes the PFS and OS endpoints are independent. There will be no accrual suspension between the phase II and III components. For the PFS endpoint we consider an increase in median PFS to 8.4 months sufficiently active to warrant a phase III study. Therefore we will power the phase II portion of the study to be able to detect a hazard ratio of 0.61. For the OS endpoint, we will consider an increase in OS to 19.4 months to be clinically relevant. Therefore, we will calculate the sample size to have high power to detect a hazard ratio of 0.75.

The total sample size for the II/III study is 675 patients (532 deaths). This sample size is based on the following assumptions and parameters: the deaths will follow an exponential distribution, the accrual rate will be 22.5 patients per month, the final test will be performed at the 1-sided .025 level, there will be 90% power to a detect hazard ratio of .75 (without the stage II PFS analysis), there will be a minimum follow up of 19 months after trial closure (30 months trial accrual). The phase II futility analysis will take place after 225 patients have been accrued (102 deaths). If the 1-sided p-value is less than .2 the study will continue to the full sample size. If the p-value is greater than 0.2, the trial will be terminated for insufficient evidence of activity with respect to PFS. The sample size for the phase II component was based on the following assumptions and parameters: PFS will follow an exponential distribution and there will be 95% power to detect a hazard ratio of .61.

In contrast, this trial could also be designed in the traditional way, such as a randomized screening phase II, followed by a randomized phase III trial if applicable. We investigate the possible benefits and drawbacks of using an integrated phase II/III design. The integrated phase II/III design leads to a relative 25% reduction on maximum of required patients compared to

separate phase II and III trials design. In addition, the expected sample size under the global null hypothesis is smaller for the integrated phase II/III design. For the integrated phase II/III design, there will be great time-saving from an administrative perspective because there will be no gap between phase II and III components, which generally takes a minimum of 1.5 years. However, extra time is needed to set up phase II/III study at the beginning of phase II/III trials due to extra complexity. In addition, it will need to be a multicenter study (extra time to set up) instead of a single or fewer centers. In contrast, for the traditional design, there exists gap between completion of phase II and launch of phase III for administrative time, and additional gap due to minimum follow up for data to mature in phase II (if applicable). With regard to study duration (from study activation of phase II component to the final analysis of phase III component if the phase II supports the launch of phase III), the whole study duration is between 4 and 4.5 years for the integrated phase II/III design, while it runs to between 7 and 8 years for the traditional design. Therefore, the integrated phase II/III design leads to a reduction of 3 years on duration of this trial compared to the traditional design.

<u>Simulation: The impact of correlation between endpoints on operating characteristics</u>
The phase II/III seamless design aims to add logistical efficiency and maximize the contribution of patient information. A critical aspect of the statistical efficiency of this design is the degree to which the effect of treatment on the Phase II endpoint can reliably predict the treatment effect on the Phase III endpoint. For simplicity here, we assume that if the two endpoints are correlated, then the treatment effect is manifest in both endpoints. We examined the impact of correlation between endpoints for the two trial stages on type I error, power, study duration and expected sample size. We considered the endpoints of PFS for phase II and OS for phase III, and assumed that individual OS and PFS follow a bivariate exponential distribution indexed by three parameters (θ_x, θ_y, δ):

$$S(X,Y) = \Pr(X > x, Y > y) = \exp(-((x/\theta_x)^{1/\delta} + (y/\theta_y)^{1/\delta})^{\delta})$$

Where θ_x and θ_y denote the scale parameter for X, Y, respectively; δ introduces a non-negative correlation between X and Y, which can be expressed in a function of θ_x, θ_y and δ; here, X denotes OS and Y denotes PFS. For our purposes, if X>Y, OS=X and PFS=Y; Otherwise, OS=PFS=X.

To evaluate the impact of correlation between endpoints under a) no treatment effect on either PFS or OS (global null); b) treatment effect on both PFS and OS (global alternative); c) treatment effect on PFS, but not on OS (PFS only alternative) and d) treatment effect on OS, but not on PFS (OS only alternative), θ_x, and θ_y, (inverse of hazard rates) were adopted from the AG trial described above. Four values of association parameters were studied: δ = 0.2, 0.45, 0.7, and 1. Accounting for the influence of replacing PFS with OS when the generated PFS was greater than the generated OS, and the impact of censoring on the correlation between PFS and OS, the four selected values approximately represented Pearson correlation (ρ) of 0.9, 0.7,0.54, and 0.33, respectively. To measure the operating characteristics of the design under the above four scenarios with varying δ, these four trial outcome summaries were obtained based on 10000 simulations per scenario: probability of going to phase III, probability of claiming positive

phase III, study duration and expected sample size. The first stage accrual of 10 months for the phase II component, additional accrual of 20 months for the phase III component, monthly accrual of 22.5 patients, and minimal follow-up of 24 months for patients in the stage III accrual were used for the simulations. The simulation results were presented in the Table 1. The simulation results demonstrated: 1) under the global null, as expected the significance level was maintained; 2) under the global alternative, a high correlation led to slightly higher probability of going to phase III and slightly higher probability of positive phase III claims; 3) there was little influence on study duration and expected sample size with varying correlations under either global hypothesis.

