Adaptive Seamless Design for An Efficacy Trial of Replication-Competent Adenovirus-mediated Suicide Gene Therapy and Radiation in Newly-diagnosed Prostate Cancer (ReCAP Trial)

Mei Lu[1], **Svend O. Freytag**[2], **Hans Stricker**[3], **Jae Ho Kim**[2], **Kenneth Barton**[2], and **Benjamin Movsas**[2]

[1] Department of Biostatistics and Research Epidemiology, Henry Ford Health System, Detroit, MI 48202

[2] Department of Radiation Oncology, Henry Ford Health System, Detroit, MI 48202

[3] Vattikuti Urology Institute, Henry Ford Health System, Detroit, MI 48202

Abstract

Purpose—Cumulative evidence has suggested investigation of the efficacy of **Re**plication-**C**ompetent **A**denovirus-mediated Suicide Gene Therapy in newly-diagnosed **P**rostate Cancer (ReCAP). There is a challenge in designing an efficacy trial for newly-diagnosed prostate cancer given its long natural history. The regulatory agency recommended a Phase II trial for safety before conducting the efficacy trial.

Experimental Design—The ReCAP trial is an adaptive seamless, multi-site open-label, randomized Phase II/III trial. Two hundred eighty men will be randomized to receive either replication-competent adenovirus-mediated suicide gene therapy followed by radiation (Arm 1) or radiation alone (Arm 2). Phase II trial component will include the first 21 patients in Arm 1 with complete toxicity through day 90 for safety evaluation. The primary efficacy endpoint is the time free from biochemical and/or clinical failure (FFF). The secondary efficacy endpoints are 2-year prostate biopsies and overall survival. Unequal spaced interim looks are proposed with the adaptive sample-size re-estimation.

Results—This trial has been approved by the FDA for the study therapy investigation and is currently recruiting patients.

Conclusions—Challenges remain in designing newly-diagnosed prostate cancer trials. Adaptive seamless design is time-saving and a cost-effective design in the development of novel medical therapies, but requires a specified statistical plan in the decision process involved.

Keywords

Adaptive Design; Prostate Cancer; Interim Analysis

Correspondence to: Mei Lu, Ph.D., Department of Biostatistics and Research Epidemiology, Henry Ford Health System, One Ford Place, 3E, Detroit, MI 48202, Telephone: (313) 874-6413, Fax: (313) 874-6730, mlu1@hfhs.org.

Introduction

Radiation therapy (RT) is an accepted treatment for localized prostate cancer; however the outcome varies with dose of radiation and the patient's risk factors. Increased radiation dose has produced encouraging results[1;2], but about 30% of patients with intermediate or high risk prostate cancer fail to respond to treatment. Androgen suppression therapy (AST) combined with RT has demonstrated a survival benefit in patients with high risk prostate cancer compared to RT alone [3]. A study has demonstrated a benefit when AST was combined with RT (~70 Gy) in patients with intermediate or high risk disease [4]. However, the role of AST in patients with intermediate risk disease in the setting of high dose radiation (~80 Gy) remains controversial and is under investigation (RTOG 0815), since AST is associated with significant morbidity and has adverse effects on the quality of life [5;6].

Replication-competent adenovirus-mediated suicide gene therapy, a single intraprostatic injection of the Ad5-yCD/*mut*TK$_{SR39}$*rep*-ADP adenovirus, followed by 5-fluorocytosine (5-FC) and valganciclovir (vGCV) prodrug, has been shown to radiosensitize tumor cells [7]. Phase I gene therapy trials of prostate cancer have shown no concerns for safety [8] and an increase in negative prostate biopsies at year 2, compared to historical controls [9]. Thus, an efficacy trial entitled: **Re**plication-**C**ompetent **A**denovirus-mediated Suicide Gene Therapy and Intensity-Modulated Radiation Therapy (IMRT) in newly-diagnosed **P**rostate Cancer (ReCAP), has been designed to test whether gene therapy in combination with IMRT could improve outcome, compared to IMRT alone. Initially, a 2-year negative prostate biopsy was proposed as the primary outcome in a randomized Phase II trial. At a pre-trial meeting with the FDA, a Phase III trial with time free from biochemical and/or clinical failure (FFF) as the primary endpoint was discussed. After presenting to the NIH Recombinant DNA Advisory Committee, the committee endorsed the design and the efficacy FFF endpoint, but suggested conducting a Phase II trial for safety before considering the Phase III trial, given the limited safety data available [8]. It normally takes at least six-to-nine months between ending a Phase II and starting a Phase III trial, not to mention the time to complete the Phase II trial. Therefore, we considered an adaptive seamless randomized Phase II/III design to better utilize our resources.

The adaptive seamless Phase II/III design combines the two traditional Phase II and Phase III trials and includes patients enrolled before and after the adaptation in the final analysis for treatment efficacy. This design has advantages of reducing required sample size and study duration and therefore time and cost. The trial may be stopped at the completion of the Phase II if there are toxicity concerns, no drug activity (i.e., Futility analysis), or a significant finding related to drug activity.

