Published in final edited form as:
Int J Radiat Oncol Biol Phys. 2010 December 1; 78(5): 1314–1322. doi:10.1016/j.ijrobp.2009.09.073.

Baseline Serum Testosterone in Men Treated with Androgen Deprivation Therapy and Radiotherapy for Localized Prostate Cancer

Mack Roach III, MD[1], **Kyounghwa Bae, PhD**[2], **Colleen Lawton, MD**[3], **B. J. Donnelly, MD**[4], **David Grignon, MD**[5], **Gerald E. Hanks, MD**[6], **Arthur Porter, MD**[7], **Herbert Lepor, MD**[8], **Varagur Venketesan, MD**[9], and **Howard Sandler, MD**[10]

[1]University of California San Francisco, San Francisco CA

[2]Department of Statistics RTOG, Philadelphia PA

[3]Medical College of Wisconsin

[4]University of Calgary

[5]Indiana University, Indianapolis IN

[6]Fox Chase Cancer Center, Health Science

[7]McGill University

[8]NYU School of Medicine NY

[9]London Regional Cancer Program and University of Western Ontario

[10]Cedar-Sinai, Los Angeles CA

Abstract

Introduction—It is believed that men diagnosed with prostate cancer and a low baseline serum testosterone (BST) may have more aggressive disease and it is frequently recommended they forego testosterone replacement therapy. We used two large phase III trials involving androgen deprivation therapy (ADT) and radiotherapy (EBRT) to assess the significance of a BST.

Materials and Methods—All patients with a BST and complete data (n=2,478) were included in this analysis and divided into four categories: "Very low BST" (VLBST) ≤ 16.5th percentile of BST (≤248 ng/dl)(n=408); "Low BST" (LBST) >16.5th percentile and ≤ 33rd percentile (>248ng/dl but ≤314 ng/dl) (n=415); "Average BST" (ABST) > 33rd percentile and ≤ 67th percentile (314 to 437 ng/dl) (n=845); and "High BST" (HBST) > 67th percentile (>437 ng/dl) (n=810). Outcomes included: overall survival (OS), distant metastasis (DM), biochemical failure (BF), and cause-specific survival (CSS). All outcomes were adjusted for the following covariates; treatment arm,

Corresponding Author: Mack Roach III, M.D., UCSF, Helen Diller Comprehensive Cancer Center, University of California San Francisco, 1600 Divisadero Street, Suite H1031, San Francisco, California 94143-1708, 415-353-7181, FAX #-353-7182, mroach@radonc.ucsf.edu.

Conflict of Interest Statement:
None of the authors have any conflicts of interest to disclose relevant to this manuscript.

BST, age [< 70 vs. ≥ 70], PSA [< 10 vs. 10 ≤ PSA < 20 vs. 20 ≤], Gleason score (GS) [2–6 vs. 7 vs. 8–10]; T stage [T1–T2 vs. T3–T4], and KPS performance status (60–90 vs. 100).

Results—On multivariable analysis age, GS and PSA were independently associated with an increased risk of BF, DM and a reduced CSS and OS ($p<<0.05$), but BST was not.

Conclusions—BST does not impact the outcomes of men treated with EBRT and ADT for prostate cancer.

Keywords

Prostate Cancer; serum testosterone; androgen deprivation therapy; radiation therapy; clinical trials; prognostic factors

Introduction

Prostate cancer is the most commonly diagnosed non-cutaneous malignancy in the American male. A significant number of these men present with low or borderline baseline serum testosterone levels and some of are using testosterone supplements at the time of diagnosis. There are three challenges concerning the management of such patients. First, is the issue of how a low baseline serum testosterone (BST) influences prognosis. Second, whether there is reason to believe hormone replacement therapy should be discontinued and finally, when treatment is completed, whether replacement therapy should be reinstituted.

A number of retrospective studies have concluded that men with clinically localized prostate cancer presenting with a low BST level have a worse outcome. Several of these studies included men with metastatic disease and suggested that their poor outcome was a manifestation of an intrinsic state of androgen independence [1–4]. Low BST levels have also been associated with poor outcomes in patients undergoing radical prostatectomy [5, 6]. These patients appeared to present with high-grade disease that was associated with a higher risk of recurrence [7, 8]. These studies supported an association between a low testosterone and more aggressive prostate cancer but did not include patients treated on prospective randomized trials [4, 7, 9]. In addition, treatment and follow-up were not standardized on these studies.

Not all studies support the assertion that a low BST is an adverse prognostic factor. For example, Armstrong et al. noted no relationship between baseline testosterone and biochemical failure 10. Among a group of 33 patients with metastatic prostate cancer treated with orchiectomy, the third with the lowest testosterone showed a better survival than the ones with a higher testosterone. Also of note, some studies suggest that a delay in the time to recovery of testosterone in patients managed with androgen deprivation therapy and radiation might be associated with a more favorable outcome [11]. This observation raises the possibility that patients with a low BST level might benefit because this low testosterone level may lead to delayed recovery of testosterone and a prolonged disease free period. This issue has particular relevance for men managed with androgen deprivation therapy (ADT) in combination with radiotherapy for intermediate and high-risk disease. For these men ADT is commonly prescribed for periods of four months to three years or even more, with little guidance as to how their testosterone levels should be managed when ADT is completed [12].

Thus, there are conflicting studies regarding the significance of BST levels in patients with adenocarcinoma of the prostate. The studies alluded to tended to be relatively small, the patients heterogeneous, non-randomized, and most of these studies had relatively short follow-up. In an attempt to address this issue definitively, we pooled data from two large phase III randomized trials from Radiation Therapy Oncology Group (RTOG) involving

ADT and external beam radiation therapy (EBRT) in an attempt to determine the prognostic significance of a low BST level.