Aim 2: Quantile comparison in randomized Phase II trials and Phase III trials: For this project, while the analytic work is now largely complete, we have not initiated simulations yet. The other major activity involves identifying example clinical trials from the RTOG database to illustrate the method. Work on this task is continuing.

Aim 3: Competing Risks Regression Models: For this project, simulation studies examining different competing risks scenarios with respect to hazards for different events, correlation among event times, and amount of censoring were completed during the reporting period.

Some highlights of the work and contents of a draft manuscript are provided here. When cancer patients are followed after treatment for any failure event from among local recurrence, distant metastases, onset of second primary cancer, or death precluding these, it is understood that appropriate methods must be applied in order to obtain a correct estimate of the cumulative probability of each event. When comparing cause-specific outcomes between groups, one must also consider the influence of competing risks and choose the test appropriate for the question of primary interest. Here, we present simulation studies illustrating the behavior of these models under various realistic scenarios and an example from a clinical trial of the Radiation Therapy Oncology Group.

Regression Models for Competing Risks Data

Modeling cause-specific hazards

The familiar Cox proportional hazards model is readily adapted to modeling cause-specific hazards (CSH). The model has the form $\lambda_k(t) = \lambda_{0k}(u)\exp(X\beta)$. Hazard ratio estimates from this model are largely interpreted in the same way as in the absence of competing risks.

Modeling cumulative incidence

A similar formulation of the Cox model can be used for a quantity known as the 'subdistribution hazard (SDH)'. Heuristically, the subdistribution hazard λ^*_k can be thought of as the hazard for an individual who either fails from cause k or does not, and in the latter case has an infinite failure time for cause k. While this may seem unusual, indeed in the case of mutually exclusive event types, those who fail from one cause are invulnerable to failure from others. This mathematical construct allows modeling of covariate related to the cumulative incidence of a specific event type. The model has the form $\lambda^*_k(t) = \lambda^*_{ok}(u)\exp(X\beta)$. Interpretation is similar to that of the cause-specific hazard Cox model, but there are additional important considerations.

A number of other approaches have been proposed, including 'constructed' estimates from Cox models for all cause-specific hazards and partially or wholly parametric models. Klein proposed using generalized linear models adapted for censoring.

Simulation Study of Competing Risks Regression Models

To investigate circumstances where results of the modeling approaches may differ, we simulated competing risks observations and computed the estimators from the two commonly used models. After specifying parameters for a bivariate survival distribution, we generate the time to a designated primary event X_1 and a competing event X_2. We also generate an independent censoring time C. Then we obtain T = minimum(X_1,X_2,C) and an event type indicator to form the competing risks observations. We compute the Cox cause-specific hazard ratio and subdistribution hazard ratio for both event types. We average these estimates over a large number of runs to illustrate behavior of the models under various scenarios of interest. For these simulations, correlation between failure times was set at zero, implying that failure hazards from the two causes are unrelated. However, even when there is dependence among times, it cannot be measured in competing risks data, since at most only one of the failure times is observed.

Scenario II – (Table 2) In this scenario, only Event 1 is related to group membership, with the hazard in Group A equaling one-half that of Group B. For the Cause-Specific Hazard (CSH) model, the Event 1 hazard ratio is approximately 2.00 as expected, while the Event 2 hazard ratio is 1.00. For the Subdistribution Hazard (SDH) model, there is also an influence of group for Event 1, with a subdistribution hazard ratio of 1.79, somewhat attenuated from that of the CSH model. However, for Event 2, the SDH model indicates that Group B is *less* likely to fail (Group B/A SDH ratio = 0.75). Since the hazards for Event 2 are identical, this is due to influence of Event 1, for which fewer Group A subjects fail, leaving them available to fail from Event 2.

Scenario III – In this case, Group A has substantially lower hazard for Event 1 and moderately lower hazard for Event 2, and the CSH model shows these effects. The Event 1 SDH ratio is attenuated while the Event 2 SDH ratio is null. This is again due to the fact that the SDH model reflects the influence of the competing event. With relatively more individuals available to fail from Event 2 in Group A (because they fail less from Event 1), the cumulative incidence of Event 2 (and thus the SDH ratio) is similar between Group A and Group B.

