# Final Progress Report for Research Projects Funded by Health Research Grants

Instructions: Please complete all of the items as instructed. Do not delete instructions. Do not leave any items blank; responses must be provided for all items. If your response to an item is "None", please specify "None" as your response. "Not applicable" is not an acceptable response for any of the items. There is no limit to the length of your response to any question. Responses should be single-spaced, no smaller than 12-point type. The report **must be completed using MS Word**. Submitted reports must be Word documents; they should not be converted to pdf format. Questions? Contact Health Research Program staff at 717-783-2548.

1. **Grantee Institution:** American College of Radiology

2. **Reporting Period (start and end date of grant award period**): 1/1/2010-12/31/2013

3. **Grant Contact Person (First Name, M.I., Last Name, Degrees):** Stephen M Marcus, MS

4. **Grant Contact Person's Telephone Number:** 267-940-9403

5. **Grant SAP Number:** 4100050889

6. **Project Number and Title of Research Project:** #1-Novel Methods for Cancer Clinical Trial Design

7. **Start and End Date of Research Project**: 1/1/2010 – 6/30/2013

8. **Name of Principal Investigator for the Research Project:** James Dignam, PhD

9. **Research Project Expenses**.

   9(A) Please provide the total amount of health research grant funds spent on this project for the entire duration of the grant, including indirect costs and any interest earned that was spent:

   $ 284,297.67

   9(B) Provide the last names (include first initial if multiple individuals with the same last name are listed) of **all** persons who worked on this research project and were supported with health research funds. Include position titles (Principal Investigator, Graduate Assistant, Post-doctoral Fellow, etc.), percent of effort on project and total health research funds expended for the position. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

| Last Name, First Name | Position Title | % of Effort on Project | Cost |
|---|---|---|---|
| | | | |
| Wang | Senior Statistician | 5% yr1; 9% yr 2; 8% yr 3 | $33,362.84 |
| Harris | Biostatistician III | 4% yr1;9% yr2; 8% yrs 3-4 | $39,647.71 |
| Boparai | Project Mgr | 14% yr2; 10% yr3 | $27,003.85 |
| Hu | Senior Statistician | 8% Yr 4 | $10,065.98 |
| Presley | Dosimetrist | 1% | $1,509.94 |
| Seiferheld | Biostatistician II | 57% Yr 4 | $74,797.05 |
| Bonanni | Senior Research Associate | 1% Yr 2 | $650.56 |

9(C) Provide the names of **all** persons who worked on this research project, but who *were not* supported with health research funds. Include position titles (Research Assistant, Administrative Assistant, etc.) and percent of effort on project. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

| Last Name, First Name | Position Title | % of Effort on Project |
|---|---|---|
| | | |
| Kocherginsky, Masha | Research Associate, University of Chicago | 10 |
| Dignam, James | Principal Investigator | <1 |
| | | |
| | | |
| | | |
| | | |

9(D) Provide a list of **all** scientific equipment purchased as part of this research grant, a short description of the value (benefit) derived by the institution from this equipment, and the cost of the equipment.

| Type of Scientific Equipment | Value Derived | Cost |
|---|---|---|
| None | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

**10. Co-funding of Research Project during Health Research Grant Award Period.** Did this research project receive funding from any other source during the project period when it was supported by the health research grant?

Yes_____ No_____X____

If yes, please indicate the source and amount of other funds:

**11. Leveraging of Additional Funds**

11(A) As a result of the health research funds provided for this research project, were you able to apply for and/or obtain funding from other sources to continue or expand the research?

Yes_____ No____X_____

If yes, please list the applications submitted (column A), the funding agency (National Institutes of Health—NIH, or other source in column B), the month and year when the application was submitted (column C), and the amount of funds requested (column D). If you have received a notice that the grant will be funded, please indicate the amount of funds to be awarded (column E). If the grant was not funded, insert "not funded" in column E.

Do not include funding from your own institution or from CURE (tobacco settlement funds). Do not include grants submitted prior to the start date of the grant as shown in Question 2. If you list grants submitted within 1-6 months of the start date of this grant, add a statement below the table indicating how the data/results from this project were used to secure that grant.

| A. Title of research project on grant application | B. Funding agency (check those that apply) | C. Month and Year Submitted | D. Amount of funds requested: | E. Amount of funds to be awarded: |
|---|---|---|---|---|
| | ☐NIH<br>☐ Other federal (specify:_____ _____)<br>☐ Nonfederal source (specify: _____) | | $ | $ |

11(B) Are you planning to apply for additional funding in the future to continue or expand the research?

Yes_____ No____X_____

If yes, please describe your plans:

**12. Future of Research Project.** What are the future plans for this research project?

The phase II/III trial design was intended to be implemented in an RTOG trial concurrently with the work within this project. Other circumstances prevented that, but we since have proposed it for trials currently in development. We will likely open a trial under this design shortly. We will continue to research advantages and opportunities to use this approach and variations on it. The conduct of trials under our proposal will allow empirical evaluation of performance.

The quantile work will be extended to estimation under statistical models that can incorporate covariates. Quantile regression is a rich research area that has not been adequately applied to survival time data.

The work on competing risks has led to pursuit of improved methods for conducting sample size calculations to better address the complexities of effect estimates in this situation. Specifically, one may take into account 'loss' of information due to competing events, and better account for it in sample size. One may also appropriately focus on cause-specific failure endpoints provided evaluation of other-cause failures is planned and adequately informative. We are currently working on sample size research in this area.

**13. New Investigator Training and Development**. Did students participate in project supported internships or graduate or post-graduate training for at least one semester or one summer?

Yes_____ No___X_____

If yes, how many students? Please specify in the tables below:

| | Undergraduate | Masters | Pre-doc | Post-doc |
|---|---|---|---|---|
| Male | | | | |
| Female | | | | |
| Unknown | | | | |
| **Total** | | | | |

| | Undergraduate | Masters | Pre-doc | Post-doc |
|---|---|---|---|---|
| Hispanic | | | | |
| Non-Hispanic | | | | |
| Unknown | | | | |
| **Total** | | | | |

| | Undergraduate | Masters | Pre-doc | Post-doc |
|---|---|---|---|---|
| White | | | | |
| Black | | | | |
| Asian | | | | |
| Other | | | | |
| Unknown | | | | |
| **Total** | | | | |

**14. Recruitment of Out-of–State Researchers**. Did you bring researchers into Pennsylvania to carry out this research project?

Yes_____ No___X_____

If yes, please list the name and degree of each researcher and his/her previous affiliation:

**15. Impact on Research Capacity and Quality**. Did the health research project enhance the quality and/or capacity of research at your institution?

Yes_____X_____ No_____

If yes, describe how improvements in infrastructure, the addition of new investigators, and other resources have led to more and better research.

The software development for these projects has provided tools for wider use among ourselves and others who may seek to collaborate on these methods. It has also motivated ideas for additional research in three areas described in the aims.

**16. Collaboration, business and community involvement.**

16(A) Did the health research funds lead to collaboration with research partners outside of your institution (e.g., entire university, entire hospital system)?

Yes_____ No___X_____

If yes, please describe the collaborations:

16(B) Did the research project result in commercial development of any research products?

Yes_____ No___X_____

If yes, please describe commercial development activities that resulted from the research project:

16(C) Did the research lead to new involvement with the community?