Materials and Methods

RTOG 9202 and 9413 have been previously published and the details of those studies are found elsewhere 13, [14]. Briefly, RTOG 9202 included approximately 1500 patients, half of whom (Arm 1) received 4 months of ADT and the other half (Arm 2) 28 months of ADT in combination with EBRT. RTOG 9413 had four arms studying the relationship between the sequence and the volume of radiotherapy (pelvic vs. no pelvic irradiation) in approximately 1300 patients treated for locally advanced prostate cancer. In both trials patients received nearly 7000 cGy. Patients who were surgically staged were ineligible for RTOG 9202, as were patients with distant metastasis. Additional criteria included: Karnofsky performance status ≥ 70, no prior hormonal therapy, radiation or chemotherapy and liver function tests ≤ 1.2 times upper limits of normal. Patients with histologically confirmed adenocarcinoma of the prostate, clinical stage T2c (bilobar) to T4, with no involved nodes in the common iliac or higher node chains, with Karnofsky performance score ≥70 and with pretreatment PSA less than 150 ng/ml were eligible.

For RTOG 9413 eligibility included histologically confirmed clinically localized (negative bone and CT scans) adenocarcinoma of the prostate with an elevated PSA ≤ 100 ng/ml. Patients were stratified by T stage: T1c, T2a vs T1b, T2b vs T2c-T4; PSA (≤ 30 vs > 30 ng/ml) and Gleason Score (GS) (GS< 7 vs 7–10) and no prior hormonal therapy was allowed. Eligible patients were required to have an estimated risk of lymph node (LN) involvement >15%, based on the equation +LN = (2/3) PSA + [(GS − 6) × 10] [15]. Patients with T2c-T4 were also eligible if their GS ≥ 6 even if by the equation their risk did not reach 15% based on their risk as reported by Partin et al 16. A BST was required at the time of entry on this study and for the purposes of this analysis, we used data from all eligible patients for each trial with an available BST to determine the impact on outcome.

Statistical Methods

We compared the pretreatment characteristics and outcomes of patients with and without missing data by Chi-square test statistics to see if there was bias between the two patients groups. Chi-square test statistics were also used to compare pretreatment characteristics. We performed heterogeneity testing for this analysis to see if one estimate could be used to represent the combined data from different trials (homogeneous) or not (heterogeneous). Tests of heterogeneity of testosterone for group differences across arms (the stratification variable) were performed for overall survival (OS), distant metastasis (DM), biochemical failure (BF) and cause-specific survival (CSS) using Chi-square test statistics. If the arms are homogeneous, the pooled hazard ratio (HR) was used as the estimator for the combined data [17, 18]. The failure event of OS was a death due to any cause. Biochemical failure was defined using the “Phoenix Definition” and the failure event of cause-specific survival was a death related to prostate cancer or as a complication of protocol treatment 19. Distant metastasis was defined as clinical evidence of distant disease by any method. Time to failure is measured from the date of randomization to the date of the first failure event.

The Kaplan-Meier method was used to estimate the OS rate, and the log-rank test was used to test the difference between the treatments or categories in the univariate analysis 20 of OS. The cumulative incidence method was used to estimate the DM rate, BF rate, and CSS rate, and Gray’s test was used to test the difference between the treatments or categories in the univariate analysis21 of these endpoints. Cox proportional hazards regression model was used for OS, and Fine and Gray’s regression model was used for DM, BF, and CSS 22 [23] to adjust for other covariates associated with outcomes in the model. Unadjusted and adjusted

hazard ratios were calculated for all covariates using either the Cox proportional hazards model or Fine and Gray's regression model with associated 95% confidence intervals (C.I.s) and p-values. . Co-variants in this analysis included baseline serum testosterone level as either a continuous or a categorical variable, treatment arm [9202 Arm 1 [reference level; RL] vs. 9202 Arm 2 vs. 9413 Arm 1 vs. 9413 Arm 2 vs. 9413 Arm 3 vs. 9413 Arm 4], age [< 70 (reference level (RL) vs. ≥ 70], PSA [PSA < 10 (RL) vs. 10 ≤ PSA < 20 vs. 20 ≤ PSA], Gleason score [2–6 (RL) vs. 7 vs. 8–10]; T stage [T1–T2 (RL) vs T3–T4], and KPS performance status [60–90 (RL) vs. 100]. Baseline testosterone (BST) levels were initially divided into tertiles, but because most published series have implied that very low baseline testosterone levels were more likely to be clinically relevant, the lowest tertile was further subdivided into two. Thus four categories were generated: "Very low baseline testosterone" (VLBST) group that had a ≤ 16.5th percentile of BST, the "Low testosterone" (LBST) group had a testosterone that >16.5th percentile and ≤ 33rd percentile, the "Average testosterone" (ABST) group had a level that was > 33rd percentile and ≤ 67th percentile, and "High testosterone" (HBST) group that was greater than > 67th percentile. As an additional exploratory exercise to determine if men with even lower levels of testosterone, i.e., the lowest 5%, might have more aggressive disease we compared this group (n=124) with the rest of patients (n=2354). However, since this did not alter our findings the data are not shown.

Statistical Analysis System (SAS Institute, Cary, NC) was used for all statistical analyses except Gray's testing and Fine and Gray's modeling which was analyzed using R software. A p-value <0.05 was considered as statistically significant.

Results

A total of 223 (8.3%) patients were excluded from 2701 patients; 129 of the 223 patients were excluded due to missing testosterone information. In total, complete data was available for 2,478 (92%) of the patients from RTOG 9202 and 9413. Table 1 shows the outcomes between the groups with and without missing data. The results of the analyses show that there are not statistically significantly differences between the two groups (p-values >0.05). To address the trend for an association between Gleason Score and BST (p=0.07, Table 1) we performed additional statistical evaluations. Pearson correlation between BST group and GS was −0.0196, this indicates that there is no strong directional correlation between the two. Also the p-value from F-test to see if there is a mean BST difference among three GS groups is 0.875, which indicates that there is no difference with respect to mean BST among three GS group. Therefore imputation was not done for this analysis.

There were 408 patients in the VLBST group. This group made up the bottom 16.5% of patients with a serum testosterone less than or equal to 248 ng/dl (mean=189.5 ng/dl). The LBST group had 415 patients and their serum testosterone was greater than 248 but less than or equal to 314 ng/dl (mean= 282.2 ng/dl). There were 845 patients in the ABST group and had a level that was greater than 314 to 437 ng/dl (mean= 372 ng/dl) and the HBST group had 810 patients whose serum testosterone was greater than 437 ng/dl (mean= 576 ng/dl) (Table 2).