Analysis by Causes of Death after Treatment for Localized Prostate Cancer

RTOG 8610 is a randomized trial comparing radiation therapy alone to radiation therapy plus hormone deprivation therapy after curative surgery for localized prostate cancer. Men with prostate cancer may benefit from hormone therapy via reduction in risk of local recurrence and distant metastases, which account for most prostate cancer deaths. On the other hand, hormone depletion can have negative consequences with respect to other diseases, thus having the potential to increase other cause mortality. Age at diagnosis is important with respect to both the behavior of the prostate tumor and risk for non-cancer deaths. A third important covariate is tumor cell differentiation, which is strongly related to prostate cancer death but should not have any direct bearing on non-cancer deaths. We examine these factors from the perspective of modeling the cause-specific hazard and the cumulative incidence of prostate cancer and other-cause deaths.

The addition of hormones reduces the hazard of prostate cancer deaths (Table 3) by approximately 35% (CSH model). The influence of treatment on the cumulative incidence of prostate cancer death is similar in this case (Fine-Gray SDH model). For age, the cause-specific hazard model indicates a statistically non-significant 9% decrease in failure risk per 10 years of increased age. However, the cumulative incidence regression indicates a statistically significant 23% reduction in cancer deaths per 10 years of increased age. This is likely largely due to the fact that with increasing age, risk of death from causes other than cancer increases greatly, even among cancer patients. Differentiation is strongly associated with prostate cancer death hazard and cumulative incidence. Results for the Klein-Anderson model are similar.

For other-cause deaths (Table 3), a similar pattern emerges. Androgen Deprivation Therapy(ADT) nominally increases hazard of other-cause death, while for the cumulative incidence model, increased risk for other-cause deaths approaches statistical significance, with a 27% greater risk. If indeed ADT reduces prostate cancer deaths, as clearly indicated for that endpoint, then there will naturally accumulate more other-cause deaths among these patients. This is an important consideration when attempting to infer the extent to which treatment is a causative factor in other-cause deaths. Age at diagnosis is similarly related to cause-specific hazard of other-cause death. For differentiation, a 'protective' effect with differentiation suggested in the cumulative incidence models is most likely due to the strong association this factor has with prostate cancer death, and not with any direct influence on death from other causes.

Table 1: Simulation results showing the impact of correlation between endpoints on phase II/III design performance.

	δ (ρ)	Probability of going to phase III	Probability of claiming positive phase III	Expected sample size	Study duration (months)
Global Null	0.2 (0.9)	0.2019	0.0135	316	18.9
	0.45 (0.7)	0.1984	0.0132	314	18.7
	0.7 (0.54)	0.2019	0.0109	316	18.9
	1 (0.33)	0.207	0.0121	318	19.1
Global Alternative	0.2 (0.9)	0.9226	0.8488	640	50.6
	0.45 (0.7)	0.9137	0.8379	636	50.2
	0.7 (0.54)	0.9127	0.8325	636	50.2
	1 (0.33)	0.916	0.8371	637	50.3
PFS only alternative	0.2 (0.9)	0.9172	0.0416	638	50.4
	0.45 (0.7)	0.8714	0.0438	617	48.3
	0.7 (0.54)	0.8357	0.0424	601	46.8
	1 (0.33)	0.8198	0.0395	594	46.1
OS only alternative	0.2 (0.9)	0.203	0.199	316	18.9
	0.45 (0.7)	0.2269	0.2202	327	20.0
	0.7 (0.54)	0.2687	0.2536	346	21.8
	1 (0.33)	0.3236	0.3009	371	24.2

Table 2. Comparison of regression models under different competing risks scenarios, with 250 subjects per group. Competing risks data simulated from a bivariate exponential failure distribution with the hazard parameters indicated and independent censoring of about 33% of lifetimes. Averages of estimated parameters are based on 3000 simulated datasets.

				Model Estimates			
	Specified Hazards			**Cox**		**Fine and Gray**	
Scenario	**Group A**	**Group B**		**Cause-Specific Hazard Ratio**	**95% CI**	**Subdistribution Hazard Ratio**	**95% CI**
II: Event 1 rate lower in Group A							
Event 1	λ_{11} = 0.50	λ_{21} = 1.00		2.013	1.428-2.839	1.794	1.286-2.503
Event 2	λ_{12} = 1.00	λ_{22} = 1.00		0.999	0.750-1.331	0.750	0.567-0.991
III: Both event rates lower in Group A							
Event 1	λ_{11} = 0.50	λ_{21} = 1.00		2.002	1.451-2.763	1.291	0.956-1.742
Event 2	λ_{12} = 0.50	λ_{22} = 1.00		2.003	1.452-2.762	1.292	0.958-1.744

Table 3. Comparison of competing risks regression models examining treatment and two covariates for competing outcomes in prostate cancer (RTOG 8610) * per 10 year increment