Yes_____ No___X____

If yes, please describe involvement with community groups that resulted from the research project:

**17. <u>Progress in Achieving Research Goals, Objectives and Aims.</u>**

List the project goals, objectives and specific aims (as contained in the grant agreement). Summarize the progress made in achieving these goals, objectives and aims <u>for the period that the project was funded (i.e., from project start date through end date).</u> Indicate whether or not each goal/objective/aim was achieved; if something was not achieved, note the reasons why. Describe the methods used. If changes were made to the research goals/objectives/aims, methods, design or timeline since the original grant application was submitted, please describe the changes. <u>Provide detailed results of the project.</u> Include evidence of the data that was generated and analyzed, and provide tables, graphs, and figures of the data. List published abstracts, poster presentations and scientific meeting presentations at the end of the summary of progress; peer-reviewed publications should be listed under item 20.

This response should be a <u>DETAILED</u> report of the methods and findings. It is not sufficient to state that the work was completed. Insufficient information may result in an unfavorable performance review, which may jeopardize future funding. If research findings are pending publication you must still include enough detail for the expert peer reviewers to evaluate the progress during the course of the project.

Health research grants funded under the Tobacco Settlement Act will be evaluated via a performance review by an expert panel of researchers and clinicians who will assess project work using this Final Progress Report, all project Annual Reports and the project's strategic plan. After the final performance review of each project is complete, approximately 12-16 months after the end of the grant, this Final Progress Report, as well as the Final Performance Review Report containing the comments of the expert review panel, and the grantee's written response to the Final Performance Review Report, will be posted on the CURE Web site.

**There is no limit to the length of your response. Responses must be single-spaced below, no smaller than 12-point type. If you cut and paste text from a publication, be sure symbols print properly, e.g., the Greek symbol for alpha (α) and beta (ß) should not print as boxes (□) and include the appropriate citation(s). DO NOT DELETE THESE INSTRUCTIONS.**

<u>Project purpose and aims, as proposed in the research plan</u>

*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.

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.

<u>Progress made in achieving the aims</u>

*<u>Aim 1 Summary</u>*

This aim had two distinct goals: a) evaluate the operating properties (power, sample size, type I error control) of the integrated phase II/III clinical trial design in the oncology setting, compared

to the traditional approach, as well as more broadly evaluate the logistical efficiencies of a single trial protocol (e.g. phase II/III) versus sequential trials to ultimately give guidance on use of such a design and b) implement such a design in the Radiation Therapy Oncology Group (RTOG) brain tumor research portfolio (as need for such a design provided the initial motivation for this work). The first goal was met, as a detailed study was performed and results published. The second goal was not, as the specific trial (described below and used for comparative evaluation) was fully developed and approved by our main sponsor (the National Cancer Institute) but interrupted due to other circumstances. Specifically, the trial sponsor (Bristol-Meyers Squibb) at a very late date declined to provide their necessary support for this trial and so it did not commence. However, the work done in evaluation of this trial design has proven useful, as we have now have under development this same trial design in multiple other disease site settings settings within our research.

Background: Motivating example for Investigation of Integrated Phase II/III trial Designs

For the trial, we were interested in testing whether bevacizumab added to lomustine is more beneficial than lomustine alone in the setting of recurrent temozolomide-resistant anaplastic glioma (AG). In this trial following Hunsberger's integrated II/III design, the primary endpoint was progression-free survival (PFS) for the phase II portion and overall survival (OS) for the phase III portion. Based on published data in this patient population we were able to find, as well as prior experience in RTOG trials, the median PFS and OS with standard treatment were estimated as 5.1 and 14.6 months, respectively. Typical calculations for these trials are based on the conservative sample size approach that assumes the PFS and OS endpoints are *independent* [this provides an upper bound on the sample size (i.e., maximum it would need to be), hence the investigation of the effect of correlation]. For the PFS endpoint we considered an increase in median PFS to 8.4 months sufficiently active to warrant a phase III study. Therefore we sought to power the phase II portion of the study to be able to detect a hazard ratio of 0.61. For the OS endpoint, we considered an increase in OS to 19.4 months to be clinically relevant. Therefore, we sought a sample size that would have high power to detect a hazard ratio of 0.75. For the trial, there would be no accrual suspension between the phase II and III components.

Addressing the desired effect sizes and power/error control considerations, 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: 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. If the 1-sided p-value is less than 0.2 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 0.61.

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. A randomized phase II screening design (e.g., Rubinstein et al 2005) with type I and II errors of 0.2 and 0.05,

respectively, would require 225 patients (102 events) to detect the PFS difference. It is important to note that the above power specification is not typical for randomized phase II trials, but we use this so that we can make comparisons of these two designs. If the phase II results support further testing through a phase III trial, in which the design parameters are kept the same as those in the above phase II/III design, 675 analyzable patients are needed for the phase III trial. Thus the total sample size would be 900 patients.

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 is on the order of a minimum of 1.5 years. However, some extra time is needed to develop the II/III study at the outset due to extra complexity. In addition, the entire study (phase II included) 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 a gap between completion of phase II and launch of phase III for administrative time, and an 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, while these parameters are approximate, the integrated phase II/III design could lead to a reduction of 3 years on trial duration compared to the traditional design.

<u>Simulation study: 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 ($\theta_x$, $\theta_y$, $\delta$):

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

Where $\theta_x$ and $\theta_y$ denote the scale parameter for X, Y, respectively; $\delta$ introduces a non-negative correlation between X and Y. 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), the parameters $\theta_x$ and $\theta_y$ (inverse of hazard rates) were adopted from the AG trial described above. Four values of association parameters were studied: $\delta$ = 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 ($\rho$) 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 $\delta$, 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 are presented in the Table 1. Some highlights of results that were demonstrated: (1) under the global null, as expected, the significance level was maintained; (2) under the global alternative, a high correlation led to slightly greater probability of continuing to phase III and slightly higher probability of positive phase III claims; (3) under either global hypothesis, there was little influence on study duration and expected sample size over a range of correlations; (4) under the PFS-only alternative, a high correlation led to a higher probability of going to the phase III, longer study duration, and larger expected sample size, but the overall significance level was maintained; and (5) under the OS-only alternative, a high correlation led to a lower probability of going to the phase III and lower probability of positive phase III claims, shorter study duration, and smaller expected sample size.

Table 1 (Aim 1 – Phase II/III Design): Simulation results showing the impact of correlation between endpoints on phase II/III design performance.

| | **Correlation $\delta$ parameter (implied $\rho$)** | **Prob(initiate phase III)** | **Prob( claim positive phase III)** | **Expected sample size (no. of pts)** | **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 effect 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 effect 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 |

The simulations tended to confirm (not surprisingly) that if the treatment is truly ineffective (on either endpoint), then there is little risk in the design with respect to inflating type I error, under any correlation structure. What was less expected was the small effect of correlation on study duration and expected sample size, but in fact we found other researchers reporting similar findings in some related multiple endpoint or multistage trials (Bryant and Day 1995, Royston et al 2003, 2011).

Conclusions

In the published manuscript (publication 1, page 26), we summarized these results, and also included a brief review of the existing literature on so-called adaptive designs, which encompass phase II/III transition designs. We also provided practical guidance on when one might favor a phase II/III transition design in terms of practical circumstances that should be in place. Briefly, one must be committed to the resources needed for the entire trial (phases II and III) from the outset. On the other hand, the design does differ from simply having a phase III trial with a futility stopping endpoint, in that based on a non-primary endpoint, the trial may not continue. Secondly, external circumstances that may eventually favor other phase III candidates outside the trial should not be obviously available. If either condition is not met, then the entire development scheme for this trial type will be disrupted. We are actively implementing this design in trials in development wherever appropriate.