There was no statistical bias among the four BST group and age, Gleason score, or PSA. However, there is a statistically significant difference in KPS and treatment arm among the four BST groups (Table 2 p-value= 0.03 and 0.04, respectively). As is shown in this Table the mean age for men, T-Stage and pretreatment PSA levels among men treated in the VLBST group was similar to other groups. Patients in the VLBST group had a similar median follow up of 7.2 years, compared to 7.1, 7.3 and 7.2 years in patients with low, average or high testosterone.

The results of heterogeneity testing are shown in Table 3. The six arms from the two studies (RTOG 9202 and 9413) were homogeneous with respect to the hazard ratios (HRs) of OS, DM, BF, and CSS (all p-values > 0.05). These tables show the pooled HR of testosterone groups VLBST (RL) vs. LBST vs. ABST vs. HBST) of each outcome with and without adjusting for other covariates. None of the outcomes show a statistically significant difference among the four BST groups because the 95% CI of each HR included 1. The results from univariate proportional hazards analyses of the overall impact of baseline testosterone level on OS, DM, BF, and CSS are summarized in Table 4. None of the outcomes show statistically significant differences (p-values >0.05), which is the same as the results from the pooled HRs.

The multivariate analysis for OS, DM, BF and CSS are summarized in Tables 5a–d. The multivariate analysis for OS is shown in Table 5a. Age < 70, Gleason score 2–6, KPS of 100 were independently associated with an OS advantage. Of note, pretreatment PSA ≥ 20 ng/ml and Arm 3 of RTOG 9413 were marginally associated with worse OS (p=0.07). The multivariate analysis for DM is shown in Table 5b. Age <70, Gleason score > 2–6, PSA ≥ 20 ng/ml were independently associated with an increased risk of DM (p<0.0001, 0.004, <0.0001, and 0.01, for age, Gleason 7 and 8–10, and PSA ≥ 20 ng/ml, respectively. In addition, reductions in the risk of DM were noted by treatment arms, for RTOG 9202 Arm 2, RTOG Arms 1, 2, 3, and 4 compared to Arm 1 of RTOG 9202 (p <0.001, <0.001, 0.01, p=0.06 and 0.002, respectively). The multivariate analysis for this BF is shown in Table 5c. Age < 70, Gleason score > 2–6, and PSA ≥ 10 ng/ml were independently associated with an increased risk of BF (p<0.0001, 0.0004 and <0.0001, and <0, <0001 and <0.0001 for age, Gleason 7 and 8–10 and PSA ≥10-<20 and ≥20 ng/ml, respectively). Differences were also noted by treatment arms, for RTOG 9202 Arm 2, and RTOG Arms 1, 3, and 4 compared to Arm 1 of RTOG 9202 (p <0.0001, 0.0005, p=0.009 and 0.0002, respectively). The multivariate analysis for this CSS is shown in Table 5d. Age < 70, Gleason score > 7, and PSA ≥ 20 ng/ml were all independently associated with a reduced CSS (p=0.0005, <0.0001, and p=0.02, for age, GS 8–10 and PSA ≥ 20 ng/ml, respectively). Differences were also noted by treatment arms, for RTOG 9202 Arm 2, and RTOG Arms 1, 2, 3, and 4 compared to Arm 1 of RTOG 9202 (p=0.03, 0.0002, p=0.04, p=0.02 and 0.001, respectively). Of note however, the BST group was not statistically significantly associated with any of the outcomes studied (p-values >0.05).

Discussion

Testosterone levels alone are not considered adequate for diagnosing hypogonadism but men with a total testosterone less than 200 ng/dl with symptoms are usually considered candidates for testosterone replacement therapy 24, 25. Other studies define androgen-deficient levels as a total testosterone less than 320 ng/dl 26. In this study we chose to define a low serum testosterone level as less than or equal to 314 ng/dl but greater than 248 ng/dl. Those with a very low baseline serum testosterone (VLBST) were defined as less than or equal to 248 ng/dl. Thus our criteria for low and very low are not inconsistent with the literature even though they were chosen based on the distribution of values (lowest tertile) observed. In this study we found no association between the baseline serum testosterone level and outcome. This observation was unexpected given the fact that some studies suggest that a low baseline serum testosterone might be associated with an increased mortality rate independent of a diagnosis of prostate cancer 27, 28. Our findings are consistent with a review of testosterone levels and aging published by the Institute of Medicine (2004) that concluded “Endogenous testosterone levels clearly decline with aging, but it is not clear if lower levels of serum testosterone affect health outcomes in older men” 24.

To our knowledge this is the largest prospective analysis of the pre-treatment significance of the baseline serum testosterone in men treated on prospective phase III randomized trials for clinically localized prostate cancer. Prior studies have been complicated by small sample sizes, retrospective selection of data, and inclusion of heterogeneous patient groups, (some with metastatic disease and others with localized disease). In this analysis we evaluated all cause survival, distant metastatic disease, biochemical (PSA) failure and prostate cancer specific survival. We observed no relationship between baseline testosterone level and OS, CSS, DM or PSA failure. We did however observe that that older men tend to have less aggressive prostate cancer. This finding is consistent with the results of a study in which men were treated by radical prostatectomy and raises unanswered questions about potential mechanisms [29]. Not surprisingly, the pretreatment PSA and Gleason score were also major predictors of cause specific survival (Table 5d).