The adaptive seamless design has become popular due to new demands for faster new product development with correlative lower cost. In 2005, Pharmaceutical Research and Manufacturers of America (PRMA) assembled a working group to evaluate adaptive designs and their usage, and to gain regulatory acceptance for their implementation [10]. The FDA sponsored a series of workshops on this topic [11;12], and began to develop a guidance that is expected to be released in the early 2010. Objectives of this paper are to describe the rationales for designing this ReCAP trial and to discuss the methodological challenges relevant to this disease population.

Material and Method

Trial design and patient population

This is an adaptive seamless, multi-site, open-label, randomized, controlled Phase II/III trial of men with newly diagnosed prostate cancer. The trial contains two treatment arms, Arm 1- Replication-competent adenovirus-mediated suicide gene therapy (Gene Therapy) +IMRT as the investigational therapy or treatment Arm and Arm 2 –IMRT as the control/standard of care Arm (see Figure 1). Funded by the National Cancer Institute (NCI), this is the first randomized, controlled gene therapy trial for non-metastatic prostate cancer known to us.

The study includes 280 patients, 140 in each arm. Given the limited availability of toxicity data with the investigational agent [8], as the phase II component, acute toxicity will be evaluated 90 days after randomization in the first 21 patients enrolled to the investigational therapy arm (total of 42 in both arms).

Patients with biopsy confirmed and clinically localized (Stage T1/T2) prostate cancer are eligible for the trial. Other eligibility criteria include 1) age (18 or older); 2) Karnofsky performance status ≥ 70; 3) negative lymph nodes confirmed by imaging or biopsy; 4) no evidence of metastatic disease; 5) no prior treatment of prostatectomy, cryosurgery, radiotherapy, systemic chemotherapy or AST and no planned AST.

Exclusion criteria includes 1) patients with low-risk disease (Gleason ≤ 6 and PSA < 10 ng/ mL and < 50% positive biopsy cores); 2) patients with high-risk disease (Stage ≥ T3 or Gleason ≥ 8 or PSA > 20 ng/mL); 3) high prostate volume (> 120 cc); 4) positive lymph node or metastatic disease; 5) prior invasive cancer, except for non-melanoma skin cancer, within 5 years of enrollment; 6) prognosis for survival of < 5 years; 7) prior radical prostatectomy, cryosurgery for prostate cancer, or bilateral orchiectomy for any reason; 8) prior radiotherapy, including brachytherapy, to the region of the study cancer that could result in overlap of radiation fields; 9) prior or planned AST, or prior systemic chemotherapy for the cancer; 10) severe, active co-morbidity; 11) positive serological test for HIV, 12) previous history of significant liver disease including hepatitis; 13) immunosuppressive therapy including systemic corticosteroids; 14) impaired immunity or susceptibility to serious viral infections; 15) allergy to any product used in the protocol; and 16) serious medical or psychiatric illness or concomitant medication, which, in the judgment of the principal investigator, might interfere with the subject's ability to respond to or tolerate the treatment or complete the trial. Patients could not have any active HIV infection per FDA. Corticosteroids suppress the immune response, which is necessary to fight off the adenovirus infection.

Randomization

The randomization schema was prepared prior to the study initiation using varying block sizes 2 to 6 and stratified by the two stratification variables: 1) clinical sites (Henry Ford Hospital served as the primary site, and one or two additional sites will be involved with the patient recruitment) and 2) prognostic risk in three categories, defined as, A) (Gleason score 5/6 and PSA 10 – 20 ng/mL) OR (Gleason score 7 and PSA < 20 ng/mL) AND < 50% positive biopsy cores; B) (Gleason score 5/6 and PSA 10 – 20 ng/mL) OR (Gleason score 7 and PSA < 20 ng/mL) AND ≥ 50% positive biopsy cores; and C) Gleason score 5/6 AND PSA < 10 ng/mL AND ≥ 50% positive biopsy cores. Both Gleason and PSA combinations and percentage of positive biopsy cores are known risk factors for prostate cancer prognosis [3;13]. For simplicity, the other risk factors (e.g., age and ethnicity) are not included.

Prior to the randomization, the clinical site principle investigator (PI) must confirm that each subject meets all inclusion/exclusion criteria of the protocol and that a signed copy of the

informed consent document has been obtained. The code with either the study treatment assignment Arm 1 or Arm 2 will be issued to the subject according to the next available sequence. Subjects are considered to be randomized in the trial upon initiation of the study treatment.

Study Treatment

The treatment schema is illustrated in Figure 1. Patients in Arm 1 receive a single intraprostatic injection of the Ad5-yCD/*mut*TK$_{SR39}$*rep*-ADP adenovirus at 1×10^{12} vp on Day 1. Two days later (Day 3), patients will receive a two week course (10 days, weekdays only) of 5-fluorocytosine (5-FC) and valganciclovir (vGCV) prodrug therapy concomitant with 80 Gy IMRT (40×2 Gy fractions) for 8 weeks. Patients in Arm 2 receive 80 Gy IMRT only.

Patient Assessments

Patient assessments prior to, during, and after treatment are summarized in Table 1. Patients will undergo a thorough evaluation prior to enrollment to ensure they meet all eligibility requirements. Blood draws are required and clinical assessments are performed during the study treatment period for patient safety. The biopsy will be performed at 2 years after the completion of the IMRT, in addition to the baseline assessment. Patients will be closely monitored for any adverse reaction during and after IMRT.