After some difficulty moving the submission through the editorial office at *Journal of Clinical Investigation* in 2010 (due to editorial circumstances beyond our control, as the reviewer comments seemed addressable), we were very pleased instead to submit to (12/2010) - and publish the work in - *Clinical Trials*, the official journal of the Society for Clinical Trials and the premier forum for clinical trials methodology discussions.

*Aim 2 Summary*

The aim of this study was to develop and test new methods of comparing summary outcome measures in survival analysis. This aim was met in that we have fully developed and tested a procedure for computing the median difference in survival between two groups, as described below. However, the manuscript summarizing this work has yet to be submitted for publication, because the development work was lengthy and a valuable illustrative example has only now become available after publication of the primary results. Publication of the method remains an important goal that we intend to achieve in 2014.

Background: Problem and Method

To motivate the use of quantile estimation and testing in time to event (i.e., survival) data, we focused initially on the median survival time, as in clinical trials, interest is often focused on the median time to event as an easily interpretable summary metric. Generation of one-sample confidence intervals for the median failure time has received considerable attention in the literature but less attention has been given to interval estimation of the difference in medians between two groups. Because in oncology we are now conducting two-sample comparative trials (Rubinstein et al 2005) in the phase II setting with much greater frequency (as opposed to traditional single-arm phase II trials), we propose a nonparametric method of comparing median survival between two treatment groups based on a multiple imputation approach. There have been a few proposals for this problem, (Su and Wei (1993), Wang and Hettmansperger (1990)), that require programming the method or obtaining software. In contrast, the proposed approach is simpler and can be implemented without specialized computer software.

To describe the approach, complete survival times for censored observations are first imputed by drawing from the conditional distribution given survival up to the time of censoring. The difference in medians and its variance for the complete data are estimated using the formulae in Price and Bonett (2002), and Rubin's multiple imputation approach (1987) is used to obtain the total variance from which confidence intervals are derived. The idea is rather simple - first, impute survival times for each censored observation conditional on the Kaplan-Meier estimates $\hat{S}_j(t)$ in each group j = 1, 2, and then use complete data methods to construct the confidence interval for the difference in medians. Using the multiple imputation approach of Rubin (1978), each censored observation is imputed M times. Estimates from these M complete data sets are then combined to obtain a point estimate of the median and its associated variance estimate.

Simulation Study Evaluating Performance of the Method

A simulation study was performed to examine the performance of this method under several scenarios. We considered a range of small to moderate sample sizes ($n1 + n2 = 50, 100, 150$) with both equal and unequal sample sizes in the two groups, $n1/n2 \in (1, 2)$. True survival times $X_i$ were generated independently in each group from exponential (Exp) and Weibull survival distributions. The censoring random variable was independently generated from $C_i \sim \text{Exp}(\lambda)$, and survival times $Y_i$ were obtained by censoring the "true" survival times $X_i$ according to $Y_i = \min(X_i, C_i)$. The censoring indicator was determined as $\delta_i = I(X_i < C_i)$. For each scenario under consideration, $\lambda$ was chosen to be such that 10%, 25% or 40% of the observations were censored in each group. K = 1000 independent data sets were generated, and M = 10 sets of imputations were done.

Data were first generated under the null hypothesis, that is, survival times in both groups were generated using the same distribution. We considered Exp(1),Weibull(1.5, 1) and Weibull(0.45, 1). Under the alternative hypothesis, data were generated from $X_{i1} \sim \text{Exp}(1), X_{i2} \sim \text{Exp}(2)$ and $X_{i1} \sim \text{Weibull}(2.5, 0.75), X_{i2} \sim \text{Weibull}(1, 0.6)$. Generating data forms are shown in Figure 1.

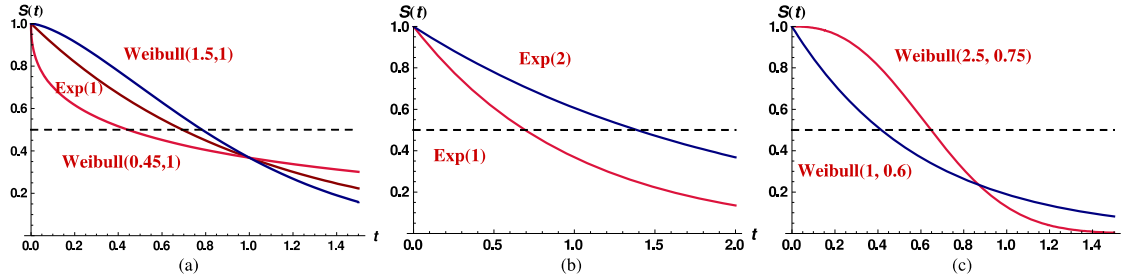


Figure 1 (Aim 2 – Quantiles for Survival Data): The data generating scenarios considered for the simulation of median difference estimation.

Rejection rates and coverage probabilities are reported in Table 2. Rejection rates for the first and third scenarios (Exp(1) and Weibull(1.5, 1)) are fairly close to the $H_0$ target $\alpha = 0.10$ and $\alpha = 0.05$, but were quite conservative for the second scenario (Weibull(0.45, 1)). Coverage probabilities for the two scenarios that were considered were also fairly close to the $H_A$ nominal 90% and 95% rate.

Table 2 (Aim 2 - Quantiles for Survival Data): Results from simulation study of proposed method. Data were simulated as described in the text under different survival models and the rejection rates (alpha levels) and coverage probabilities (for difference in median estimates) computed.

| | | | Under $H_0$ (Rejection Rates) | | | | | | Under $H_A$ (Coverage Rates) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n1 | n2 | Cens. | Exp(1) | | Weibull(0.45, 1) | | Weibull(1.5, 1) | | Exp(1) vs. Exp(2) | | Weibull(2.5, 0.75) vs. Weibull(1, 0.6) | |
| | | Prop. | 10% | 5% | 10% | 5% | 10% | 5% | 90% | 95% | 90% | 95% |
| 25 | 25 | 0.10 | 0.085 | 0.043 | 0.050 | 0.024 | 0.109 | 0.053 | 0.885 | 0.937 | 0.889 | 0.927 |
| | | 0.25 | 0.078 | 0.038 | 0.046 | 0.020 | 0.080 | 0.043 | 0.901 | 0.951 | 0.902 | 0.949 |
| | | 0.40 | 0.074 | 0.034 | 0.062 | 0.038 | 0.093 | 0.047 | 0.894 | 0.933 | 0.913 | 0.950 |
| 34 | 16 | 0.10 | 0.106 | 0.054 | 0.056 | 0.028 | 0.093 | 0.045 | 0.909 | 0.947 | 0.898 | 0.947 |
| | | 0.25 | 0.090 | 0.053 | 0.059 | 0.034 | 0.102 | 0.056 | 0.905 | 0.941 | 0.919 | 0.950 |
| | | 0.40 | 0.095 | 0.054 | 0.103 | 0.061 | 0.098 | 0.057 | 0.866 | 0.917 | 0.896 | 0.941 |
| 50 | 50 | 0.10 | 0.066 | 0.031 | 0.054 | 0.016 | 0.089 | 0.049 | 0.908 | 0.957 | 0.907 | 0.956 |
| | | 0.25 | 0.075 | 0.032 | 0.052 | 0.028 | 0.091 | 0.047 | 0.905 | 0.954 | 0.916 | 0.956 |
| | | 0.40 | 0.075 | 0.032 | 0.050 | 0.025 | 0.081 | 0.042 | 0.921 | 0.960 | 0.922 | 0.965 |
| 67 | 33 | 0.10 | 0.118 | 0.056 | 0.078 | 0.032 | 0.097 | 0.049 | 0.887 | 0.929 | 0.880 | 0.933 |
| | | 0.25 | 0.077 | 0.034 | 0.048 | 0.025 | 0.097 | 0.057 | 0.919 | 0.962 | 0.899 | 0.947 |
| | | 0.40 | 0.092 | 0.041 | 0.065 | 0.037 | 0.070 | 0.041 | 0.888 | 0.939 | 0.912 | 0.947 |
| 75 | 75 | 0.10 | 0.081 | 0.040 | 0.064 | 0.032 | 0.075 | 0.032 | 0.923 | 0.960 | 0.913 | 0.960 |
| | | 0.25 | 0.058 | 0.027 | 0.060 | 0.030 | 0.080 | 0.040 | 0.920 | 0.963 | 0.926 | 0.959 |
| | | 0.40 | 0.061 | 0.028 | 0.040 | 0.019 | 0.078 | 0.037 | 0.903 | 0.951 | 0.915 | 0.958 |
| 100 | 50 | 0.10 | 0.085 | 0.048 | 0.058 | 0.025 | 0.097 | 0.049 | 0.898 | 0.940 | 0.927 | 0.950 |
| | | 0.25 | 0.082 | 0.035 | 0.047 | 0.018 | 0.087 | 0.041 | 0.898 | 0.947 | 0.911 | 0.955 |
| | | 0.40 | 0.079 | 0.037 | 0.058 | 0.026 | 0.089 | 0.031 | 0.913 | 0.956 | 0.918 | 0.961 |