Recent studies support the health benefits of addressing hypogonadism. For example, Coward et al. recently described the outcomes of 81 hypogonadal men (mean age 56.8 years) treated with testosterone replacement therapy 30. With a mean follow-up of 33.8 months (range, 6–144) they noted that after starting testosterone replacement therapy the total cholesterol improved from 203.8 to 166.6 mg/dL ($P < 0.05$) after 36 months, and the incidence of prostate cancer among men appears to be no greater than that in the general population. Our findings provide additional support for the growing body of literature questioning the merits of withholding testosterone replacement therapy in men with prostate cancer 31. For example, Morgentaler recently reviewed the scientific and ethical considerations associated with testosterone replacement therapy. He highlighted the fact that due to the saturation of the androgen receptor binding capacity, higher concentrations of androgen do not results in greater androgen binding and discussed a body of literature that demonstrates a lack of correlation with PSA levels and the risk of prostate cancer. He also summarized the findings from 3 series that reported the outcomes of patients treated (surgery, in two series and brachytherapy) for prostate cancer who subsequently under went testosterone replacement therapy 32–34. In his review he noted biochemical recurrences in only 1.8% of men (2 of 111) who received testosterone replacement therapy, which would appear to be a lower rate of recurrence than usually observed 35. More recently still, Morales et al. have reported on five men with testosterone deficiency syndrome (TDS) after external beam radiotherapy (EBRT) for localized prostate cancer who were treated with testosterone once their PSAs had reached their nadir 36. With a follow-up of 14.5 months (6–27) one of the patients had a transitory increase in PSA level but none had levels of >1.5 ng/mL. The patients generally reported marked improvement in hot flushes, fatigue, and libido, with two reporting improved erectile function.

There are a number of potential limitations to our study. First, it is possible that only men with extremely low levels of testosterone might have more aggressive disease. To test this hypothesis we re-ran our analysis defining very low risk as the lowest 5% (n=124) but also saw no evidence that this subset was any different than the bottom 16.5% used in this analysis (data not shown). Perhaps with longer follow-up an impact on survival would appear, but this seems unlikely since patients with androgen independent usually have a much shorter survival than those with androgen dependent disease. The fact that no differences were seen in biochemical (PSA) failure based on the baseline testosterone level would also seem to make this possibility very unlikely.

Another potential limitation is that this study may not be generalized to all men with prostate cancer because the current study can only help answer the questions regarding baseline testosterone levels in prostate cancer patients undergoing EBRT and ADT. It is possible that these findings should not be extrapolated to men managed with radiation therapy alone. However, since androgen deprivation therapy is thought to be neither indicated nor

beneficial in low risk patients there is no obvious reason to suspect that the findings would be different for this subset of patients. The continuation of testosterone replacement therapy in low-risk patients would theoretically be safer than in intermediate-risk and high-risk patients because of their lower risk of death from prostate cancer. However, since low-risk patients were not included in this study, we have no direct evidence that low risk patient can safely continue testosterone replacement therapy during radiotherapy. Based on our data, one can simply conclude that BST levels did not affect prognosis. Of note however, is the recent report by Taira et al. that suggests that our conclusions may be equally relevant to low risk patients [37]. To our knowledge no one in the RTOG plans to test the hypothesis that testosterone replacement therapy during radiotherapy improves the quality of life of low-risk prostate cancer patients.

Another limitation of this study is that based on our data we are unable to answer the questions as to whether hormone replacement therapy should be discontinued in newly diagnosed patients or whether when treatment is completed, whether replacement therapy should be reinstituted. It seems logical to assume however, that since BST does not impact outcome, then a newly diagnosed patient with low risk disease (for whom ADT would not be indicated) could probably safely continue testosterone replacement therapy. The issue of re-initiating testosterone after completion of treatment in a patient whose treatment includes ADT is more problematic because it is sometime difficult to determine disease free status in a patient with a low serum testosterone.

In conclusion, we were unable to confirm the assertion that men with very low baseline testosterone levels have more aggressive prostate cancer. Since baseline serum testosterone levels do not appear to influence clinically significant outcomes in men with intermediate to high risk disease and in light of other recent observations we question the common practice of recommending discontinuation of testosterone replacement therapy in all men with newly diagnosed low risk prostate cancer. The point is, a mildly hypogonadal state is probably inadequate protection against the growth of occult metastatic disease, while a normal level may be no more dangerous and may be associated with an improvement in quality of life. In contrast however, patients with metastatic disease may require castrate levels to be spared the sequel of disease progression.

Acknowledgments

Supported by RTOG U10 CA21661, CCOP U10 CA37422, and Stat U10 CA32115 grants from the NCI. This manuscript's contents are the sole responsibility of the authors and do not necessarily represent the official views of the NCI. The human investigations were performed after approval by a local Human Investigations Committee and in accordance with an assurance filed with and approved by the Department of Health and Human Services

References

1. Imamoto T, Suzuki H, Fukasawa S, et al. Pretreatment serum testosterone level as a predictive factor of pathological stage in localized prostate cancer patients treated with radical prostatectomy. Eur Urol 2005;47:308–312. [PubMed: 15716191]
2. Iversen P, Rasmussen F, Christensen IJ. Serum testosterone as a prognostic factor in patients with advanced prostatic carcinoma. Scand J Urol Nephrol Suppl 1994;157:41–47. [PubMed: 7939452]
3. Ribeiro M, Ruff P, Falkson G. Low serum testosterone and a younger age predict for a poor outcome in metastatic prostate cancer. Am J Clin Oncol 1997;20:605–608. [PubMed: 9391550]
4. Daniell HW. A worse prognosis for men with testicular atrophy at therapeutic orchiectomy for prostate carcinoma. Cancer 1998;83:1170–1173. [PubMed: 9740082]
5. Yamamoto S, Yonese J, Kawakami S, et al. Preoperative serum testosterone level as an independent predictor of treatment failure following radical prostatectomy. Eur Urol 2007;52:696–701. [PubMed: 17412490]

