Title

Impact of biochemical failure classification on clinical outcome: A secondary analysis of Radiation Therapy Oncology Group 9202 and 9413

Permalink

Journal

ISSN

Authors

Hamstra, Daniel A
Bae, Kyounghwa
Hanks, Gerald
et al.

Publication Date

DOI

Peer reviewed

Impact of Biochemical Failure Classification on Clinical Outcome: A Secondary Analysis of Radiation Therapy Oncology Group 9202 and 9413

Daniel A. Hamstra, MD, PhD[1], **Kyounghwa Bae, PhD**[2], **Gerald Hanks, MD**[3], **Chen Hu, PhD**[2], **William U. Shipley, MD**[4], **Charlie C. Pan, MD, MS**[5], **Mack Roach III, MD**[6], **Colleen A. Lawton, MD**[7], and **Howard M. Sandler, MD**[8]

[1]Department of Radiation Oncology, University of Michigan Medical Center, Ann Arbor, Michigan

[2]Radiation Therapy Oncology Group Statistical Center, Philadelphia, Pennsylvania

[3]Fox Chase Cancer Center, Philadelphia, Pennsylvania

[4]Massachusetts General Hospital, Boston, Massachusetts

[5]Bellin Cancer Center, Green Bay, Wisconsin

[6]University of California San Francisco, San Francisco, California

[7]Medical College of Wisconsin, Milwaukee, Wisconsin

[8]Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California

Abstract

BACKGROUND—Biochemical failure (BF) after radiation therapy is defined on the basis of a rising prostate-specific antigen (PSA) level (A1 failure) or any event that prompts the initiation of salvage androgen-deprivation therapy without PSA failure (A2). It was hypothesized that A2 failure may have a different prognosis.

METHODS—Data for 2799 eligible patients from Radiation Therapy Oncology Group (RTOG) 9202 and RTOG 9413 were analyzed. BF was defined according to the 1997 American Society for Therapeutic Radiology and Oncology consensus definition as A1 for PSA failure or as A2 for the start of salvage hormone therapy before 3 consecutive PSA rises.

RESULTS—Rates of all-cause mortality (hazard ratio [HR], 1.7; 95% confidence interval [CI], 1.5-2.0; *P* <.0001) and distant metastasis (DM; HR, 1.6; 95% CI, 1.3-2.0; *P* <.0001) were greater

with A2 failure. The 5-year overall survival (OS) rates were 88.2% and 74.6% for A1 and A2, respectively ($P <.0001$), and the DM rates were 15.7% and 29.0%, respectively ($P <.0001$). The DM rate was greater at 5 years for A2 patients with DM as the first sign of failure versus patients with other A2 failures (87.3% vs 11.7%, $P <.001$), and this also correlated with worse OS at 5 years: 81.1% for A2 failure without DM and 52.8% with DM ($P <.001$). After the removal of patients with DM, the difference between A1 and A2 BF persisted for OS ($P =.002$) but not for DM ($P =.16$)

CONCLUSIONS—These results suggest that patients with rising PSA levels alone have less risk than those with A2 failures; although DM was the largest contributor of adverse risk to A2 failure, it did not account for all excess risk in A2 failure.

Keywords

prostate-specific antigen; prostate-specific antigen failure; radiotherapy

INTRODUCTION

Because of the long natural history of prostate cancer, for many studies, the use of prostate-specific antigen (PSA) to define biochemical failure (BF) has replaced the use of clinical failure. Because of the multiple different definitions of BF in use, in 1997, the American Society for Therapeutic Radiology and Oncology (ASTRO) convened a panel that established a consensus definition of PSA failure, and this definition was in use for 10 years.[1] This consensus definition initially described failure as 3 consecutive rises in posttreatment PSA after the achievement of a nadir, with the date of failure recorded as the time midway between the nadir and the first rising PSA level (ASTRO1 or A1). Subsequently, a second type of failure, defined as the initiation of salvage androgen-deprivation therapy (ADT) for any reason other than 3 consecutive rises (ASTRO2 or A2), was added to capture what would otherwise be nonevaluable patients. A second ASTRO consensus conference subsequently described a new "phoenix" definition for BF and amended the original PSA-based definition (A1) to a rise of 2 ng/mL or more above the nadir PSA level, with the date of failure at call.[2] In addition, the documentation of local or distant failure or the initiation of salvage ADT before the nadir + 2 ng/mL value was also included in the phoenix definition (A2). However, the prognostic significance of failure as indicated by salvage ADT without an express cutoff for rising PSA being met, whether it is based on the old or new ASTRO definitions, has not been well explored.