	Model Effect Estimates					
	Cox CSH		**Fine-Gray SDH**		**Klein-Andersen**	
Event type (death)	**HR**	**95% CI**	**SDH**	**95% CI**	**?**	**95% CI**
A. Prostate Cancer						
ADT (vs RT only)	0.67	0.49 – 0.92	0.66	0.48 – 0.91	0.67	0.49 – 0.93
Age*	0.89	0.71 – 1.13	0.75	0.60 – 0.95	0.79	0.63 – 1.00
Grade 2 vs 1	1.84	1.04 – 3.23	1.83	1.05 – 3.17	1.87	1.06 – 3.31
Grade 3 vs 1	2.87	1.66 – 4.98	2.83	1.65 – 4.87	2.94	1.70 – 5.08
B. Other causes						
ADT (vs RT only)	1.13	0.85 – 1.51	1.26	0.95 – 1.68	1.20	0.89 – 1.61
Age	2.02	1.60 – 2.57	1.93	1.54 – 2.43	1.88	1.49 – 2.38
Grade 2 vs 1	0.87	0.59 – 1.28	0.75	0.52 – 1.08	0.82	0.56 – 1.20
Grade 3 vs 1	0.91	0.62 – 1.35	0.60	0.41 – 0.87	0.61	0.41 – 0.90
All deaths						
ADT (vs RT only)	0.88	0.71 – 1.09	-	-	-	-
Age	1.36	1.15 – 1.61	-	-	-	-
Grade 2 vs 1	1.13	0.83 – 1.55	-	-	-	-
Grade 3 vs 1	1.44	1.06 – 1.97	-	-	-	-

Research Project 2: Project Title and Purpose

Exploration of the RTOG Clinical Trial Database – Beyond Protocol-Specified Endpoints - For over 40 years, the Radiation Therapy Oncology Group (RTOG) has been funded by the National Cancer Institute (NCI) to conduct clinical trials seeking to improve the survival and quality of life of cancer patients. Drawing upon this vast resource of demographic, treatment, outcome, and patient-reported data, the researchers will develop hypotheses and explore correlations that were not defined in the treatment protocols for patients with brain, cervix, gastrointestinal, head and neck, lung, and prostate cancer. These analyses may lead to future protocols and/or better ways to identify high-risk subgroups and screen patients for specific treatment regimens.

Anticipated Duration of Project

1/1/2010 - 12/31/2013

Project Overview

RTOG investigators complete analyses and report on the endpoints specified in each NCI-approved protocol. Frequently these analyses raise questions or point to other potential hypotheses that were not included in the original protocol. Likewise, current literature and new research may point to areas of interest or possible correlations that were unknown during the design of the original protocol. The broad objectives of this research proposal are to (i) generate hypotheses and explore correlations that may lead to more efficient clinical trials and more patient-targeted treatments, and (ii) explore novel ways of analyzing the demographic (age, gender, race), treatment (including dose, volume, duration), outcome (survival, disease-free survival, time-to-progression), and quality of life (frequency/severity of adverse events, patient-reported outcomes) data in the RTOG database to potentially develop new tools for determining the best treatment regimen for each patient based upon their personal profile.

Principal Investigator

Kathryn A. Winter, MS
RTOG Director, Statistics
Radiation Therapy Oncology Group
American College of Radiology
1818 Market Street, Suite 1600
Philadelphia, PA 19103

Other Participating Researchers

Kyounghwa Bae, PhD, Daniel Hunt, PhD, Jonathan Harris, MS, Jennifer Moughan, MS, Rebecca Paulus, MS – employed by American College of Radiology

Expected Research Outcomes and Benefits

The identification of pre-treatment patient characteristics and demographics associated with better or worse outcome for cancer patients may allow future researchers to generate new hypotheses to address outcome disparities due to age, race, ethnic origin or gender. Exploring the radiation therapy dose volume histogram data in more detail will help with better definitions of dose constraints in future trials. In addition to aiding in the conduct of clinical trials, this type of research may help to tailor treatments to individual patients based on their demographic and/or treatment characteristics profile.

Summary of Research Completed

During this reporting period, work has been done on 3 analyses exploring the RTOG clinical trials database. Details are listed below.

Analysis # 1

This analysis evaluated the impact of per protocol defined PSA complete response (PSA-CR), defined at the end of external beam radiotherapy (EBRT) and short-term hormonal therapy (STHT), on treatment outcomes for patients treated on RTOG 9413 *"A Phase III Trial Comparing Definitive Whole Pelvic Irradiation Followed by a Conedown Boost to Boost Irradiation Only and Comparing Neoadjuvant to Adjuvant Total Androgen Suppression (TAS)"* and was presented at the 2010 Annual American Society for Radiation Oncology (ASTRO) Meeting in November 2010.

Abstract: In RTOG 9413, STHT consisted of flutamide 250 mg three times a day plus a luteinizing hormone releasing hormone (LHRH) agonist for a total of 4 months and, per protocol, a PSA measurement was required at the end of treatment. PSA-CR was defined as undetectable if serum PSA levels ≤0.3 ng/ml. All patients treated on RTOG 9413 for whom post treatment PSA data was available are included in this analysis. The Kaplan-Meier method was used to estimate overall survival (OS) and disease-free survival (DFS). Cumulative incidence was used to estimate the endpoints of biochemical failure (BF), distant metastasis (DM), and disease-specific survival (DSS). Univariate and multivariate analyses were performed to correlate PSA-CR after TAS with all endpoints, and the following variables were considered for analysis: PSA at baseline, Gleason score, treatment arm, age, and baseline testosterone. Phoenix consensus was used to define PSA failure.