6. Massengill JC, Sun L, Moul JW, et al. Pretreatment total testosterone level predicts pathological stage in patients with localized prostate cancer treated with radical prostatectomy. J Urol 2003;169:1670–1675. [PubMed: 12686805]
7. Schatzl G, Madersbacher S, Thurridl T, et al. High-grade prostate cancer is associated with low serum testosterone levels. Prostate 2001;47:52–58. [PubMed: 11304729]
8. D'Amico AV, Chen MH, Malkowicz SB, et al. Lower prostate specific antigen outcome than expected following radical prostatectomy in patients with high grade prostate and a prostatic specific antigen level of 4 ng/ml. Or less. J Urol 2002;167:2025–2030. discussion 2030-2021. [PubMed: 11956431]
9. Denberg TD, Flanigan RC, Kim FJ, et al. Self-reported volume of radical prostatectomies among urologists in the USA. BJU Int 2007;99:339–343. [PubMed: 17155974]
10. Armstrong AJ, Halabi S, de Wit R, et al. The relationship of body mass index and serum testosterone with disease outcomes in men with castration-resistant metastatic prostate cancer. Prostate Cancer Prostatic Dis. 2008
11. D'Amico AV, Renshaw AA, Loffredo B, et al. Duration of testosterone suppression and the risk of death from prostate cancer in men treated using radiation and 6 months of hormone therapy. Cancer 2007;110:1723–1728. [PubMed: 17828774]
12. Roach M 3rd. Hormonal therapy and radiotherapy for localized prostate cancer: who, where and how long? J Urol 2003;170:S35–S40. discussion S40-31. [PubMed: 14610408]
13. Hanks GE, Pajak TF, Porter A, et al. Phase III trial of long-term adjuvant androgen deprivation after neoadjuvant hormonal cytoreduction and radiotherapy in locally advanced carcinoma of the prostate: the Radiation Therapy Oncology Group Protocol 92-02. J Clin Oncol 2003;21:3972–3978. [PubMed: 14581419]
14. Roach M 3rd, DeSilvio M, Lawton C, et al. Phase III trial comparing whole-pelvic versus prostate-only radiotherapy and neoadjuvant versus adjuvant combined androgen suppression: Radiation Therapy Oncology Group 9413. J Clin Oncol 2003;21:1904–1911. [PubMed: 12743142]
15. Roach M, Marquez C, Yuo H, et al. Predicting the risk of lymph node involvement using the pre-treatment prostate specific antigen and Gleason score in men with clinically localized prostate cancer. International Journal Radiation Oncology Biology Physics 1993;28:33–37.
16. Partin AW, Yoo J, Carter HB, et al. The use of prostate specific antigen, clinical stage and Gleason score to predict pathologic stage in men with localized prostate cancer. J Urol 1993;150:10–14.
17. Smith CT, Williamson PR, Marson AG. Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. Stat Med 2005;24:1307–1319. [PubMed: 15685717]
18. Collett, D. Modelling Survival Data in Medical Research. Great Britain: Chapman & Hall; 1994.
19. Roach M 3rd, Hanks G, Thames H Jr, et al. Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: Recommendations of the RTOG-ASTRO Phoenix Consensus Conference. Int J Radiat Oncol Biol Phys 2006;65:965–974. [PubMed: 16798415]
20. Kaplan EL, Meier P. Non-parameteric estimation from incomplete observations. Journal American Statistical Association 1958;53:457–481.
21. Gray R. A class of K-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat 1988;16:1141–1154.
22. Cox DR. Regression models and life tables. Journal of the Royal Statistical Society 1972;34:187–202.
23. Fine J GR. A proportional hazards model for the subditribution of a competing risk. J Am Statist Assoc 1999;94:496–509.
24. Liverman, CT.; Blazer; Dan, G., editors. Testosterone and Aging: Clinical Research Directions. 1st ed.. Washington, D.C.: The National Academies Press; 2004.
25. Bhasin S, Cunningham GR, Hayes FJ, et al. Testosterone therapy in adult men with androgen deficiency syndromes: an endocrine society clinical practice guideline. J Clin Endocrinol Metab 2006;91:1995–2010. [PubMed: 16720669]
26. Heaton JP. Hormone treatments and preventive strategies in the aging male: whom and when to treat? Rev Urol 2003;5 Suppl 1:S16–S21. [PubMed: 16985938]

27. Shores MM, Matsumoto AM, Sloan KL, et al. Low serum testosterone and mortality in male veterans. Arch Intern Med 2006;166:1660–1665. [PubMed: 16908801]
28. Laughlin GA, Barrett-Connor E, Bergstrom J. Low serum testosterone and mortality in older men. J Clin Endocrinol Metab 2008;93:68–75. [PubMed: 17911176]
29. Bill-Axelson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. N Engl J Med 2005;352:1977–1984. [PubMed: 15888698]
30. Coward RM, Simhan J, Carson Iii CC. Prostate-specific antigen changes and prostate cancer in hypogonadal men treated with testosterone replacement therapy. BJU Int. 2008
31. Morgentaler A. Testosterone Therapy in Men With Prostate Cancer: Scientific and Ethical Considerations. J Urol. 2009
32. Sarosdy MF. Testosterone replacement for hypogonadism after treatment of early prostate cancer with brachytherapy. Cancer 2007;109:536–541. [PubMed: 17183557]
33. Agarwal PK, Oefelein MG. Testosterone replacement therapy after primary treatment for prostate cancer. J Urol 2005;173:533–536. [PubMed: 15643240]
34. Kaufman JM, Graydon RJ. Androgen replacement after curative radical prostatectomy for prostate cancer in hypogonadal men. J Urol 2004;172:920–922. [PubMed: 15310998]
35. Kupelian PA, Potters L, Khuntia D, et al. Radical prostatectomy, external beam radiotherapy <72 Gy, external beam radiotherapy > or =72 Gy, permanent seed implantation, or combined seeds/ external beam radiotherapy for stage T1–T2 prostate cancer. Int J Radiat Oncol Biol Phys 2004;58:25–33. [PubMed: 14697417]
36. Morales A, Black AM, Emerson LE. Testosterone administration to men with testosterone deficiency syndrome after external beam radiotherapy for localized prostate cancer: preliminary observations. BJU Int 2009;103:62–64. [PubMed: 18671790]
37. Taira AV, Merrick GS, Galbreath RW, et al. Pretreatment serum testosterone and androgen deprivation: effect on disease recurrence and overall survival in prostate cancer patients treated with brachytherapy. Int J Radiat Oncol Biol Phys 2009;74:1143–1149. [PubMed: 19117691]

Table 1

Missing Data Analysis Outcomes of Patients With or Without Missing Data (n=2701)