METHODS AND MATERIALS

Patients, Treatment, and Follow-Up

As part of an institutional review board approved secondary analysis data from 2799 eligible patients from Radiation Therapy Oncology Group (RTOG) 9202 and RTOG 9413 were analyzed. The details regarding patient selection and treatment have been described previously.[3,4] In RTOG 9202, patients had locally advanced prostate cancer (T2c-T4) and pretreatment PSA levels < 150 ng/mL. All patients received external-beam radiation therapy (RT) to the whole pelvis, and this was followed by a boost to the prostate to a total dose of

67.5 to 70 Gy. Patients were randomly assigned to receive short-term ADT for 4 months starting 2 months before RT or to the same regimen with an additional 2 years of ADT (long-term ADT). In RTOG 9413, patients had localized prostate cancer with a presumed risk of lymph node involvement   15% on the basis of the Roach formula.[5] Patients were randomized into 4 groups according to whether they started to receive short-term ADT before or after RT and whether the prostate only or the whole pelvis was treated.

Follow-up was scheduled every 3 to 4 months during years 1 and 2, every 6 months during years 3 to 5, and annually thereafter. During follow-up, any patient presenting with bone pain not attributable to any intercur-rent disease underwent a bone scan. The protocol did not require a bone scan in asymptomatic patients with elevated PSA levels.

Study Endpoints

BF according to the original ASTRO consensus definition was categorized into 1 of 2 types: 3 consecutive rises in posttreatment PSA levels after the achievement of a nadir (A1) or the initiation of salvage hormone therapy for any reason other than A1 failure after the end of RT (A2). A PSA rise was defined as a minimum of 0.2 ng/mL at a minimum of 3 months after the end of RT. The date of failure was defined as the halfway point between the nadir date and the first rise for A1 and the start date of salvage hormone therapy for any reason other than A1 (before the 3 consecutive rises) for A2. The PSA nadir was defined as the lowest PSA level achieved after the completion of RT. The time to BF was measured from the randomization date to the date of failure. Overall survival (OS) was defined as death by any cause. Local failure was defined as tumor recurrence (positive rebiopsy at least 2 years after study entry or tumor regrowth of 50%) or a tumor never cleared by digital rectal examination. Distant metastasis (DM) was defined as the documentation of clinical evidence of distant disease.

Statistical Methods

The chi-square test was used to compare pretreatment characteristics, whereas the Kaplan-Meier method[6] and the log-rank test were used to estimate the rates for OS.[7] The cumulative incidence method[8] was used to estimate the time to local failure and DM, with Gray's test[9] used to compare the cumulative incidence rates over time between treatment arms. For multivariate analyses, Cox proportional hazard regression analysis[10] was used for OS, and Fine and Gray's regression analysis[11] was used for local failure and DM. The proportional hazards assumptions were assessed with both graphical approaches and goodness-of-fit tests.[12] The following categorical covariates were considered for the combined data of RTOG 9202 and RTOG 9413: age (<70 [reference level] vs   70 years), combined Gleason score (2-6 [reference level] vs 7-10), PSA (30 [reference level] vs >30 ng/mL), clinical stage (T2 [reference level] vs >T2), and study (RTOG 9202 [reference level] vs RTOG 9413). In addition to these covariates, the PSA level at its nadir after RT (continuous), the time to the PSA nadir from randomization (continuous), and the PSA change ([PSA at failure – PSA at nadir]/[date of failure – date of PSA nadir]; continuous) were considered as covariates for each study and in the combined data. All statistical comparisons were 2-sided, and a *P* value < .05 was considered statistically significant. SAS software (SAS Institute, Cary, NC) and R software were used for all analyses.

RESULTS

Pretreatment Patient Characteristics

The median follow-up times were 9.0 and 6.5 years for RTOG 9202 and RTOG 9413, respectively. The pre-treatment characteristics for these studies have been previously described and are summarized in Table 1 for patients with BF categorized as A1 or A2 failure. There were no differences in age, PSA, T classification, Gleason score, or lymph node status between those with A1 failure and those with A2 failure (all $P > .05$).