Blinding/unblinding

Blinding to study treatment is intended to limit bias in a trial conduct and its interpretation. Due to the complicated procedures involved with the ultrasound-guided intraprostatic injection of adenovirus, it is difficult to conceal the treatment information (e.g., no ultrasound-guided intraprostatic sham injections for controls); therefore, the open-label approach is used. Nevertheless, steps are taken in the trial to minimize such bias, including 1) to blind the physician/patient with the study treatment until a randomization code is issued when there is an absolute need for patient scheduling, 2) to have the same follow-up evaluations on all randomized patients, 3) to blind pathologists who evaluate pathology results and the laboratory personal who perform the PSA laboratory tests with the treatment allocation, and 4) to blind the treating physicians to interim efficacy analyses results, if any. However, the data collection regarding patient safety and the safety results will not be blinded to the study investigators and staff.

Hypothesis, endpoint and analysis—Toxicity up to 90 days is the primary endpoint for the Phase II component that counts any gastrointestinal and genitourinary events including re-hospitalization and death. All adverse events are defined by Common Terminology Criteria for Adverse Events v3.0 (CTCAE) of the NCI. All possible toxicity grades and serious adverse events (SAEs) will be presented by the study treatment arms and analysis will be descriptive with distribution of AE grades, the percentage of SAEs and its 95% confidence limits. The hypothesis, endpoint and analysis for Phase III component are described in the following.

The primary hypothesis (P1): the gene therapy in combination with IMRT has an effect on time to freedom from biochemical/clinical failure (FFF) relative to the control therapy of IMRT.

Biochemical failure is measured as the first date of a rise of ≥ 2 ng/mL above the nadir PSA [14] after the study treatment initiation, where a nadir PSA is the lowest PSA value after the randomization and prior to the call date PSA. Clinical failures include any incidence of 1) local progression, measured by a $\geq 25\%$ increase in the product of the two largest

perpendicular diameters of the prostate tumor at any time, 2) tumor recurrence, defined as a redevelopment of a palpable abnormality after complete disappearance of previous abnormalities, 3) metastatic disease, confirmed by CT, MRI or biopsy, and/or 4) death from any cause.

Kaplan-Meier and log rank test will be used to examine FFF (time to freedom from biochemical/clinical failure) differences between the two arms, adjusting for the stratification variables with estimation of the median FFF assuming that stratification prognosis risk contributions to the treatment effect are quantitative rather than qualitative . The significance of the treatment effect will be adjusted for multiple interim looks (detailed in sample size calculation section). The primary analysis will be conducted using the intention-to-treat (ITT) algorithm. Intention-to-treat (ITT) is required for efficacy trials. It compares the treatment groups as assigned, based on the randomization, irrespective of the treatment they actually receive, based on all patients randomized. If a patient fails to reach the endpoint assessment, the endpoint will be computed using a proper analysis approach (e.g., last observation carries over or the worst outcome), followed by a sensitivity analysis.

<u>Secondary Hypothesis (S1 for safety)</u> – there is no safety concerns in the gene therapy in combination with IMRT. Toxicity is also a secondary endpoint for the Phase III component and will be analyzed using the same analysis proposed for the safety endpoint in Phase II component.

<u>Secondary Hypothesis (S2 for efficacy)</u> – the gene therapy in combination with IMRT has an effect on local tumor control relative to IMRT.

Local tumor control is defined as having a negative prostate biopsy at year 2 after completion of IMRT and no evidence of biochemical or clinical failure (FFF within 2 years after IMRT). Biopsy slides of all subjects will be examined by two GU pathologists. Both pathologists must agree on the diagnosis in order for the 2-year biopsy to be scored as either positive or negative. Two-year biopsy specimens that are equivocal will be reviewed by a third GU pathologist, with no affiliation with either clinical site, for the final determination of absence/presence of adenocarcinoma. Mantel-Haenszel test will be used to compare the proportional negative biopsy differences between two Arms adjusting for stratification variables

<u>Secondary Hypotheses (S3–S5 for efficacy)</u> – the gene therapy in combination with IMRT has an effect on time free from 1) biochemical (PSA) /clinical failure including <u>disease</u> related death, 2) distant metastases including disease related death, and 3) on overall survival, relative to IMRT. A similar analysis approach proposed for the primary hypothesis (P1) will be used to test the treatment effect on each outcome of interest.

<u>Secondary Hypothesis (S6 for efficacy)</u> – the gene therapy in combination with IMRT has an effect on quality of life (QOL), relative to the control therapy of IMRT only. QOL will be assessed based on two self-report survey instruments, EPIC [15] and EQ-5D [16] at baseline, 0.5, 1, 2 and 5 years after completion of IMRT.

An EPIC questionnaire will be completed by the patient. The questionnaire will be summarized into 4 endpoints of function and difficulty in urinary, bowel, hormone and sexual domains with a score of 0 to 100. EQ-5D will have a self-report score of 1 to 3 in 5 domains (mobility, self-care, usual activity, pain/discomfort and anxiety/depression), and an overall health score 10 to 100. All the patients will be included in the analysis. The worst score will be used for a patient who dies before the follow-up assessments.