Outcome	Inclusion Status	n	Failures	7-Year Survival/Failure Rate* (95% CI)	Unadjusted HR** (95% CI)	p-value†
OS	With Missing Data	223	121	70.6% (64.0, 76.1)	RL 1.04 (0.86, 1.25)	
	Without Missing Data	2478	1102	69.4% (67.4, 71.2)		0.70
BF	With Missing Data	223	122	49.0% (42.2, 55.6)	RL 0.89 (0.74, 1.07)	
	Without Missing Data	2478	1206	45.8% (43.8, 47.8)		0.21
DM	With Missing Data	223	48	14.1% (9.5, 18.7)	RL 0.78 (0.58, 1.05)	
	Without Missing Data	2478	390	13.4% (12.1, 14.8)		0.11
CSS	With Missing Data	223	30	7.8% (4.2, 11.4)	RL 0.98 (0.67, 1.43)	
	Without Missing Data	2478	279	7.9% (6.8, 9.0)		0.93

*Actuarial estimates for overall survival (OS) were calculated using the Kaplan-Meier method, and the cumulative incidence method was used to estimate biochemical failure (BF), distant metastasis (DM), and cause-specific survival (CSS).

**Hazard ratio: A hazard ratio quantifies how much more or less risk patients at some level have than those at the reference level (RL). A confidence interval that includes 1 indicates no difference between these two subgroups.

†OS : p-value from Chi-square test using the Cox-proportional hazards model; BF, DM, CSS : p-value from log-rank/Gray's test statistics for testing whether the two groups have the same survival distribution or not.

Int J Radiat Oncol Biol Phys. Author manuscript; available in PMC 2011 December 1.

Table 2

Pretreatment Characteristics By Baseline Serum Testosterone (n=2478)

Characteristics	**Very Low Testosterone (n=408)**		**Low Testosterone (n=415)**		**Average Testosterone (n=845)**		**High Testosterone (n=810)**		
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**p-value***
Age (years)									
< 70	186	46	172	41	374	44	377	47	0.38
≥ 70	222	54	243	59	471	56	433	53	
Median (Range)	70 (49–86)		71 (45–87)		70 (46–88)		70 (43–88)		0.59
Mean (standard error)	69.2 (0.3)		69.8 (0.3)		69.6 (0.2)		69.7 (0.2)		
Baseline TL									
Median (Range)	201.7 ng/dl (4.7–248)		282 ng/dl (249–314)		370 ng/dl (314.1–437)		531.5 ng/dl (438–481.5)		
Mean (Standard error)	189.5 ng/dl (2.5)		282.2 ng/dl (0.9)		372.0 ng/dl (1.2)		576.4 ng/dl (8.5)		
Combined Gleason Score†									
2–6	139	34	130	31	295	35	282	35	0.07
7	167	41	166	40	297	35	331	41	
8–10	102	25	119	29	253	30	197	24	
T- Stage									
T1/T2	236	58	226	54	456	54	456	56	0.55
T3/T4	172	42	189	46	389	46	354	44	
PSA									
PSA < 10	93	23	95	23	155	18	160	20	0.42
10 ≤ PSA <20	99	24	103	25	229	27	219	27	
20 ≤ PSA	216	53	217	52	416	55	431	53	
KPS									

Characteristics	Very Low Testosterone (n=408)		Low Testosterone (n=415)		Average Testosterone (n=845)		High Testosterone (n=810)		
	n	%	n	%	n	%	n	%	p-value*
60–90	216	53	213	51	478	57	400	49	0.03
100	192	47	202	49	367	43	410	51	
Arm									
9202 Arm 1	112	27	101	24	226	27	198	24	0.04
9202 Arm 2	86	21	100	24	251	30	229	28	
9413 Arm 1	46	11	60	14	101	12	90	11	
9413 Arm 2	53	13	54	13	84	10	96	12	
9413 Arm 3	49	12	57	14	93	11	92	11	
9413 Arm 4	62	15	43	10	90	11	105	13	

*p-value for continuous variables is from analysis of variance; for categorical variables it is from Chi-square test statistics.

Int J Radiat Oncol Biol Phys. Author manuscript; available in PMC 2011 December 1.

Table 3

Heterogeneity Testing: Proportional Hazards Regression Model By Baseline Serum Testosterone (BST) group

Stratification Variable: Arm	**Unadjusted**		**Adjusted****	
BST Group				
Overall Survival	Chi-Square T.S. (Q)	p-value	Chi-Square T.S. (Q)	p-value
VLBST LBST ABST HBST	RL 7.9 6.6 3.6	- 0.84 0.75 0.40	RL 6.03 6.49 3.27	- 0.70 0.74 0.34
BST Group	Unadjusted Pooled HR† (95% CI*)		Adjusted Pooled HR† (95% CI*)	
VLBST LBST ABST HBST	RL 1.06 (0.86, 1.30) 1.04 (0.87, 1.25) 1.03 (0.86, 1.23)		RL 1.03 (0.84, 1.27) 1.02 (0.85, 1.22) 1.07 (0.89 1.28)	
Distant Metastasis	Chi-Square T.S. (Q)	p-value	Chi-Square T.S. (Q)	p-value
VLBST LBST ABST HBST	RL 5.2 1.9 2.4	- 0.61 0.13 0.21	RL 5.2 1.6 2.3	- 0.61 0.10 0.19
BST Group	Unadjusted Pooled HR† (95% CI*)		Adjusted Pooled HR† (95% CI*)	
VLBST LBST ABST HBST	RL 1.01 (0.69, 1.47) 1.12 (0.82, 1.53) 1.25 (0.92, 1.70)		RL 0.96 (0.66, 1.40) 1.03 (0.75, 1.41) 1.23 (0.90, 1.66)	
Biochemical Failure	Chi-Square T.S. (Q)	p-value	Chi-Square T.S. (Q)	p-value
VLBST LBST ABST HBST	RL 4.8 2.4 2.5	- 0.56 0.21 0.23	RL 7.0 3.9 3.3	- 0.78 0.44 0.35
BST Group	Unadjusted Pooled HR† (95% CI*)		Adjusted Pooled HR† (95% CI*)	
VLBST LBST ABST HBST	RL 1.06 (0.87, 1.29) 1.06 (0.89, 1.26) 1.14 (0.96, 1.36)		RL 1.06 (0.87, 1.31) 1.01 (0.85, 1.21) 1.15 (0.96, 1.36)	
Cause-specific Survival††	N/A		N/A	

*CI = Confidence Interval; RL = Reference Level.