For 1069 evaluable patients, the median PSA at the end of STHT was 0.2 ng/mL. A total of 744 patients (69.6%) had a PSA-CR per the protocol definition. At a median follow-up of 7.2 years, failure to obtain a PSA-CR was significantly associated with worse DSS (hazard ratio (HR)=1.95; 95% CI [1.33 - 2.85]) and DFS (HR=1.28; 95% CI [1.09 - 1.50]), as well as with a higher incidence of DM (HR=1.92; 95% [CI, 1.37 - 2.69]) and BF (HR=1.58; 95% CI [1.30 - 1.92]). The other factors significantly associated with worse DSS were Gleason score ≥8 (HR=3.06; 95% CI [1.72 - 5.47]) and baseline PSA >20ng/mL (HR=1.55; 95% CI [1.03 - 2.32]).

Failure to obtain a post STHT and EBRT PSA-CR, (≤0.3 ng/mL) appears to be an independent predictor of unfavorable outcomes, and may help identify patients who might benefit from the addition of long-duration androgen ablation.

Analysis # 2

This analysis evaluated associations between V10 and V20 (volume of bone marrow receiving at least 10 and 20 Gy respectively) and hematologic toxicities for patients treated on RTOG 0418 *"A Phase II Study of Intensity Modulated Radiation Therapy (IMRT) to the Pelvis +/- Chemotherapy for Post-operative Patients with either Endometrial or Cervical Carcinoma"* and was presented at the 2010 Annual ASTRO Meeting in November 2010.

Abstract: Intensity Modulated Radiation Therapy (IMRT) can reduce the volume of bone marrow irradiated as compared to conventional 4-field treatment. Pelvic bone marrow sparing has been reported to produce a clinically significant reduction of hematologic toxicity and increase the number of cycles of chemotherapy delivered in single institution retrospective studies. This analysis investigates hematologic toxicity on RTOG 0418, a prospective study, in a multi-institutional setting.

Patients on RTOG 0418 were treated with post-operative IMRT to 50.4 Gy to the pelvic lymphatics and vagina. Endometrial cancer patients received IMRT alone, while patients with cervical cancer also received weekly cisplatin (40 mg/m^2). Hematologic toxicity during concurrent chemoradiation was compared to patients treated with conventional fields and chemoradiation on RTOG 9708 in which cisplatin (50 mg/m^2) was delivered on weeks 1 and 5. Hematologic toxicity which occurred during subsequent adjuvant chemotherapy on RTOG 9708 was excluded for this analysis. Pelvic bone marrow was defined within the treatment field using a CT-density based auto-contouring algorithm. V10 and V20, dosimetric parameters correlated with hematologic toxicity (HT) in previous studies, were determined for 39 patients (30 with endometrial and 9 with cervical cancer). Hematologic toxicity was graded using the Common Terminology Criteria for Adverse Events (CTCAE) v 3.0 criteria. Chi-squared tests were used to compare proportions between studies and to evaluate associations between toxicity and dosimetric variables.

Eighty-three patients were eligible for analysis (43 patients with endometrial cancer and 40 patients had cervical cancer). Patients with cervical cancer treated with weekly cisplatin and pelvic IMRT on RTOG 0418 had (23%, 33%, 25%, 0%, 0%) grade 1-5 hematologic toxicity. Ninety percent of patients with cervical cancer received at least 4 cycles of cisplatin. While there was no difference in grade 3 or greater toxicity between the cervical arm of RTOG 0418 and RTOG 9708 (25% vs. 31%, p=NS), there was significantly less grade 4 or higher toxicity on the cervical arm of RTOG 0418 than RTOG 9708 (0% vs. 18%, p=0.002), despite the more frequent weekly chemotherapy on 0418. The mean V10 for evaluated patients was 88.0 (min-max: 75.8% to 96.3%) and mean V20 was 74.3 (min-max: 56.9%-90.8%). Thirty-nine percent (11/28) of patients with V10>85% had Grade 1 or worse hematologic toxicity compared to 27% (3/11) of patients with a V10 ≤85%.

Pelvic IMRT with weekly cisplatin is associated with lower rates of hematologic toxicity than historical controls treated with conventional 4-field plans. Although all patients were treated

with IMRT to the same anatomical target, the extent of bone marrow sparing varied widely, suggesting that pelvic IMRT can be optimized to limit bone marrow dose. Expanded analysis of this cohort will focus on identifying the dosimetric parameter which best predicts hematologic toxicity.