Data will be evaluated for normality, and data transformation or the nonparametric approach will be considered if data is not normal. For each QOL domain variable, a mixed model will be used. Analysis will include the baseline covariates of the pre-treatment domain score, the stratification variables, the time (of post-treatment assessment) and the independent treatment variable. The analysis will start testing for treatment by time interaction, followed by testing the treatment effect at each time point, if the interaction is detected at critical value 0.10, or testing the overall treatment effect or time effect at critical value of 0.05, if otherwise.

Sample Size Calculation

The adaptive design yields the first interim look when the first 21 patients in the investigational therapy arm (42 patients in both arms) have completed the toxicity evaluation through day 90 after the randomization (Phase II component). Two additional interim looks are proposed after the Phase II component, when 1/3 and 2/3 of the total events are observed. O'Brien-Fleming spending function [17;18] is used to control type I errors.

No investigational treatment-related SAEs were observed in our completed Phase I trials on a total of 52 patients. With 21 patients in this Phase II component, assuming no investigational treatment related SAEs, the upper 95% confidence limit of SAE will be no more than 0.14 [19].

The time free from biochemical/clinical failure (FFF) in controls was estimated based on results of the Phase III trial [1;20;21], in which patients with either intermediate or high risk prostate cancer were randomized to receive either 70 or 78 Gy radiotherapy and 59% freedom of biochemical/clinical failure excluding deaths at 6 years was observed in 78 Gy. Considering 80 Gy and death rate about 10% and intermediate risk group for the proposed study, we expect a 65% FFF at 6 years in the control group.

The effect size for time-to-event endpoint (e.g. FFF) can be expressed as

$$\delta = Ln\left(\frac{\lambda_{treatment}}{\lambda_{control}}\right)$$

(1)

where $\gamma_{treatment}$ and $\gamma_{control}$ are the proportion of responses in the treatment treated and control treated groups. The ratio of the proportions in Equation (1) originally refers to the relative risk (RR) for a binary-endpoint assuming binominal distribution, but is extended to the hazard ratio (HR) for a time-to-event-endpoint assuming a constant hazard ratio [22]. Since the time-to-event-endpoint includes the time when the event occurs, the sample size calculation for such an endpoint depends not only on those above mentioned factors, but also on accrual rate or study duration, defined as the time of the first patient enrolled to the last patient to have a complete follow-up. A 65% time free from biochemical/clinical failure (FFF) rate at 6 years yields a constant hazard rate $\gamma_{control}$ = 0.0718 for controls, assuming exponential distribution. Given the feature of the time to event outcome, a patient will be assessed either to have an event with the event time or not have an event with the last follow-up time. The time and event or last follow-up will be included into the analysis. Therefore, patient drop out is not expected.

Considering alpha = 0.05 (Type I error) while using the O'Brien-Fleming spending function [17;18] to control multiple looks, a two-sided test and an exponential distribution of FFF (time free from biochemical/clinical failure), with 4 unequal spaced looks, it is proposed that

at power of 80%, a 15% absolute increase in FFF in the investigational therapy group (Gene Therapy + IMRT) will be clinically meaningful relative to 65% FFF in controls (IMRT alone), which yielded an effect size δ as 0.6578 in Equation (1). Assuming an accrual rate of 90 patients per year from multiple clinical sites with a total of 280 patients, we are able to detect such effect size with a maximum of 79 events at expected study duration 6.8 years (Figure 2). Cross boundaries by Z-value or P-value based on time free from biochemical/ clinical failure (FFF) are presented in Figure 3. If the test statistics and p-value cross into the blue zone, the trial may be stopped in favor of the investigational therapy. It is estimated that there are only 14 events at the completion of patient recruitment (Year 3.1), 27 events (Year 4.5) at the second interim analysis and 54 events (Year 5.6) at the third interim analysis. It should be noted that the sample size/power calculation is based on the total number of events (of failure), rather than the number of patients, a special case for time-to-event endpoint compared to the usual binary endpoint. The simulation study using EAST 5.1 shows no boundary at the first interim analysis for treatment efficacy, ± 3.7 at the second interim analysis with a critical value of <0.0002 on a maximum 27 events, ± 2.51 at the third interim analysis with a critical value of 0.0123 on a maximum of 54 events, and ± 1.99 with a critical value of 0.0462 on the final analysis with a maximum of 79 events (Figure 3). The power estimation under 20,000 simulations is 80.23%. The detailed power justifications are provided for the secondary outcomes, which are included in the clinical protocol for FDA submission.

Interim Analysis

An interim analysis will be performed at each interim look as proposed and then presented to the external Data Safety Monitor Board (DSMB) committee. The interim analysis will not withhold subject recruitment and follow-up. The interim analysis reports will include, but not be limited to, the following items 1) evaluation of the study inclusion and exclusion criteria; 2) recruitment status; 3) data collection status and quality; 4) protocol deviation and validation reports for all patients; 5) patient safety report and 6) the study efficacy results, if any.

Given that 0% treatment-related serious adverse events (SAEs) and 5.4% ≥ Grade 3 toxicity were observed in the previous Phase I trials with cumulative 52 patients, we are expecting less than 5% treatment-related SAEs and less than 15% ≥ Grade 3 events at 90 days. Depending on the patients' response to treatment, the external DSMB may have to make an early determination to stop/continue the trial if 1) there are safety concerns at any point in the trial; or 2) there is evidence of treatment efficacy with respect to time free from biochemical/clinical failure (FFF) when the test statistic or p-value at the interim analysis crosses into the blue zone (Figure 3).