**Adjusted for age (<70 vs. > 70), Gleason (2–6 vs. 7 vs. 8–10), PSA (<10 vs. ≥10 and <20 vs. ≥ 20), and KPS (60–90 vs. 100),

†This is a pooled estimate

††This can not be calculated because the number of failures in RTOG 9413 is too few.

Int J Radiat Oncol Biol Phys. Author manuscript; available in PMC 2011 December 1.

Table 4

Univariate Proportional Hazards Models by Baseline Serum Testosterone (BST) (n=2478)

Endpoint	Baseline Serum Testosterone	n	Failures	7-Year Survival/Failure Rate* (95% CI)	Unadjusted HR** (95% CI)	p-value†
Overall Survival	VLBST	408	176	69.0% (64.0, 73.4)	RL	
	LBST	415	180	67.8% (62.9, 72.2)	1.04 (0.84, 1.27)	0.75
	ABST	845	387	68.9% (65.5, 72.0)	1.05 (0.88, 1.26)	0.58
	HBST	810	359	70.9% (67.5, 74.0)	1.02 (0.85, 1.22)	0.81
		p-value from log-rank test statistic = 0.95				
Distant Metastasis	VLBST	408	58	12.3% (9.1, 15.6)	RL	
	LBST	415	56	11.4% (8.2, 14.5)	0.95 (0.66, 1.38)	0.80
	ABST	845	135	13.4% (11.0, 15.7)	1.13 (0.83, 1.54)	0.44
	HBST	810	141	15.1% (12.6, 17.6)	1.25 (0.92, 1.69)	0.15
		p-value from Fine Gray's test statistic = 0.29				
Biochemical Failure	VLBST	408	190	44.8% (39.8, 49.7)	RL	
	LBST	415	202	45.5% (40.6, 50.5)	1.05 (0.86, 1.28)	0.61
	ABST	845	403	45.8% (42.3, 49.2)	1.03 (0.86, 1.22)	0.77
	HBST	810	411	46.6% (43.1, 50.1)	1.11 (0.94, 1.32)	0.22
		p-value from Fine Gray's test statistic = 0.57				
Cause-specific Survival	VLBST	408	47	8.4% (5.6, 11.2)	RL	
	LBST	415	44	7.7% (5.0, 10.3)	0.93 (0.62, 1.40)	0.72
	ABST	845	100	8.9% (6.9, 10.8)	1.01 (0.72, 1.43)	0.95
	HBST	810	88	6.9% (5.1, 8.7)	0.93 (0.66, 1.33)	0.71

Endpoint	Baseline Serum Testosterone	n	Failures	7-Year Survival/Failure Rate* (95% CI)	Unadjusted HR** (95% CI)	p-value†
		p-value from Fine Gray's test statistic = 0.94				

*The Kaplan-Meier method was used to estimate overall survival, and the cumulative incidence method was used to estimate local failure, distant metastasis, and cause-specific survival.

**Hazard ratio: A hazard ratio quantifies how much more (less) risk patients at some level have than those at the reference level (RL). A confidence interval that includes 1 indicates no difference between these two subgroups.

†For overall survival, the p-value is from the Chi-square test using the Cox-proportional hazards model; for all other endpoints, the p-value is from log-rank/Gray's test statistics for testing whether the two groups have the same survival distribution or not.

Table 5

a Multivariate Proportional Hazards Model Overall Survival Categorized Testosterone Level (n=2478)

Covariate	Comparison	Adjusted HR* (95% CI)	p-value†
Baseline Serum Testosterone group	VLBST	RL	
	LBST	1.02 (0.83, 1.25)	0.86
	ABST	1.03 (0.86, 1.23)	0.78
	HBST	1.06 (0.88, 1.27)	0.56
Age	< 70	RL	
	≥ 70	1.46 (1.29, 1.65)	<0.0001
Gleason	2–6	RL	
	7	1.18 (1.02, 1.37)	0.02
	8–10	1.55 (1.34, 1.81)	<0.0001
PSA	PSA < 10	RL	
	10 ≤ PSA < 20	1.02 (0.85, 1.22)	0.83
	PSA ≥ 20	1.16 (0.99, 1.35)	0.07
KPS	60–90	RL	
	100	0.72 (0.64, 0.82)	<0.0001
Study / Treatment Arm	9202 Arm 1	RL	
	9202 Arm 2	0.89 (0.77, 1.04)	0.14
	9413 Arm 1	0.93 (0.74, 1.17)	0.54
	9413 Arm 2	0.92 (0.72, 1.16)	0.47
	9413 Arm 3	1.22 (0.98, 1.52)	0.07
	9413 Arm 4	0.99 (0.79, 1.24)	0.92

b Multivariate Proportional Hazards Model Distant Metastasis Categorized Testosterone Level (n=2478)

Covariate	Comparison	Adjusted HR* (95% CI)	p-value†
BST group	VLBST	RL	
	LBST	0.94 (0.65, 1.35)	0.73
	ABST	1.08 (0.79, 1.47)	0.62
	HBST	1.25 (0.92, 1.69)	0.15
Age	< 70	RL	
	≥ 70	0.62 (0.51, 0.76)	<0.0001
Gleason	2–6	RL	
	7	1.48 (1.14, 1.94)	0.004
	8–10	2.67 (2.06, 3.46)	<0.0001
PSA	PSA < 10	RL	
	10 ≤ PSA < 20	1.09 (0.79, 1.51)	0.59
	PSA ≥ 20	1.44 (1.09, 1.90)	0.01

b Multivariate Proportional Hazards Model Distant Metastasis Categorized Testosterone Level (n=2478)

Covariate	Comparison	Adjusted HR* (95% CI)	p-value†
KPS	60–90	RL	
	100	1.01 (0.83, 1.23)	0.93
Study / Treatment Arm	9202 Arm 1	RL	
	9202 Arm 2	0.64 (0.50, 0.83)	0.0007
	9413 Arm 1	0.52 (0.36, 0.76)	0.0008
	9413 Arm 2	0.62 (0.43, 0.90)	0.01
	9413 Arm 3	0.72 (0.51, 1.02)	0.06
	9413 Arm 4	0.56 (0.38, 0.81)	0.002

c Multivariate Proportional Hazards Model Biochemical Failure Categorized Testosterone Level (n=2478)