Applications to Oncology Trials

Following methodological developments and simulations, we sought and developed clinical examples to illustrate the method. Specifically, we sought data with specific properties such as non-proportionality and early censoring (before the median) in order to challenge the new method and illustrate its utility. Two of these are presented briefly here, and the final additional and most important example is currently analyzed, but must await publication of the primary results (February 2014) before it can be included in a manuscript.

The first example is from NSABP R-01, a randomized trial in rectal cancer (Fig 2). The survival curves show reasonable proportionality that diminishes with time, and a minimal amount of censoring before the median difference, our target quantity for estimation. We focus on the subset of male patients, in whom a significant advantage for chemotherapy was noted. The median difference and associated variance were estimated by the proposed method and compared to two others [Wang and Hettmansperger (shift model) 1990 and Su and Wei (purely nonparametric) 1993]. As seen in Figure 2, the proposed method has a slightly narrower confidence interval length than the Su and Wei method, but like it includes zero. The Wang and Hettmansperger method has the narrowest length and does not include zero. The second example uses a much smaller dataset from a University of Chicago clinical trial of bevacizumab in mesothelioma, where there was a large difference in outcomes based on vascular endothelial growth factor (VEGF) status that was statistically significant according to the logrank test. Figure 3 shows the three methods applied to this dataset. Here, the proposed method has the narrowest confidence interval length.

The additional example, involving a brain tumor trial currently in press (data from Gilbert et al 2014), involves much more extensive censoring before the median, one of the principal motivations for derivation of this method. Preliminary results on the third example again indicate good performance. The proposed estimator has a confidence interval that is about 5% narrower than the Su and Wei method as well as the property of being easier to compute.

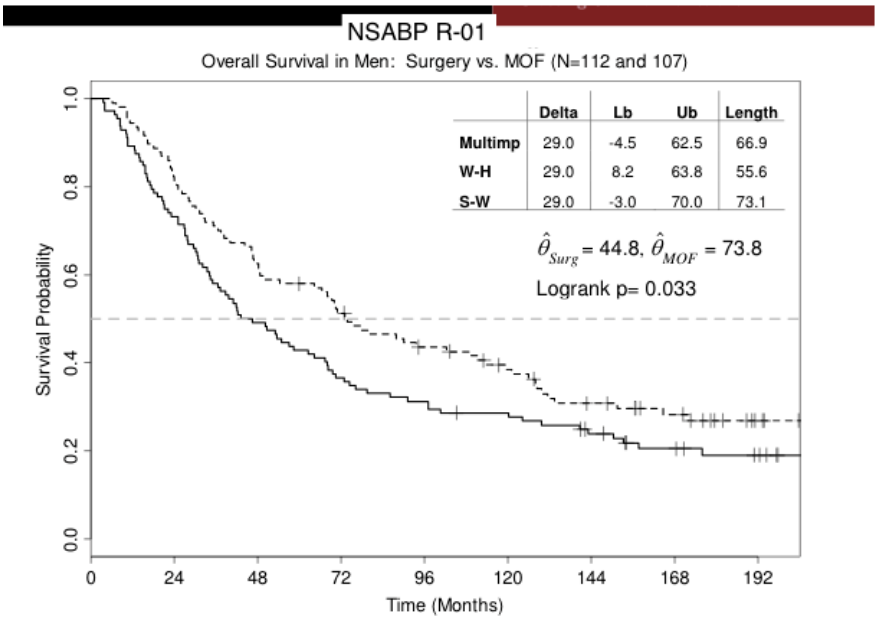


Figure 2 (Aim 2 – Quantiles for Survival Data): Median difference estimates for NSABP R-01 data. Shown are estimates for the difference in median survival time by treatment arm according to the newly derived methods and two previously proposed methods.

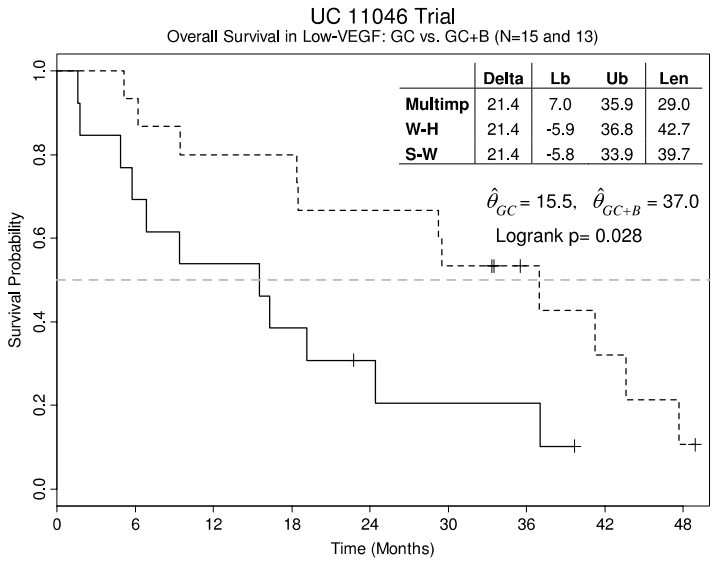


Figure 3 (Aim 2 – Quantiles for Survival Data): Median difference estimates for the UC mesothelioma data. Shown are estimates for the difference in median survival time by treatment arm according to the newly derived methods and two previously proposed methods.

<u>Conclusions</u>

This easy-to-compute method performs comparably with less restrictive assumptions than the Wang and Hettmansperger method and similar confidence interval width to the Su and Wei, and

appears to do particularly well in small sample situations. We anticipate completion of these and the manuscript in early 2014, and using this method where appropriate in our clinical research reporting.

*Aim 3 Summary*

In clinical oncology, competing risks are very common, and ubiquitous in the disease scenarios we most often deal with in RTOG (potentially curable or controllable single site disease, as opposed to widespread metastatic disease). While competing risks methods abound, correct use and interpretation of these is less common. This project aimed to explore empirically how different approaches, either of which may be correct for a specific question, nonetheless can produce seemingly different answers, and thus the critical issue is using and interpreting the right method. Through a comprehensive study of the application of competing risks regression models in the oncology setting, the project would provide justification for which model should be applied according to the research questions of interest. This aim was achieved, as described below and in the publication arising from this work (publication #2, page 26).