Type of BF

From both studies, there were 1181 BF events according to the ASTRO consensus definition (663 of 1521 for RTOG 9202 and 518 of 1278 for RTOG 9413). Overall, 42% of the patients experienced BF, and among the patients who experienced BF, 56% (664 of 1181) were diagnosed according to 3 rises in PSA (A1), whereas a substantial minority (44% [517 of 1181]) experienced A2 failure (47% [311 of 663] for RTOG 9202 and 40% [206 of 518] for RTOG 9413). Salvage ADT was given to 34% (951 of 2799) of all patients from the 2 studies (36% [553 of 1521] in RTOG 9202 and 31% [398 of 1278)] in RTOG 9413); this rate was higher for patients with BF defined as A2 (100% [517 of 517]) versus patients with BF defined by rising PSA alone (A1; 65% [434 of 664]).

Survival Outcomes

At 5 years, the metastasis rate was greater for patients with A2 failure versus those with A1 failure (29.0% vs 15.7%; hazard ratio [HR], 1.60; 95% confidence interval [CI], 1.32-1.95; $P < .0001$; Fig. 1A). Among patients with A2 failure, those with DM before or within 1 month of the initiation of ADT had substantially greater DM at 5 years in comparison with those with all other A2 failures (87.3% vs 11.7%, $P < .001$; Fig. 1B and Table 2), whereas there was no statistical difference in DM between those with A1 failure and those with A2 failure without initial DM ($P = .15$). OS at 5 years was also lower for those with A2 failure (88.2% vs 74.6%; HR, 1.68; 95% CI, 1.48-1.99; $P < .0001$; Fig. 1C), and this again was worst for those with initial DM (52.8%) versus those with other A2 failures (81.1%, $P < .001$; Fig. 1D). However, A2 failure without initial DM was still associated with worse OS in comparison with A1 failure (5-year rate: 88.2% vs 81.1%, $P = .0002$). Local failure was not different between BF types (19.6% vs 21.3%; HR, 1.01; 95% CI, 0.81-1.27; $P = .92$) or by type of A2 failure. The impact of A2 failure was similar in RTOG 9413 and RTOG 9202 (Table 3).

Because of the heterogeneity identified in patients with A2 failure, men with A2 failure were divided into 4 groups: those with a rising PSA level after RT without decreasing PSA (pattern 1, n = 54); those who had an initially decreasing PSA level, achieved a PSA nadir, and subsequently experienced an increase but did not meet the definition for A1 failure (pattern 2, n = 284); those who had irregular PSA patterns other than the previous 2 patterns (pattern 3, n = 61); and finally those who had DM as the first event or within 1 month of the initiation of ADT (pattern 4, n = 118; Table 2). Those with DM as the first sign of failure had the highest rate of DM at 5 years (pattern 4, 87.3%), and those with a rising PSA level after the completion of RT also had a higher rate of DM at 5 years (pattern 1, 37.3%) in

comparison with those with rising PSA levels who did not meet the definition of A1 failure (pattern 2, 7.9%) and those with another PSA pattern (pattern 3, 6.7%). The 5-year OS rate was also worse for those with DM as the first sign of failure (52.8%) and those with initially rising PSA after RT (51.0%) in comparison with those with rising PSA (90.3%) or another pattern (63.9%).

Stepwise multivariate regression models (Table 5) identified an age 70 years ($P < .00001$), a Gleason score of 7 to 10 ($P < .0001$), and A2 BF ($P < .0001$) as associated with an increased risk of all-cause mortality. Similarly, a Gleason score of 7 to 10 ($P = .0012$), a T classification > T2 ($P < .0001$), treatment in RTOG 9202 ($P = .00065$), and A2 BF ($P < .0001$) all predicted higher rates of metastasis. However, only a greater T classification ($P = .0016$) and treatment in RTOG 9202 ($P = .0024$) predicted greater local failure, and neither the Gleason score ($P > .05$) nor A2 BF ($P = .78$) predicted increased local failure.