Analysis # 3

This analysis evaluated the impact of prophylactic cranial irradiation (PCI) on combined neurocognitive functioning and quality of life (QOL) for non-small cell lung cancer (NSCLC) patients treated on RTOG 0214 *"A Phase III Comparison of Prophylactic Cranial Irradiation versus Observation in Patients with Locally Advanced Non-Small Cell Lung Cancer"* and for small-cell lung cancer (SCLC) patients treated on RTOG 0212 *"A Phase II/III Randomized Trial of Two Doses (Phase III-Standard Vs. High) and Two High Dose Schedules (Phase II-Once Vs. Twice Daily) for Delivering Prophylactic Cranial Irradiation for Patients with Limited Disease Small Cell Lung Cancer (Companion study to the International Cranial Irradiation Trial, PCI 01-EULINT1)"* and was presented at the 2011 Annual American Society of Clinical Oncology (ASCO) Meeting in June 2011.

Abstract: This analysis focused on the association of deterioration (D) in patient-reported quality of life (QOL) with survival following curative-intent treatment for limited-disease small cell lung cancer (LD-SCLC) and locally advanced non-small cell lung cancer (LA-NSCLC).
Patients with LD-SCLC and LA-NSCLC were enrolled on RTOG trials 0212 and 0214, respectively. QOL was assessed as a secondary endpoint using the European Organization for Research and Treatment of Cancer - Quality of Life Cancer Questionnaire (EORTC QLQ-C30) and the brain cancer module (BN20) at baseline (following loco-regional therapy but before PCI or observation) and at 6 and 12 months follow-up. Deterioration was defined compared to baseline using the reliable change index method and correlated with overall survival and the cumulative incidence of brain and distant metastases using the Cox proportional hazards and Fine and Gray's proportional hazards models, respectively. Covariates included age, gender, race, partner status, education, Zubrod performance score, histology, PCI use, and PCI dose.

Five-hundred and fifty-two (91%) of the 604 eligible patients accrued to RTOG 0212 and RTOG 0214 had baseline QOL assessments and were included in this pooled analysis. Compliance with QOL assessments (66% at 6 months and 54% at 12 months) did not differ between RTOG 0212 and 0214. After adjusting for baseline QOL score and covariates, deterioration in global health status/QOL at 6 months (adjusted hazard ratio (AHR)=2.41; 95% CI [1.70-3.43]; $p<0.0001$) and 12 months (AHR=2.22; 95% CI [1.39-3.56]; $p=0.0009$); fatigue at 6 months (AHR=1.71; 95% CI [1.20-2.43]; $p=0.003$) and 12 months (AHR=1.87; 95% CI [1.15-3.04]; $p=0.01$), and cognitive functioning at 12 months (AHR=1.88; 95% CI [1.18-3.00]; $p=0.008$) were all associated with worse overall survival, compared to no deterioration/improvement. Deterioration in global health status/QOL (AHR=1.82; 95% CI [1.28-2.59], $p=0.0009$) and fatigue (AHR=1.61; 95% CI [1.15-2.27], $p=0.006$) at 6 months were also associated with increased risk of distant metastasis.

Longitudinal deterioration in global health status/QOL, fatigue and cognitive functioning at 6 and 12 months after definitive treatment for LD-SCLC or LA-NSCLC is potentially an early surrogate marker for poor overall survival.

Research Project 3: Project Title and Purpose

Emerging Imaging Technology Clinical Trials in PA: Comparison of Full Field Digital Mammography with Digital Breast Tomosynthesis Imaging: Comparison of Recall Rates - The purpose of this multi-center study, to be conducted as part of the American College of Radiology Imaging Network – Pennsylvania, is to evaluate the digital breast tomosynthesis screening recall rates compared to routine 2D projection digital mammography. The goal is to understand if a hybrid combination of 3D tomosynthesis and low dose 2D digital mammography can significantly reduce the recall rate of women from screening mammography without a concomitant reduction of sensitivity of cancer detection.

Anticipated Duration of Project

1/1/2010 - 12/31/2013

Project Overview

Previous C.U.R.E. funding established a network of 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. A multi-institutional clinical trial is proposed to evaluate the impact of breast tomosynthesis on the recall rate of screening mammography.

Study Hypothesis: Digital breast tomosynthesis (DBT) will improve the specificity of breast cancer screening as measured by a reduction in the recall rate while maintaining the sensitivity of cancer detection. This improved accuracy will be achieved by the optimization of the imaging sequence and number of views obtained at a capped radiation dose in the combined DBT and 2D screening sequence.