To expand on why competing risks are so common, 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. We used simulation studies examining different competing risks scenarios with respect to hazards for different events, correlation among event times, and amount of censoring to illustrate the key issues, and then followed with a data example from the RTOG trials.

Background: 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 $\lambda^*_k$ 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 (Klein and Anderson 2005).

The two above methods are often used interchangeably, when in fact they address different aspects of the failure process. Conceptually speaking, the CSH model naturally conditions on not failing from the competing cause, that is, the effect of some covariate (treatment, some patient feature, etc.) is among those having not failed from the competing cause. The SDH model, being a function of the cumulative incidence function (which measures the cumulative risk of a given event, which is influenced by the hazard for that event and competing events), measures the 'net' effect of that covariate on failing from the given cause, whether due to greater or lesser risk for that cause, or concurrent greater or lesser risk for other causes. This will become more concrete in the simulations and data example.

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 most commonly used regression models (Prentice et al 1978, Fine and Gray 1999). After specifying parameters for a bivariate survival distribution (Hougaard 1986), 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.

Highlights of the findings are summarized here (Table 3).

*Scenario II* – (Table 3) 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.

Table 3 (Aim 3 – Competing Risks Regression Models): Comparison of regression models under different competing risks scenarios. Competing risks data simulated from a bivariate exponential failure distribution with hazard parameters indicated and independent censoring of

about 33% of lifetimes. Averages of estimated parameters based on 3000 simulated datasets, with 250 subjects per group.

| | | | | Model Estimates | | | |
|---|---|---|---|---|---|---|---|
| | Specified Hazards | | | Cox | | Fine and Gray | |
| **Scenario** | **Group A** | **Group B** | | **CHR** | **95% CI** | **SDR** | **95% CI** |
| | | | | | | | |
| **II: Event 1 rate lower in Grp A** | | | | | | | |
| Event 1 | $\lambda_{11} = 0.50$ | $\lambda_{21} = 1.00$ | | 2.01 | 1.43-2.84 | 1.79 | 1.29-2.50 |
| Event 2 | $\lambda_{12} = 1.00$ | $\lambda_{22} = 1.00$ | | 0.99 | 0.75-1.33 | 0.75 | 0.57-0.99 |
| | | | | | | | |
| **III: Both event rates lower in Grp A** | | | | | | | |
| Event 1 | $\lambda_{11} = 0.50$ | $\lambda_{21} = 1.00$ | | 2.00 | 1.45- 2.76 | 1.29 | 0.96- 1.74 |
| Event 2 | $\lambda_{12} = 0.50$ | $\lambda_{22} = 1.00$ | | 2.00 | 1.45- 2.76 | 1.29 | 0.96- 1.74 |
| | | | | | | | |
| **IV: Large and small effects in Group A** | | | | | | | |
| Event 1 | $\lambda_{11} = 0.50$ | $\lambda_{21} = 1.00$ | | 2.58 | 1.76 -3.79 | 2.13 | 1.48 - 3.07 |
| Event 2 | $\lambda_{12} = 0.75$ | $\lambda_{22} = 1.00$ | | 1.13 | 0.83 -1.54 | 0.79 | 0.59 - 1.07 |

We also examined the effect of censoring on how the two approaches may differ (Table 4). Heuristically, if all patients fail of one cause or the other (no censoring), then the SDH estimate will show a ‘teeter-totter’ effect in that the estimates should be on opposite sides of the null value of 1.0. The CSH estimate should not show this effect except under strong correlation between event times. The extent to which such a phenomenon is observable will thus depend on how many patients are event-free. A simple simulation confirmed this result, and explains in part why in many practical instances (where many patients are censored) the CSH and SDH hazard ratios are quite similar.

Table 4 (Aim 3 – Competing Risks Regression Models): Effect of censoring on the degree to which CHR and SHR estimates differ when the covariate is only related to cause 1 failure.

| | **Event 2 Effect Estimates** | | | | **Event 1 Effect Estimates** | | | |
|---|---|---|---|---|---|---|---|---|
| | **Cox** | | **Fine-Gray** | | **Cox** | | **Fine-Gray** | |
| **II: No event 2 hazard difference** | **CHR** | **95% CI** | **SHR** | **95% CI** | **CHR** | **95% CI** | **SHR** | **95% CI** |
| **% cases censored** | | | | | | | | |
| 0% | 0.99 | 0.83–1.20 | 0.69 | 0.57– 0.83 | 2.00 | 1.60-2.50 | 1.76 | 1.41-2.19 |
| 10% | 1.00 | 0.82–1.22 | 0.71 | 0.58– 0.86 | 2.01 | 1.58–2.54 | 1.77 | 1.40-2.23 |
| 25% | 0.99 | 0.81–1.12 | 0.73 | 0.5 – 0.90 | 2.00 | 1.55-2.58 | 1.78 | 1.38-2.28 |
| 50% | 0.99 | 0.77–1.29 | 0.79 | 0.61– 1.01 | 2.01 | 1.49-2.72 | 1.81 | 1.35-2.43 |
| 67% | 0.99 | 0.72–1.37 | 0.84 | 0.62– 1.15 | 2.01 | 1.37-2.94 | 1.86 | 1.28-2.69 |
| 85% | 0.98 | 0.60–1.63 | 0.92 | 0.56– 1.51 | 2.02 | 1.11-3.66 | 1.95 | 1.08-3.49 |

In general, these findings are commensurate with a prior study examining nonparametric estimators and tests in the competing risks setting (Dignam and Kocherginsky 2008).

*Data Example: 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 androgen deprivation therapy (ADT) 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.

Hormone treatment 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 5), a similar pattern emerges. Androgen deprivation therapy (ADT) only nominally increases the *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 5 (Aim 3 – Competing Risks Regression Models): 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** | **SDH** | **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 | - | - | - | - |

Conclusions

We summarized these results for publication, and in our discussion provided guidance on the basis for choosing one modeling approach versus another.

We submitted the manuscript to *Clinical Cancer Research,* a high impact journal from the American Association for Cancer Research (AACR) that frequently publishes more technical articles on statistical methodology practice in oncology (publication #2, page 26). We were very pleased to obtain a positive review, and the article was accepted upon resubmission with revisions. A complementary editorial accompanied the article.

Posters/Presentations Related to All Aims

*Aim 1:*

Oral presentation (2011 Society for Clinical Trials) and abstract: Wang M Dignam J. The Impact of Association Between Endpoints on Performance in Seamless Phase II/III Clinical Trials. *Clinical Trials* 2011; 8:490

*Aim 2:*

None

*Aim 3:*

Oral Presentation: Competing Risk Analysis. *Department of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago*, Chicago, IL, March 4, 2011 (invited).

Oral Presentation: Competing Risks Analysis Methods in Cancer Clinical Research. *Division of Biostatistics, Department of Pharmacology and Experimental Therapeutics, Thomas Jefferson*, Philadelphia, PA, January 10, 2013 (invited).

Literature Cited (all aims)

Hunsberger S, Zhao Y, Simon R. A Comparison of Phase II Study Strategies. *Clinical Cancer Research*. 2009;15:5950-5955

Rubinstein L, Korn E, Freidlin B, *et al*. Design issues of randomized phase II trials and a proposal for phase II screening trials. *J Clin Oncol* 2005; 23: 7199–206.

Hougaard P. A class of multivariate failure time distributions. *Biometrika* 1986;73:671-678

Bryant J, Day R. Incorporating toxicity considerations into the design of two-stage phase II clinical trials. *Biometrics* 1995; 51: 1372–83.