Sample Size Adjustment

As we can see from the time-to-event endpoint, the number of events is more important to sample size calculation than the number of patients. The trial will be monitored closely at the time of the interim analyses when 1/3 and 2/3 of the maximum number of events have been achieved, with reassessment of the conditional power, control event rate, and the effect size. If the expected time free from biochemical/clinical failure (FFF) rate (c.a., 65% at Year 6) in controls is considerably different from the observed rate, or the recruitment rate is different from what is anticipated at the interim analysis, especially at the third interim look, the number of events or the study duration will be re-assessed.

Obviously, the maximum 79 events only yield 28% of the entire study population. There is a sufficient sample to observe more events if the patients are observed for a longer period of

time. Therefore, the study is powered to observe more than 79 events with an extended study period.

Discussion and Conclusion

The adaptive design was first developed in the 1920s, and applied to clinical research in the 1970s~80s on new demands by pharmaceutical companies to look at trials in an early stage for futility and trial modification. In response, FDA regulations continue to evolve parallel with the pharmaceutical industry's demands. The adaptive interim analysis approach was included in the FDA Guidance in 1988 and the adaptive approach for sample size adjustment was included ten years after, in 1998 [12].

The adaptive seamless design has since been introduced for new demands to develop new products faster, with more certainty and lower cost, which is regarded as a paradigm shift in clinical development. This is due to its unique time saving features, such as no pause in patient recruitment when proceeding from phase to phase, that the same clinical sites may be used in both phases and all required approvals for both phases can be obtained prior to initiation of the trial. Patients enrolled before and after adaptation will be used in the final efficacy analysis, thereby reducing the required sample size and cost. However, there are some restrictions in designing adaptive seamless trials. Basically, a similar protocol must be used for both phases with only limited adjustments made at the end of first phase, based on the results obtained. If the goal of a phase II trial is to determine what primary endpoint might be carried forward into phase III, then the adaptive seamless design may not be feasible. In other words, the adaptive seamless design requires a well-understood efficacy endpoint. Draft FDA guidance of adaptive design clinical trials was completed in February 2010 for input, but contains nonbinding recommendations [23]

There are challenges in designing clinical trials for newly-diagnosed prostate cancer. The first major challenge is the primary efficacy endpoint selection. In the oncology field, overall survival (OS) is generally considered the "gold-standard" efficacy endpoint for a Phase III trial. Given the average of 8 years from the first rise in PSA to the development of bony metastases and 5 years from development of bony metastases to death [24], the length of 10 to 20 years survival makes OS an unrealistic endpoint for newly-diagnosed patients. Thus, very few Phase III prostate clinical trials have been conducted by the pharmaceutical industry. This drawback has significantly hindered the development of novel prostate cancer therapies. In 2004, National Prostate Cancer Coalition (NPCC) and the FDA sponsored a forum to seek possible surrogate endpoints for prostate cancer. A set of PSA related measurements, PSA doubling time (PSADT), a biochemical failure based on rising PSA, and a composite endpoint of biochemical or clinical failure were discussed [25–27]. Some of these have been implemented in Phase III efficacy trials for this population [1;2;28]. Unfortunately, none of these endpoints have reached the high threshold needed for FDA acceptance, although, PSADT recently met three of Prentice's four criteria to be a valid surrogate endpoint compared to overall survival [29].

For the ReCAP trial, we chose the composite endpoint FFF as time to biochemical and/or clinical failure including death. Although the FFF can not be used to support product licensure at the present time, it has been used in multiple randomized, controlled trials that have changed the standard of care for prostate cancer radiotherapy [20;30]. The negative year 2 prostate biopsy was initially considered as the primary endpoint because it occurs much earlier than FFF and OS and is prognostic for clinical failure [31;32]. However, a major drawback of the prostate biopsy endpoint is its high false negative rate [21;33]. Increasing biopsy cores to 12, proposed in our ReCAP trial, can reduce error rate; however, the pathological determination of adenocarcinoma remains challenging [33]. Therefore, it

was decided to make the 2-year prostate biopsy a secondary endpoint. It is our hypothesis that the ReCAP trial will generate evidence that the gene therapy can improve the response to prostate radiotherapy, and provide important knowledge regarding the possible surrogate endpoints of FFF and OS.

A second challenge in designing clinical trials for newly-diagnosed prostate cancer relates to the futility analysis. The purpose of the futility analysis is to examine at an early point in the trial (i.e., after Phase II) evidence of no efficacy to determine whether to continue or stop the trial early. This is very important for preventing future patients from undergoing unnecessary procedures and therefore unnecessary risks when there is little chance of the treatment benefit, and was initially suggested by the FDA. A futility analysis was not included in this trial because: 1) there are no validated, early surrogate endpoints for the primary endpoint of the time free from biochemical/clinical failure (FFF), and 2) it is expected that less than one event (actual 0.8) will be observed at the completion of Phase II component. Moreover, the futility analysis was not included at the second (Year 4.5) or the third (Year 5.6) interim analyses because the recruitment will have been completed at Year 3. It simply may not be feasible to include a futility analysis in clinical trials of newly-diagnosed prostate cancer.