Principal Investigator

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

Other Participating Researchers

Emily Conant, MD – employed by University of Pennsylvania
Constantine Gatsonis, PhD – employed by Brown University

Expected Research Outcomes and Benefits

Screening mammography has been extensively criticized for the high rate of false positive interpretations, a subgroup of which is the recall of patients for additional diagnostic imaging for "pseudolesions" or superimpositions of normal tissue, perceived on screening mammography to be potentially significant lesions that on additional imaging prove to be normal. With competing parameters of specificity and sensitivity, mammographic screening must both limit missed cancers and reduce false positive call-backs. Tomosynthesis, a new emerging technology that allows the 3D reconstruction of images, has shown early evidence suggesting that it could significantly reduce the rate of false positive recalls from screening without a loss of sensitivity or breast cancer detection.

There are few published trials on breast cancer screening with tomosynthesis partly because the optimal procedural metrics for tomosynthesis have not been fully defined. Manufacturers have different platforms that offer different views, different angles, and different dose and exposure levels. The exact number of tomosynthesis views of the mediolateral-oblique (MLO) view only or both MLO and cranio-caudal (CC) tomosynthesis views varies while the screening imaging sequence with or without 2D digital mammography remains controversial. This disparity in image number and image acquisition parameters may alter the balance between specificity and sensitivity and significantly affect radiation dose. The expected outcome of this research is to show that the incorporation of tomosynthesis in the screening paradigm can reduce the number of false positive interpretations without a loss of cancer detection. This improvement in screening specificity must be gained while limiting both the number of imaging views and the radiation dose to the patient.

Summary of Research Completed

<u>Site Readiness</u>

September – October 2010

Site activation was still in progress during this time. Site activation included obtaining IRB approvals, finalizing site contracts and meeting the study-required physics qualification. Each site was required to perform a phantom image qualification and submission, as well as participate in a visit by the study physicist, before opening the study. During this time the trial team continued to finalize the study specific imaging manuals and the case report forms.

The trial team also finalized the process for the extraction of the Digital Imaging and Communications in Medicine (DICOM) data, the method of transference and the quality assurance process.

A site training session was held at ACRIN Headquarters on October 6, 2010 for the site Principal Investigators (PIs) and Research Associates (RAs.) The session consisted of training for the data management requirements including form completion guidelines and website review.

Transfer of Images and Data (TRIAD) software was installed at Albert Einstein Medical Center (AEMC).

November 2010 – June 2011

The Imaging manuals and qualification process were reviewed and finalized on December 10, 2010.

Site Activation

Albert Einstein Medical Center (AEMC) obtained IRB approval on 10/18/2010 and opened to enrollment on 11/18/2010. The site delayed opening from November 15th to November 18th was due to Informed Consent Form (ICF) issues and finalization of the qualification of the equipment.

The Hospital of the University of PA (HUP) received IRB approval on 9/20/2011 and opened the study on 2/15/2011. HUP was delayed in opening because of contract issues and the necessary equipment (Hologic) upgrades.

Study Enrollment – as of June, 30, 2011

The total enrollment for the Tomosynthesis Trial (ACRIN PA 4006) is 302/550 subjects with 170 subjects enrolled at the Hospital of the University of PA and 132 at the Albert Einstein Medical Center. The excellent progress shown in the table below is a result of strong patient willingness to participate on the trial. The protocol team meets monthly to review trial progress, assess quality and resolve issues.

Institution	Total Accrual
Hospital of the University of Pennsylvania	170
Albert Einstein Medical Center Hospital	132
Total accrual as of June 30,2011	302

Research Project 4: Project Title and Purpose

Investigation and Analyses of Patient Co-Morbidities in a Survey of Radiation Oncology Facilities in the USA and their Association with Treatment Decisions in Radiation Oncology - The purpose of this project is to describe the distribution of co-morbidities by socio-demographic characteristics such as age, race, geographic region, insurance status and socio-economic status in patients diagnosed with cancer of the breast, cervix, stomach, lung and prostate, to investigate the association of the prevalence of co-morbidities with treatment decisions and variations in compliance with recommended disease management guidelines for such patients, and to examine the interaction of co-morbidities by site and stage of disease with gender, race, and age.

Anticipated Duration of Project

1/1/2010 - 12/31/2013

Project Overview

Since 1973 the American College of Radiology (ACR) has conducted retrospective surveys of the processes of care in radiation oncology through the Quality Research in Radiation Oncology (QRRO, formerly Patterns of Care Study). Detailed information is collected from chart reviews on patient and tumor characteristics, imaging, treatment planning, surgery, radiation and systemic therapy with the purpose of measuring quality of care and comparing care actually received by patients to well-established clinical guidelines. These guidelines base treatment recommendations on tumor and patient characteristics, but provide little guidance on including patient co-morbidities in the treatment decision.

Although co-morbidities are not part of the scope of the QRRO study, the current data collection has included detailed data on co-morbidities for patients treated for breast, cervix, gastric and prostate cancers and non-small cell and limited stage small cell lung cancers. This project will investigate co-morbidity data in detail including interaction with other patient characteristics and association with treatment decisions.