Wang M, Day R. An adaptive Bayesian approach to jointly modeling response and toxicity in phase I dose finding trials. *J Biopharm Stat* 2010; 20(1):125–44.

Royston P, Parmar MKB, Qian W. Novel designs for multi-arm clinical trials with survival outcomes, with an application in ovarian cancer. *Stat Med* 2003; 22(14): 2239–56.

Royston P, Barthel FMS, Parmar MKB, Choodari-Oskooei B, Isham V. Designs for clinical trials with time-to-event outcomes based on stopping guidelines for lack of benefit. *Trials* 2011; 12(81).

Hettmansperger T.P. Two-sample inference based on one-sample sign statistics. *Applied Statistics* 1985, 33, 45-51.

Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*, 1958; 53, 457-481.

Price R.M. and Bonett, D.G. Estimating the variance of the sample median. *Journal of Statistical Computation and Simulation*, 2001; 68, 295-305.

Price, R.M. and Bonett, D.G. Distribution-free confidence intervals for difference and ratio of medians. *Journal of Statistical Computation and Simulation*, 2002; 72, 119-124.

Rubin, D.B. Multiple imputations in sample surveys - a phenomenological Bayesian approach to nonresponse. Imputation and Edition of Faulty or Missing Survey Data, 1978, U.S. Department of Commerce, pp. 1- 23.

Rubin D.B. and Schenker N. Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *J Am Stat Assoc* 1986; 81, 366-374.

Su, J.Q. and Wei, L.J. Nonparametric estimation for the difference or ratio of median failure times. *Biometrics*, 1993; 49, 603-607.

Wang JL, Hettmansperger TP. Two-sample inference for median survival times based on one-sample procedures for censored survival data. *J Am Stat Assoc*, 1993; 85,529-536.

Fisher B, Wolmark N, Rockette H, et al. Postoperative adjuvant chemotherapy or radiation therapy for rectal cancer: results from NSABP protocol R-01. *J Natl Cancer Inst*. 1988;80:21-9.

Kindler HL, Karrison TG, Gandara DR, et al. Multicenter, double-blind, placebo-controlled, randomized phase II trial of gemcitabine/cisplatin plus bevacizumab or placebo in patients with malignant mesothelioma. *J Clin Oncol* 2012 30(20), pp. 2509–15.

Gilbert M, Dignam JJ, Armstrong TS, et al. A randomized clinical trial of bevacizumab for newly diagnosed glioblastoma. *New England Journal of Medicine*, in press, Feb 2014

Prentice RL, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. *Biometrics* 1978;34:541-554

Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc* 1999;94:496-509

Dignam JJ, Kocherginsky MN. Choice and interpretation of statistical tests used when competing risks are present. *J Clin Oncol* 2008;26:4027-4034

Klein JP, Andersen PK. Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. *Biometrics* 2005;61:223-229

Pilepich MV, Winter K, John MJ, et al. Phase III radiation therapy oncology group (RTOG) trial 86-10 of androgen deprivation adjuvant to definitive radiotherapy in locally advanced carcinoma of the prostate. *Int J Radiat Oncol Biol Phys* 2001;50:1243-1252

**18. Extent of Clinical Activities Initiated and Completed**. Items 18(A) and 18(B) should be completed for all research projects. If the project was restricted to secondary analysis of clinical data or data analysis of clinical research, then responses to 18(A) and 18(B) should be "No."

18(A) Did you initiate a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
__X__No

18(B) Did you complete a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

_____Yes
__X__No

**If "Yes" to either 18(A) or 18(B), items 18(C) – (F) must also be completed.** (Do NOT complete 18(C-F) if 18(A) and 18(B) are both "No.")

18(C) How many hospital and health care professionals were involved in the research project?

_____Number of hospital and health care professionals involved in the research project

18(D) How many subjects were included in the study compared to targeted goals?

_____Number of subjects originally targeted to be included in the study
_____Number of subjects enrolled in the study

**Note**: Studies that fall dramatically short on recruitment are encouraged to provide the details of their recruitment efforts in Item 17, Progress in Achieving Research Goals, Objectives and Aims. For example, the number of eligible subjects approached, the number that refused to participate and the reasons for refusal. Without this information it is difficult to discern whether eligibility criteria were too restrictive or the study simply did not appeal to subjects.

18(E) How many subjects were enrolled in the study by gender, ethnicity and race?

Gender:
_____Males
_____Females

_____Unknown

Ethnicity:
_____Latinos or Hispanics
_____Not Latinos or Hispanics
_____Unknown

Race:
_____American Indian or Alaska Native
_____Asian
_____Blacks or African American
_____Native Hawaiian or Other Pacific Islander
_____White
_____Other, specify:_____
_____Unknown

18(F) Where was the research study conducted? (List the county where the research study was conducted. If the treatment, prevention and diagnostic tests were offered in more than one county, list all of the counties where the research study was conducted.)

**19. Human Embryonic Stem Cell Research.** Item 19(A) should be completed for all research projects. If the research project involved human embryonic stem cells, items 19(B) and 19(C) must also be completed.

19(A) Did this project involve, in any capacity, human embryonic stem cells?
_____Yes
__X__No

19(B) Were these stem cell lines NIH-approved lines that were derived outside of Pennsylvania? N/A
_____Yes
_____ No

19(C) Please describe how this project involved human embryonic stem cells:

**20. Articles Submitted to Peer-Reviewed Publications**.

20(A) Identify all publications that resulted from the research performed during the funding period and that have been submitted to peer-reviewed publications. Do not list journal abstracts or presentations at professional meetings; abstract and meeting presentations should be listed at the end of item 17. **Include only those publications that acknowledge the Pennsylvania Department of Health as a funding source** (as required in the grant agreement). List the title of the journal article, the authors, the name of the peer-reviewed

publication, the month and year when it was submitted, and the status of publication (submitted for publication, accepted for publication or published.). Submit an electronic copy of each publication or paper submitted for publication, listed in the table, in a PDF version 5.0.5 (or greater) format, 1,200 dpi. Filenames for each publication should include the number of the research project, the last name of the PI, and an abbreviated title of the publication. For example, if you submit two publications for Smith (PI for Project 01), one publication for Zhang (PI for Project 03), and one publication for Bates (PI for Project 04), the filenames would be:

Project 01 – Smith – Three cases of isolated
Project 01 – Smith – Investigation of NEB1 deletions
Project 03 – Zhang – Molecular profiling of aromatase
Project 04 – Bates – Neonatal intensive care

If the publication is not available electronically, provide 5 paper copies of the publication.

**Note:** The grant agreement requires that recipients acknowledge the Pennsylvania Department of Health funding in all publications. Please ensure that all publications listed acknowledge the Department of Health funding. If a publication does not acknowledge the funding from the Commonwealth, do not list the publication.

| Title of Journal Article: | Authors: | Name of Peer-reviewed Publication: | Month and Year Submitted: | Publication Status (check appropriate box below): |
|---|---|---|---|---|
| 1. Integrated phase II/III clinical trials in oncology: A case study | Wang M, Dignam JJ, Zhang QE, Degroot JF, Mehta MP, Hunsberger S. | Clinical Trials | 12/2010 | ☐Submitted<br>☐Accepted<br>☒ Published |
| 2. The use and interpretation of competing risks regression models. | Dignam JJ, Zhang QE, Kocherginsky M. | Clinical Cancer Research | 08/2011 | ☐Submitted<br>☐Accepted<br>☒Published |

20(B) Based on this project, are you planning to submit articles to peer-reviewed publications in the future?

Yes_____X_____ No_____

If yes, please describe your plans:

The manuscript for aim 3 is nearing completion and we plan to submit it for publication in the first half of 2014.