Despite the challenges mentioned above, the theme of adaptive approaches were utilized whenever possible. As required by the NIH Recombinant DNA Advisory Committee for safety evaluation, unequally spaced interim looks were proposed which were not only built in for the first look when the initial 21 patients in Arm 1 completed the 90-day toxicity assessment (for the Phase II for patient's safety), but also incorporated two additional interim looks assuming a maximum 79 events of FFF from the entire study for possible treatment efficacy.

A cohort of 21 patients for the Phase II component is based on no treatment-related SAEs observed in the completed Phase I trials, with a cumulative 50 patients to date (not shown), which will ensure the upper 95% confidence limit <0.14 [19] assuming no treatment related SAEs.

The second interim look is expected at the study duration of 4.5 years when 27 events are observed and at least 180 patients have completed the 2-year biopsy assessments. Based on recent data from the Memorial Sloan Kettering Cancer Center (MSKCC) [34] with an average 3.1 year assessment on 6 biopsy cores , a 27% positive prostate biopsy rate at year 2 is expected in controls after adjustment for a shorter assessment time, 12 biopsy cores and 80 Gy compared to MSKCC, and about a 10% positive biopsy rate in the gene-therapy + IMRT treated group. It is more likely to have sufficient power (e.g. 80% or more) to detect the treatment effect on the 2-year biopsy at the second interim look. Another challenge will be encountered: the second efficacy endpoint for this trial requires a smaller sample size within a much shorter time period, compared to the primary endpoint. We will seek for external DSMB and/or FDA approval at that time to unblind the 2-year biopsy results before we conduct the second interim analysis, and will follow the external DSMB guidance on this issue.

The third interim look is expected at the study duration 5.6 years (Figure 2) when 54 events are observed and will include the events estimation/sample size re-adjustment, which is an adaptive approach and is especially useful for a trial with long study duration. It is not uncommon that the initial estimated parameters may differ from what is actually observed in the trial, which may lead to an under- or over-powered trial. The ReCAP trial is powered based on events rather than the number of patients enrolled. From the current design, only 28% of patients are expected to have an event at the proposed study duration of 6.8 years.

The trial should have sufficient samples to obtain more events, given the nature of the endpoint. At the third interim analysis, we will present event re-estimation to the external DSMB for approval and we will follow the DSMB's recommendation on time of study duration to conduct the final efficacy analysis.

We are pleased that both the RAC committee and the FDA were in favor of this adaptive seamless Phase II/III design for the ReCAP trial and that the trial was approved by the FDA for enrollment. Although, at this point, the efficacy endpoint FFF can not be an acceptable endpoint for a registration trial to support product licensure; it can be used as an endpoint in a "proof of principle", non-registration trial. We remain optimistic that the composite biochemical and clinical endpoints (FFF), rather than overall survival, may become an acceptable alternative endpoint in the future to support product licensure in clinical trials of newly-diagnosed prostate cancer. In summary, the adaptive seamless, Phase II/III design facilitates the development of novel therapies by shortening the interval between phases of a traditional trial design, and reduces the number of required patients and long-term follow-up. Therefore, it is a time-saving and cost-effective design. However, it requires a specified statistical plan to guide the decision making processes necessary in this trial design.

Acknowledgments

Reference List

1. Pollack A, Zagars GK, Starkschall G, Antolak JA, Lee JJ, Huang E, von Eschenbach AC, Kuban DA, Rosen I. Prostate cancer radiation dose response: results of the M. D. Anderson phase III randomized trial. Int J Radiat Oncol Biol Phys. 2002; 53:1097–1105. [PubMed: 12128107]

2. Zietman AL, DeSilvio ML, Slater JD, Rossi CJ Jr, Miller DW, Adams JA, Shipley WU. Comparison of conventional-dose vs high-dose conformal radiation therapy in clinically localized adenocarcinoma of the prostate: a randomized controlled trial. JAMA. 2005; 294:1233–1239. [PubMed: 16160131]

3. Bolla M, Gonzalez D, Warde P, Dubois JB, Mirimanoff RO, Storme G, Bernier J, Kuten A, Sternberg C, Gil T, Collette L, Pierart M. Improved survival in patients with locally advanced prostate cancer treated with radiotherapy and goserelin. N Engl J Med. 1997; 337:295–300. [PubMed: 9233866]

4. D'Amico AV, Manola J, Loffredo M, Renshaw AA, DellaCroce A, Kantoff PW. 6-month androgen suppression plus radiation therapy vs radiation therapy alone for patients with clinically localized prostate cancer: a randomized controlled trial. JAMA. 2004; 292:821–827. [PubMed: 15315996]

5. Holzbeierlein JM, Castle E, Thrasher JB. Complications of androgen deprivation therapy: prevention and treatment. Oncology (Williston Park). 2004; 18:303–309. [PubMed: 15065701]

6. Sanda MG, Dunn RL, Michalski J, Sandler HM, Northouse L, Hembroff L, Lin X, Greenfield TK, Litwin MS, Saigal CS, Mahadevan A, Klein E, Kibel A, Pisters LL, Kuban D, Kaplan I, Wood D, Ciezki J, Shah N, Wei JT. Quality of life and satisfaction with outcome among prostate-cancer survivors. N Engl J Med. 2008; 358:1250–1261. [PubMed: 18354103]