SPECIFIC AIMS

1. To describe the distribution of co-morbidities by socio-demographic characteristics such as age, race, geographic region, insurance status and socio-economic status in patients diagnosed with cancer of the breast, cervix, stomach, lung and prostate.
2. To investigate the association of the prevalence of co-morbidities with treatment decisions and variations in compliance with recommended disease management guidelines for such patients.
3. To examine the interaction of co-morbidities by site and stage of disease with gender, race, and age.

Principal Investigator

Jean B. Owen, PhD
Senior Director

American College of Radiology
1818 Market St., Suite 1600
Philadelphia, PA 19103

Other Participating Researchers

Alex Ho, MS, MA, Najma Khalid, MS – employed by American College of Radiology

Expected Research Outcomes and Benefits

This project will provide new information on the effect of co-morbidities on treatment decisions for cancer patients and the interaction of co-morbidities with other patient and tumor factors. This will help fill a knowledge gap in the application of nationally recognized treatment guidelines that currently allow vague exceptions to the established standard of care for patients who have multiple confounding medical problems. By providing analyses that help increase understanding of the impact of co-morbidities on treatment decisions, this project will help improve the standard of care for these patients.

Summary of Research Completed

This project used data collected from a stratified sample of radiation oncology facilities in the United States, and then a random sample of eligible patients for each disease site study from each participating facility. Trained research associates performed on-site reviews of the medical records of selected cases and collected data on demographics, history, workup, staging, and treatment. The specific additional value for this project was the collection of data on patient co-morbidities. The evaluation of co-morbidities brought a new dimension to this survey that had not been included or reported in previous surveys of radiation oncology care. The Adult Co-morbidity Evaluation Index (ACE-27) was selected as the instrument to measure co-morbidities because of its clinical relevance and sensitivity. This index reflects not only a wide range of co-existing conditions relevant to cancer therapy choice and outcome, but also the severity of these conditions. When treatment was contra-indicated or changed due to co-morbidities, this was also recorded. The ACE-27 form and Piccirillo Index have been used in publications for other cancer sites and modalities.

Data from the 27-item Co-morbidity Data Collection Form were summarized into 12 body systems. Severity was recorded as Grade 1 -- Mild decompensation; Grade 2 -- Moderate decompensation, and Grade 3 -- Severe decompensation. The overall co-morbidity score was computed according to the highest ranked single ailment, except in the case where two or more Grade 2 ailments occurred in the different organ systems. In that situation, the overall co-morbidity score was designated as Grade 3.

Computation of weights

Weights were computed to reflect the relative contribution of each patient in the analysis of this complex survey. The weights for each stratum within a disease site were the product of two factors:

$$\frac{\textit{Estimated number of eligible cases in population}}{\textit{Number of eligible cases in sample}} \quad X \quad \frac{\textit{Proportion of eligible cases in the population}}{\textit{Proportion of eligible cases in the sample}}$$

Statistical Analyses

Univariate and bivariate analyses and tests of inference were performed using SAS v9.2. National estimates were calculated from the survey data using SUDAAN statistical software (Research Triangle Institute, Research Triangle Park, NC), which incorporates the design elements and weights that reflect the relative contribution of each patient in the analysis of this complex survey. The strength of the associations between the overall co-morbidity score and the treatment given, whether treatment was completed or not, treatment duration, and demographic variables will form the bases for building more complex multivariate models.

For each study, distributions of co-morbidities for each disease site were computed for each of the 12 body systems and for the overall co-morbidity score. More detailed analyses investigated the association of co-morbidities by stage of disease, treatment variables, treatment dose, and demographic variables. In some analyses, the categories for no co-morbidities and mild co-morbidities were collapsed into one category.

In April 2011, the results of analyses done for three disease sites, Lung, Gastric, and Breast, were prepared as abstracts and submitted for presentation to the 2011 Annual Meeting of the American Society for Therapeutic Radiation Oncology (ASTRO) and all three were accepted. Each lead author had different approaches to the investigation of co-morbidities with other treatment characteristics.

Summary of progress toward milestone(s) for 7/1/2010-6/30/2011:

1) Compute weights to reflect sample design.
Status -- This task was completed.
2) Compute descriptive statistics for co-morbidity data elements for each study.
Status -- This was completed for all studies.
3) Compute multivariate statistics to investigate the interaction of co-morbidity and other patient and tumor factors.
Status -- Investigations of the associations of co-morbidities with stage and treatment characteristics were completed. The results of these analyses form the bases for building more detailed multi-variate models to study more complex interactions. This is in progress.
4) Prepare abstracts for presentation of scientific results at national meetings.
Status -- Abstracts were prepared for three disease sites, Gastric, Breast and Lung cancers, submitted for presentation at the 2011 Annual meeting of the American Society for Therapeutic Radiation Oncology (ASTRO), and accepted.
5) Prepare manuscript of results for first disease site. Status --The manuscript based on the abstract for Stage III NSCLC is in preparation but not yet completed.