**21. Changes in Outcome, Impact and Effectiveness Attributable to the Research Project.**
Describe the outcome, impact, and effectiveness of the research project by summarizing its impact on the incidence of disease, death from disease, stage of disease at time of diagnosis,

or other relevant measures of outcome, impact or effectiveness of the research project. If there were no changes, insert "None"; do not use "Not applicable." Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

As this project involves methodological work in study design and analysis, impact or effectiveness is not measured in a single instance but rather the impact would be reflected in changes to how we approach research in the future. All of the aims concerned with better ways to design and analyze cancer clinical trials, as well as interpret findings in a more informative way. As such, we expect this work to have substantive impact going forward. Specifically, as mentioned, we plan to use the phase II/III design (aim 1) wherever advantageous. For the quantile work (aim 2), our growing portfolio of comparative phase II trials will require median comparisons, and this approach will serve us well. Competing risks (aim 3) are ubiquitous in cancer clinical research and of particular relevance in several of our disease sites of interest. Better application of the methods available and extension of this information to study design will lead to more efficient and informative clinical trials.

**22. Major Discoveries, New Drugs, and New Approaches for Prevention Diagnosis and Treatment**. Describe major discoveries, new drugs, and new approaches for prevention, diagnosis and treatment that are attributable to the completed research project. If there were no major discoveries, drugs or approaches, insert "None"; do not use "Not applicable." Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

None

**23. Inventions, Patents and Commercial Development Opportunities.**

23(A) Were any inventions, which may be patentable or otherwise protectable under Title 35 of the United States Code, conceived or first actually reduced to practice in the performance of work under this health research grant? Yes_____ No___X_____

If "Yes" to 23(A), complete items a – g below for each invention. (Do NOT complete items a - g if 23(A) is "No.")

a. Title of Invention:

b. Name of Inventor(s):

c. Technical Description of Invention (describe nature, purpose, operation and physical, chemical, biological or electrical characteristics of the invention):

d. Was a patent filed for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?

Yes_____ No___

If yes, indicate date patent was filed:

e. Was a patent issued for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
Yes_____ No___
If yes, indicate number of patent, title and date issued:
Patent number:
Title of patent:
Date issued:

f. Were any licenses granted for the patent obtained as a result of work performed under this health research grant? Yes_____ No___

If yes, how many licenses were granted?_____

g. Were any commercial development activities taken to develop the invention into a commercial product or service for manufacture or sale? Yes__ No___

If yes, describe the commercial development activities:

23(B) Based on the results of this project, are you planning to file for any licenses or patents, or undertake any commercial development opportunities in the future?

Yes_____ No___X_____

If yes, please describe your plans:

**24. Key Investigator Qualifications.** Briefly describe the education, research interests and experience and professional commitments of the Principal Investigator and all other key investigators. *In place of narrative you may insert the NIH biosketch form here; however, please limit each biosketch to 1-2 pages. For Nonformula grants only – include information for only those key investigators whose biosketches were not included in the original grant application.*

## BIOGRAPHICAL SKETCH

| NAME | POSITION TITLE |
|---|---|
| James J. Dignam, Ph.D. | Associate Professor, Biostatistics |
| eRA COMMONS USER NAME (credential, e.g., agency login) | |
| JDIGNAM | |

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)*

| INSTITUTION AND LOCATION | DEGREE *(if applicable)* | MM/YY | FIELD OF STUDY |
|---|---|---|---|
| Saint Xavier College, Chicago, IL | BA | 05/86 | Mathematics |
| University of Pittsburgh, Pittsburgh, PA | MS | 06/91 | Biostatistics |
| University of Pittsburgh, Pittsburgh, PA | PhD | 05/94 | Biostatistics |

### A. Personal Statement

I have been involved in cancer clinical trials design, conduct, monitoring, and analysis for over 20 years. I currently serve as the Group Statistician (Senior Director of statistical and data management activities) for the Radiation Therapy Oncology Group, a National Cancer Institute sponsored consortium that conducts clinical trials via a distributed network of investigators. I have also conducted collaborative cancer clinical and biologic research with University of Chicago investigators, and have been involved in as well as led a large number of studies examining factors that may influence cancer outcomes. These have included studies of race/ethnicity, breast density, and obesity. Major areas of statistical methodology research have included clinical trial design and interim monitoring, competing risks analysis, hazard function estimation, and prognostic factor modeling.

### B. Positions and Honors

**Positions and Employment**

1986-1987 Scientific Assistant, Argonne National Laboratory, Argonne, IL
1994-1998 Research Associate, Department of Biostatistics, University of Pittsburgh, PA
1998-2000 Research Assistant Professor, Department of Biostatistics, University of Pittsburgh, PA
2000 Statistician, American College of Surgeons Oncology Group, Chicago, IL
2001-2007 Assistant Professor, Department of Health Studies, University of Chicago, IL
2007-Present Associate Professor, Department of Health Studies, University of Chicago, IL
2004-Present Investigator, University of Chicago Cancer Research Center
2009-Present Group Statistician, Radiation Therapy Oncology Group

### C. Selected Peer-reviewed Publications

1. **Dignam JJ,** Bryant J, Wieand HS, Fisher B, Wolmark N. Early stopping of a clinical trial when there is evidence of no treatment benefit: Protocol B-14 of the National Surgical Adjuvant Breast and Bowel Project. *Controlled Clinical Trials* 19:575-588, 1998

2. Bouman P, Meng XL, **Dignam J,** Dukic V. A multiresolution hazard model for multicenter survival studies: Application to tamoxifen treatment in early stage breast cancer. *J Am Stat Assoc* 102:1145-57, 2007.
3. **Dignam JJ,** Wieand K, Johnson KA, Fisher B, Xu L, Mamounas EP. Obesity, tamoxifen use, and outcomes in women with estrogen receptor positive early stage breast cancer. *J Natl Cancer Inst* 95:1467-76, 2003.
4. Bryant J, **Dignam JJ.** Semiparametric models for cumulative incidence functions. *Biometrics* 60:182-190, 2004.
5. **Dignam JJ,** Polite B, Yothers G, et al. Body mass index and outcomes in patients who receive adjuvant chemotherapy for colon cancer. *J Natl Cancer Inst* 98:1647-54, 2006.
6. **Dignam JJ,** Wieand K, Rathouz PJ. A missing data approach to semi-competing risks problems. *Stat Med* 26:837-56, 2007.
7. Dukic V, **Dignam J.** Bayesian hierarchical multiresolution model for the study of time-dependent patterns of failure in early stage breast cancer. *Bayesian Analysis* 2:591-610, 2007
8. **Dignam JJ,** Kocherginsky MN. Choice and interpretation of statistical tests used when competing risks are present. *J Clin Oncol* 26:4027-34, 2008. PMCID: PMC265431
9. **Dignam JJ**, Huang L, Ries L, Reichman M, Mariotto A, Feuer E. Estimating breast cancer-specific and other cause mortality in clinical trial and population-based cancer registry cohorts. *Cancer* 115:5272-5283, 2009
10. **Dignam JJ,** Zhang Q, Kocherginsky MN. The use and interpretation of competing risks regression models. *Clinical Cancer Research*, 18:2301-2308, 2012.
11. Wang M, **Dignam JJ,** Zhang QE, Degroot JF, Mehta MP, Hunsberger S. Integrated phase II/III clinical trials in oncology: A case study. *Clin Trials*. 9:741-747, 2012
12. Lee M, Cronin K, Gail MH, **Dignam JJ**, Feuer EJ. Multiple imputation methods for inference on cumulative incidence with missing cause of failure. *Biometrical Journal*, 53:974-993, 2011
13. Chung CH, **Dignam JJ,** Hammond ME, et al. Glioma-associated oncogene family zinc finger 1 expression and metastasis in patients with head and neck squamous cell carcinoma treated with radiation therapy (RTOG 9003). *J Clin Oncol*. 29:1326-34, 2011.
14. Lawrence RL, Vikram B, **Dignam JJ,** Chakravarti A, et al. NCI-RTOG translational-program strategic guidelines for the early stage development of radiosensitizers. *Journal of the National Cancer Institute*, 2013 Jan 3 [Epub ahead of print]. PMCID: PMC3536642
15. Huo N, Hong S, Wang W, Olopade OI, **Dignam JJ,** Huo D. Hormone replacement therapy and breast cancer: heterogeneous risks by race, weight, and breast density. *J Natl Cancer Inst* 105:1365-1372, 2013