7. Freytag SO, Rogulski KR, Paielli DL, Gilbert JD, Kim JH. A novel three-pronged approach to kill cancer cells selectively: concomitant viral, double suicide gene, and radiotherapy. Hum Gene Ther. 1998; 9:1323–1333. [PubMed: 9650617]

8. Freytag SO, Movsas B, Aref I, Stricker H, Peabody J, Pegg J, Zhang Y, Barton KN, Brown SL, Lu M, Savera A, Kim JH. Phase I trial of replication-competent adenovirus-mediated suicide gene therapy combined with IMRT for prostate cancer. Mol Ther. 2007; 15:1016–1023. [PubMed: 17375076]

9. Freytag SO, Stricker H, Movsas B, Kim JH. Prostate cancer gene therapy clinical trials. Mol Ther. 2007; 15:1042–1052. [PubMed: 17406342]

10. Gallo P, Chuang-Stein C, Dragalin V, Gaydos B, Krams M, Pinheiro J. Adaptive designs in clinical drug development--an Executive Summary of the PhRMA Working Group. J Biopharm Stat. 2006; 16:275–283. [PubMed: 16724485]

11. Kuehn BM. Industry, FDA warm to "adaptive" trials. JAMA. 2006; 296:1955–1957. [PubMed: 17062852]

12. Gallo P, Maurer W. Challenges in implementing adaptive designs: comments on the viewpoints expressed by regulatory statisticians. Biom J. 2006; 48:591–597. [PubMed: 16972710]

13. Spalding AC, Daignault S, Sandler HM, Shah RB, Pan CC, Ray ME. Percent positive biopsy cores as a prognostic factor for prostate cancer treated with external beam radiation 2. Urology. 2007; 69:936–940. [PubMed: 17482938]

14. Horwitz EM, Thames HD, Kuban DA, Levy LB, Kupelian PA, Martinez AA, Michalski JM, Pisansky TM, Sandler HM, Shipley WU, Zelefsky MJ, Hanks GE, Zietman AL. Definitions of biochemical failure that best predict clinical failure in patients with prostate cancer treated with external beam radiation alone: a multi-institutional pooled analysis. J Urol. 2005; 173:797–802. [PubMed: 15711272]

15. Wei JT, Dunn RL, Litwin MS, Sandler HM, Sanda MG. Development and validation of the expanded prostate cancer index composite (EPIC) for comprehensive assessment of health-related quality of life in men with prostate cancer. Urology. 2000; 56:899–905. [PubMed: 11113727]

16. Rabin R, de CF. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001; 33:337–343. [PubMed: 11491192]

17. O'Brien PC, Fleming T. A Multiple Testing Procedure for Clinical Trials. Biomatric. 1979; 35:549–556.

18. Lan K, DeMets D. Discrete Sequential Boundaries for Clinical Trials. Bimmetrika. 1983; 70:659–663.

19. Newman TB. If almost nothing goes wrong, is almost everything all right? Interpreting small numerators, 2. JAMA. 1995; 274:1013. [PubMed: 7563445]

20. Pollack A, Zagars GK, Smith LG, Lee JJ, von Eschenbach AC, Antolak JA, Starkschall G, Rosen I. Preliminary results of a randomized radiotherapy dose-escalation study comparing 70 Gy with 78 Gy for prostate cancer. J Clin Oncol. 2000; 18:3904–3911. [PubMed: 11099319]

21. Pollack A, Zagars GK, Antolak JA, Kuban DA, Rosen II. Prostate biopsy status and PSA nadir level as early surrogates for treatment failure: analysis of a prostate cancer randomized radiation dose escalation trial. Int J Radiat Oncol Biol Phys. 2002; 54:677–685. [PubMed: 12377318]

22. Rothman, K.; Greenland, S. Modern Epidemiology. Lippincott-Raven; Philadelphia: 1998.

23. FDA. Guidance for Industry Adaptive Design Clinical Trials for Drugs and Biologics (DRAFT GUIDANCE). 2010.

24. Pound CR, Partin AW, Eisenberger MA, Chan DW, Pearson JD, Walsh PC. Natural history of progression after PSA elevation following radical prostatectomy. JAMA. 1999; 281:1591–1597. [PubMed: 10235151]

25. Renshaw AA, Schultz D, Cote K, Loffredo M, Ziemba DE, D'Amico AV. Accurate Gleason grading of prostatic adenocarcinoma in prostate needle biopsies by general pathologists. Arch Pathol Lab Med. 2003; 127:1007–1008. [PubMed: 12873176]

26. Sandler HM, DeSilvio ML. Surrogate end points for prostate cancer: what is prostate-specific antigen telling us? J Natl Cancer Inst. 2003; 95:1352–1353. [PubMed: 13130102]

27. Collette L. Prostate-Specific Antigen (PSA) as a Surrogate End Point for Survival in Prostate Cancer Clinical Trials. European Urology. 2008; 53:6–9. [PubMed: 17764823]