Covariate	Comparison	Adjusted HR* (95% CI)	p-value†
BST group	VLBST	RL	
	LBST	1.05 (0.86, 1.28)	0.63
	ABST	1.02 (0.85, 1.21)	0.84
	HBST	1.13 (0.95, 1.35)	0.15
Age	< 70	RL	
	≥ 70	0.69 (0.61, 0.77)	<0.0001
Gleason	2–6	RL	
	7	1.28 (1.11, 1.46)	0.0004
	8–10	1.60 (1.38, 1.86)	<0.0001
PSA	PSA < 10	RL	
	10 ≤ PSA < 20	1.45 (1.20, 1.76)	0.0001
	PSA ≥ 20	2.02 (1.71, 2.39)	<0.0001
KPS	60–90	RL	
	100	1.07 (0.96, 1.20)	0.24
Study / Treatment Arm	9202 Arm 1	RL	
	9202 Arm 2	0.58 (0.50, 0.68)	<0.0001
	9413 Arm 1	0.70 (0.57, 0.86)	0.0005
	9413 Arm 2	0.88 (0.72, 1.06)	0.18
	9213 Arm 3	0.77 (0.63, 0.94)	0.009
	9213 Arm 4	0.67 (0.55, 0.83)	0.0002

d Multivariate Proportional Hazards Model Cause-Specific Survival Categorized Testosterone Level (n=2478)

Covariate	Comparison	Adjusted HR* (95% CI)	p-value†
BST group	VLBST	RL	
	LBST	0.89 (0.59, 1.34)	0.57
	ABST	0.94 (0.66, 1.32)	0.71
	HBST	0.91 (0.64, 1.29)	0.60

d Multivariate Proportional Hazards Model Cause-Specific Survival Categorized Testosterone Level (n=2478)

Covariate	Comparison	Adjusted HR* (95% CI)	p-value†
Age	< 70	RL	
	> 70	0.66 (0.52, 0.83)	0.0005
Gleason	2–6	RL	
	7	1.39 (1.00, 1.92)	0.048
	8–10	3.05 (2.25, 4.13)	<0.0001
PSA	PSA < 10	RL	
	10 < PSA < 20	0.97 (0.66, 1.45)	0.90
	PSA > 20	1.50 (1.08, 2.10)	0.02
KPS	60–90	RL	
	100	0.94 (0.74, 1.19)	0.62
Study / Treatment Arm	9202 Arm 1	RL	
	9202 Arm 2	0.74 (0.56, 0.97)	0.03
	9413 Arm 1	0.35 (0.21, 0.60)	0.0002
	9413 Arm 2	0.63 (0.40, 0.99)	0.04
	9213 Arm 3	0.57 (0.36, 0.90)	0.02
	9213 Arm 4	0.44 (0.27, 0.72)	0.001

*Hazard ratio: A hazard ratio quantifies how much more (less) risk patients at some level have than those at the reference level (RL). A confidence interval that includes 1 indicates no difference between these two subgroups. HR adjusted for treatment arm (9202 Arm 1 [reference level; RL] vs. 9202 Arm 2 vs. 9413 Arm 1 vs. 9413 Arm 2 vs. 9413 Arm 3 vs. 9413 Arm 4), age (<70 vs. > 70), Gleason (2–6 vs. 7 vs. 8–10), PSA (<10 vs. ≥10 and <20 vs. ≥ 20), and KPS (60–90 vs. 100).

†The p-value is from the Chi-square test using the Cox-proportional hazards model. Statistically significant is at the significance level of 0.05.

*Hazard ratio: A hazard ratio quantifies how much more (less) risk patients at some level have than those at the reference level (RL). A confidence interval that includes 1 indicates no difference between these two subgroups. HR adjusted for treatment arm (9202 Arm 1 [reference level; RL] vs. 9202 Arm 2 vs. 9413 Arm 1 vs. 9413 Arm 2 vs. 9413 Arm 3 vs. 9413 Arm 4), age (<70 vs. > 70), Gleason (2–6 vs. 7 vs. 8–10), PSA (<10 vs. ≥10 and <20 vs. ≥ 20), and KPS (60–90 vs. 100).

†The p-value is from the log-rank/Gray's test statistics for testing whether the two groups have the same survival distribution or not. Statistically significant at the significance level of 0.05.

*Hazard ratio: A hazard ratio quantifies how much more (less) risk patients at some level have than those at the reference level (RL). A confidence interval that includes 1 indicates no difference between these two subgroups. HR adjusted for treatment arm (9202 Arm 1 [reference level; RL] vs. 9202 Arm 2 vs. 9413 Arm 1 vs. 9413 Arm 2 vs. 9413 Arm 3 vs. 9413 Arm 4), age (<70 vs. > 70), Gleason (2–6 vs. 7 vs. 8–10), PSA (<10 vs. ≥10 and <20 vs. ≥ 20), and KPS (60–90 vs. 100).

†The p-value is from the log-rank/Gray's test statistics for testing whether the two groups have the same survival distribution or not. Statistically significant at the significance level of 0.05.

*Hazard ratio: A hazard ratio quantifies how much more (less) risk patients at some level have than those at the reference level (RL). A confidence interval that includes 1 indicates no difference between these two subgroups. HR adjusted for treatment arm (9202 Arm 1 [reference level; RL] vs. 9202 Arm 2 vs. 9413 Arm 1 vs. 9413 Arm 2 vs. 9413 Arm 3 vs. 9413 Arm 4), age (<70 vs. > 70), Gleason (2–6 vs. 7 vs. 8–10), PSA (<10 vs. ≥10 and <20 vs. ≥ 20), and KPS (60–90 vs. 100).

†The p-value is from the log-rank/Gray's test statistics for testing whether the two groups have the same survival distribution or not. Statistically significant at the significance level of 0.05.