# BIOGRAPHICAL SKETCH

| NAME | POSITION TITLE |
|---|---|
| Masha Kocherginsky | Research Associate (Associate Professor) |
| eRA COMMONS USER NAME | |
| MASHAK | |

EDUCATION/TRAINING

| INSTITUTION AND LOCATION | DEGREE *(if applicable)* | YEAR(s) | FIELD OF STUDY |
|---|---|---|---|
| University of Illinois, Urbana-Champaign | BS | 1996 | Statistics & |
| University of Illinois, Urbana-Champaign | MS | 1998 | Statistics |
| University of Illinois, Urbana-Champaign | PhD | 2003 | Statistics |

## A. Personal Statement

I am a biostatistician with extensive experience in experimental design and application of a wide variety of statistical methods for the analysis of biomedical data, including clinical trial design, power analysis and sample size estimation, survival analysis, analysis of competing risks data, categorical data analysis, mixed and longitudinal regression models, and analysis of microarray data. As a member of the Biostatistics Laboratory, which provides University of Chicago researchers with biostatistical and epidemiological expertise in a collaborative setting, I have been involved in the design and analysis of numerous basic science and pre-clinical studies, and of many clinical trials. My statistical methods research interests include the use of multiple imputation methods for censored data in survival analysis, imputation in survey data, analysis of data in the presence of competing risks, and the optimal design of clinical trials for the multinomially distributed outcome.

## B. Positions and Honors

**Positions:**

| | |
|---|---|
| 1998-2001 | Teaching Assistant, Dept. of Statistics, University of Illinois |
| 1998-2003 | Statistical Consultant, Illinois Statistics Office, University of Illinois |
| 2001-2003 | Research Assistant, University of Illinois |
| 2003-2004 | Research Associate (Instructor), Dept. of Health Studies, University of Chicago |
| 2004-2011 | Research Associate (Assistant Professor), Dept. of Health Studies, University of Chicago |
| 2011-present | Research Associate (Associate Professor), Dept. of Health Studies, University of Chicago |

**Honors:**

| | |
|---|---|
| April 1999-2000 | Outstanding Rating, University of Illinois Incomplete List of Teachers Ranked as Excellent by their Students |
| June 2001 | First Prize, Capital One Financial Data Challenge |

## C. Selected peer-reviewed publications (in chronological order):

**Statistical Methods**

1. **Kocherginsky M**, He XM, Mu YM. Practical confidence intervals for regression quantiles. Journal of Computational and Graphical Statistics. 2005; 14(1): 41-55.
2. **Kocherginsky M** and He XM. Extensions of the Markov chain marginal bootstrap. Statistics & Probability Letters. 2007; 77(12): 1258-1268.

3. Dignam JJ and **Kocherginsky MN**. Choice and interpretation of statistical tests used when competing risks are present. Journal of Clinical Oncology. 2008; 26(24): 4027-34.
4. **Kocherginsky M**, Cohen EE, Karrison T. Design of Phase II cancer trials for evaluation of cytostatic/cytotoxic agents. Journal of Biopharmaceutical Statistics. 2009; 19(3): 524-9.
5. Dignam JJ, Zhang Q, **Kocherginsky M**. The use and interpretation of competing risks regression models. Clin Cancer Res. 2012;18(8):2301-8. PMCID: 3328633.
6. Pang D, **Kocherginsky M**, Krausz T, Kim SY, Conzen SD. Dexamethasone decreases xenograft response to Paclitaxel through inhibition of tumor cell apoptosis. Cancer Biology & Therapy. 2006; 5(8): 933-40.
7. Cohen EE, Haraf DJ, List MA, **Kocherginsky M**, Mittal BB, Rosen F, Brockstein B, Williams R, Witt ME, Stenson KM, Kies MS, Vokes EE. High survival and organ function rates after primary chemoradiotherapy for intermediate-stage squamous cell carcinoma of the head and neck treated in a multicenter phase II trial. Journal of Clinical Oncology. 2006; 24(21): 3438-44.
8. Artz AS, Pollyea DA, **Kocherginsky M**, Stock W, Rich E, Odenike O, Zimmerman T, Smith S, Godley L, Thirman M, Daugherty C, Extermann M, Larson R, van Besien K. Performance status and comorbidity predict transplant-related mortality after allogeneic hematopoietic cell transplantation. Biology of Blood & Marrow Transplant. 2006; 12(9): 954-64.
9. Artz AS, Wickrema A, Dinner S, Godley LA, **Kocherginsky M**, Odenike O, Rich ES, Stock W, Ulaszek J, Larson RA, van Besien K. Pretreatment C-reactive protein is a predictor for outcomes after reduced-intensity allogeneic hematopoietic cell transplantation. Biology of Blood & Marrow Transplant. 2008; 14(11): 1209-16.
10. Belova L, Delgado B, **Kocherginsky M**, Melhem A, Olopade OI, Conzen SD. Glucocorticoid receptor expression in breast cancer associates with older patient age. Breast Cancer Research & Treatment. 2009; 116(3): 441-7.
11. Williams JB, Pang D, Delgado B, **Kocherginsky M**, Tretiakova M, Krausz T, Pan D, He J, McClintock MK, Conzen SD. A model of gene-environment interaction reveals altered mammary gland gene expression and increased tumor growth following social isolation. Cancer Prevention Research. 2009; 2(10): 850-61.
12. Pan D, **Kocherginsky M**, Conzen SD. Activation of the glucocorticoid receptor is associated with poor prognosis in estrogen receptor-negative breast cancer. Cancer Res. 2011;71(20):6360-70.
13. Cohen EE, Wu K, Hartford C, **Kocherginsky M**, Eaton KN, Zha Y, Nallari A, Maitland ML, Fox-Kay K, Moshier K, House L, Ramirez J, Undevia SD, Fleming GF, Gajewski TF, Ratain MJ. Phase I studies of sirolimus alone or in combination with pharmacokinetic modulators in advanced cancer patients. Clin Cancer Res. 2012;18(17):4785-93.
14. Muffly, L. S., Boulukos, M., Swanson, K., **Kocherginsky, M.**, Cerro, P. Del, Schroeder, L., Pape, L., et al. (2012). Pilot Study of Comprehensive Geriatric Assessment (CGA) in Allogeneic Transplant: CGA Captures a High Prevalence of Vulnerabilities in Older Transplant Recipients. Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation.
15. Volden PA, Wonder EL, Skor MN, Carmean CM, Patel FN, Ye H, **Kocherginsky M**, McClintock MK, Brady MJ, Conzen SD. Chronic social isolation is associated with metabolic gene expression changes specific to mammary adipose tissue. Cancer Prevention Research. 2013 Jul;6(7):634-45.