PSA Kinetics in the Patients Who Failed by BF: A1 Type

PSA-related variables—PSA nadir after RT, time to the PSA nadir after RT, and rate of the PSA rise after its nadir—were each analyzed to assess their effects on the time to BF for those with A1 BF (Table 4). The multivariate analysis controlled for the following: age, PSA at the baseline, T classification, Gleason score, and treatment in RTOG 9202 or RTOG 9413. Among men who had conventional BF due to rising PSA (as defined by A1), a longer time to the PSA nadir in months ($P < .00001$; HR, 0.99; 95% CI, 0.99-0.995) was associated with an improved outcome, whereas a faster rate of the PSA rise after its nadir ($P < .0001$; HR, 1.19; 95% CI, 1.12-1.27) was associated with an earlier time to A1 BF. The PSA nadir itself, however, was not a predictor of the time to A1 BF across all patients or in the subset of patients treated in RTOG 9202, but as previously documented, the PSA nadir was prognostic for those treated in RTOG 9413 (data not shown).[13,14]

PSA Kinetics and Outcomes in Patients Who Failed by BF: A2 Type

Finally, the influence of PSA kinetics was assessed independently in those who had A2 BF (Table 2). The median PSA at the time of A2 BF was higher for those with an initially rising PSA level after RT (pattern 1, 10.9 ng/mL) and those with DM (pattern 4, 11.2 ng/mL) versus those with patterns 2 and 3 (2.2 and 1.9 ng/mL, respectively). In addition, the PSA doubling times were also shorter for those with an initially rising PSA level (7.8 vs 22.6, 18.7, and 11.4 months for patterns 2, 3, and 4, respectively).

Salvage Therapy

Finally, the use of salvage therapy could potentially have influenced the clinical outcomes because all patients with A2 failure by definition started ADT, whereas it might have been delayed for those with A1 failure. The median time after BF to salvage ADT in all A1 failures was 23.8 months, with 50.5% starting salvage ADT within the first 2 years after A1 BF. The rate of salvage ADT use did not differ when it was based on the variables prognostic for outcomes in the multivariate (MTV) analysis (Table 5), including the Gleason score (2-6 vs 7-10, $P = .32$), T classification (T2 vs T3/T4, $P = .54$), and age (<70 vs 70 years, $P = .34$), although those treated in RTOG 9413 did have a shorter median time to

salvage ADT (21.4 vs 26.3 months in RTOG 9202; HR, 1.40; 95% CI, 1.15-1.70; *P* = .0007).

DISCUSSION

Prostate cancer patients treated in RTOG 9202 and RTOG 9413 who were classified with BF due to the initiation of salvage ADT for any reason other than the study definition of PSA failure with 3 consecutive rises in posttreatment PSA levels after the achievement of a nadir appeared to have worse OS and DM rates than those who had PSA rise–defined failure. The majority of this increased risk was captured by those with DM as the first sign of failure (Fig. 1C), although those with A2 failure without initial DM did still have worse OS than those with rising PSA alone (Fig. 1D). The corollary of this is that BF documented by a rising PSA level alone (A1 failure), which occurred in 56% of the patients with BF, is less risky than one might assume from an analysis of outcomes for all patients with BF that does not account for A1, A2, or DM as the first sign of failure. The apparent lower risk with rising PSA alone as the cause of BF reinforces the growing concept that the early initiation of salvage ADT may not provide clinical benefit.[15] Nevertheless, because salvage ADT use was very prevalent in these patients after any BF, it is difficult to sort out the relative clinical benefit of salvage ADT.

This is the first report of the significance of failure according to the non–PSA rise definition, and failure was observed across 2 different clinical trials with significantly different ADT treatment schedules (4 vs 28 months). In addition, this finding was seen in both univariate and multivariate analyses, and this suggests that patients whose failure is due only to a rising PSA level may be a distinctly better performing group. Nevertheless, in the current analysis, BF was defined on the basis of 3 consecutive rises (the 1997 ASTRO consensus definition), and this was similar to the standard when the 2 phase 3 trials in question were initiated and to how patients were frequently managed after treatment. The more modern phoenix definition (PSA nadir + 2 ng/mL) was not used clinically in these patients and was not analyzed herein.[2] However, because many patients were started on ADT before they reached the phoenix definition threshold (the median PSA level was 2.2 ng/mL for those with rising PSA but not 3 rises; Table 2), the current analysis, if performed with the phoenix definition, would be more problematic in that more patients would be classified as failing for causes other than a rise of 2 ng/mL.