28. Kuban DA, Tucker SL, Dong L, Starkschall G, Huang EH, Cheung MR, Lee AK, Pollack A. Long-term results of the M. D. Anderson randomized dose-escalation trial for prostate cancer. Int J Radiat Oncol Biol Phys. 2008; 70:67–74. [PubMed: 17765406]

29. Valicenti RK, DeSilvio M, Hanks GE, Porter A, Brereton H, Rosenthal SA, Shipley WU, Sandler HM. Posttreatment prostatic-specific antigen doubling time as a surrogate endpoint for prostate cancer-specific survival: an analysis of Radiation Therapy Oncology Group Protocol 92–02. Int J Radiat Oncol Biol Phys. 2006; 66:1064–1071. [PubMed: 16979837]

30. Peeters ST, Heemsbergen WD, Koper PC, van Putten WL, Slot A, Dielwart MF, Bonfrer JM, Incrocci L, Lebesque JV. Dose-response in radiotherapy for localized prostate cancer: results of

the Dutch multicenter randomized phase III trial comparing 68 Gy of radiotherapy with 78 Gy. J Clin Oncol. 2006; 24:1990–1996. [PubMed: 16648499]

31. Wong WW, Schild SE, Vora SA, Halyard MY. Association of percent positive prostate biopsies and perineural invasion with biochemical outcome after external beam radiotherapy for localized prostate cancer. Int J Radiat Oncol Biol Phys. 2004; 60:24–29. [PubMed: 15337536]

32. Vance W, Tucker SL, de CR, Kuban DA, Cheung MR. The predictive value of 2-year posttreatment biopsy after prostate cancer radiotherapy for eventual biochemical outcome. Int J Radiat Oncol Biol Phys. 2007; 67:828–833. [PubMed: 17161554]

33. Crook J, Malone S, Perry G, Bahadur Y, Robertson S, Abdolell M. Postradiotherapy prostate biopsies: what do they really mean? Results for 498 patients. Int J Radiat Oncol Biol Phys. 2000; 48:355–367. [PubMed: 10974448]

34. Zelefsky MJ, Reuter VE, Fuks Z, Scardino P, Shippy A. Influence of local tumor control on distant metastases and cancer related mortality after external beam radiotherapy for prostate cancer. J Urol. 2008; 179:1368–1373. [PubMed: 18289585]
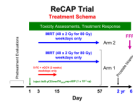
Figure 1. Trial design

Eligible patients are randomized to one of two treatment arms with the investigated treatment (the adenovirus injection on Day 1 and prodrugs beginning on Day 3 for 10 days concomitant with IMRT for 8 weeks) in Arm 1 or IMRT only in Arm 2. Study endpoints including prostate biopsy at 2 years and freedom from failure at 6 years are described in the text.

Figure 2. Four unequal spaced pre-specified interim looks
The study is designed to include four unequally spaced pre-specified interim looks. The first will occur when 21 patients in Arm 1 complete the toxicity assessment at the end of 90 days. The second and third looks are scheduled when 1/3 and 2/3 of the maximum number of events (79) are observed. The last look will occur when the number of expected or maximum event are observed (final analysis).
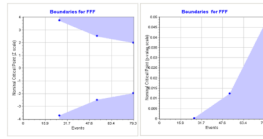
Figure 3. Cross boundaries of Z-value/P-value for FFF using
EAST 5.1 based on 280 patients, an accrual rate of 90 per year, a 2-sided test with alpha = 0.05, four unequal spaced looks using O'Brien-Fleming boundary, and a power of 80%. If the test statistic or *p*-value crosses into the blue zone, the trial will be stopped and the alterative hypothesis is accepted.

Table 1

ReCAP Trial Patients' Assessments

Parameters	Pre-Study Therapy	During Study Therapy	Follow Up
History, Physical Examination, DRE	X[a]		X[h]
EPIC and EQ-5D Instruments	X[a]		X[i]
Hematology (CBC/DP)	X[b]	X[f]	X[h]
Coagulations (PT, PTT)	X[b]		
Blood Chemistries	X[b]	X[f]	X[h]
Urinalysis	X[b]		
Diagnostic Serum PSA	X[b]		X[h]
Serology- Hepatitis B and C, HIV	X[b]		
Toxicity Assessment	X[b]	X[f]	X[h]
Molecular Studies	X[b]	X[g]	X[g]
Chest Radiographs	X[c]		
TRUS-Guided Needle Biopsy	X[d]		X[j]
Pathology Assessment	X[d]		X[j]
Bone Scan	X[e]		X[k]
CT Scan	X[e]		X[l]
On-site Physician Assessment		X[f]	

[a] Within 90 days of registration.

[b] Within 30 days of registration.

[c] Within 90 days of registration if the subject has history of pulmonary disease; otherwise, within 180 days.

[d] Within 180 days of registration.

[e] Within 90 days of registration.

[f] Once a week during study therapy/IMRT course.

[g] At every scheduled blood draw until not detected in two consecutive measurements.

[h] At every follow-up visit after completion of IMRT: 1, 3, 6, 9, 12, 18, 24 months and then annually thereafter.

[i] At 6, 12, 24, 36 and 60 months after completion of IMRT.

[j] At 24 months after completion of IMRT.

[k] In patients who experience bone pain or at PSA relapse and then annually thereafter.

[l] At PSA relapse and then annually thereafter.