Because the definition of 3 consecutive rises is known to be more susceptible to PSA bounces or benign rises in PSA after the cessation of ADT, that is one factor that might underlie the more favorable outcomes for those whose failure was determined by 3 consecutive rises.[16] Nevertheless, if patients were diagnosed with BF on the basis of 3 consecutive rises in the setting of false-positive BF, it also seems likely that patients would also have high false-positive rates if they had rising PSA levels but had not reached 3 consecutive rises and were started on ADT as part of the A2 group, and indeed this group (pattern 2; Table 2) appeared to have very favorable clinical outcomes, with a 90.3% 5-year survival rate and only a 7.9% DM rate. Interestingly, for those with A1 BF who had a rising PSA level after a nadir, the absolute nadir achieved was not prognostic for outcomes, whereas the rate of the PSA rise after a nadir was prognostic. Some have suggested that that

the absolute PSA nadir after RT and ADT is prognostic for those treated with short-term ADT[13,14,17]; however, it is possible that for those treated with long-term ADT, the influence of the absolute PSA nadir is smaller.[18] Furthermore, in those with an A1 failure, the time to achieve a nadir was also prognostic, with a longer time to a nadir being more favorable; however, this variable is susceptible to an immortality bias because failure cannot be called until a nadir has been reached, and this artificially biases results without adjustments such as landmark or time-dependent analyses.[17]

This analysis also suffers from the fact that the A2 group was rather heterogeneous (Table 2). The reasons that a patient could be classified as A2 included a continually rising PSA level after treatment, a rise in PSA (that did not meet the definition of 3 rises), the appearance of symptomatic metastasis, local progression on clinical examination, and the initiation of salvage ADT for an unspecified reason. Although this group was heterogeneous, in the multivariate analysis, which included significant covariates for treatment failure (age, PSA, T classification, Gleason score, and ADT duration), it was still found that the patients with A2 failure had worse OS and DM rates than those with 3 consecutive rises. Overall, this lends credence to the adverse impact of A2 failure despite its heterogeneous nature and the better outcomes for those with rising PSA alone.

Patients from this analysis were treated exclusively with external-beam RT, and the results may not be directly applicable to patients treated with either prostate brachytherapy or prostatectomy. The use of PSA for the determination of BF after brachytherapy, however, appears to correlate, as with external-beam therapy, with ultimate clinical failure,[19] so one might expect that non–PSA-rising brachytherapy failures may act similarly to non–PSA-rising external-beam failures.[20] In addition, because the prognostic benefit of A2 failure appears primarily for distant recurrences, the local treatment modality may not be significant. Although this study did identify similar local failure rates in the A1 and A2 failure groups, the evaluation of local failure was limited because rebiopsy was not required or performed in the vast majority of these patients. Nevertheless, one would assume that this lack of evaluation would have been relatively equally distributed throughout the entire study population; therefore, although local failure rates may be underestimated, the relative impact of local failure should not be different on the basis of the definition of BF.

The PSA pattern most commonly seen after A2 failure was one of PSA decreasing to a nadir after treatment and then rising (approximately 55% of all A2 patients; Table 2). In this group, the median PSA doubling time was 22.6 months (mean, 98.9 months). Because of the wide range of PSA doubling times (3.0-2718 months), it appears that the PSA doubling time is not useful in the A2 population in contrast to A1 patients, as seen in other studies.[21,22] The median PSA nadir for this group was 0.2 ng/mL (mean, 0.43 ng/mL), and although others have demonstrated that patients with a PSA nadir less than 0.8 or 0.2 ng/mL may have a good prognosis,[13,14,17,18] this does not seem to be the case for A2 failures, for which a PSA nadir was not prognostic. Additional work regarding prognostic factors in the A2 population is warranted, although it appears that patients with DM as the first failure (pattern 4) and patients with a rising PSA level after RT without a nadir (pattern 1) were the 2 groups with the greatest risk of overall mortality, and they accounted for 33.2% of all A2 failures. In a previous analysis of RTOG 8610, those who started salvage ADT after the development

of DM had substantially worse survival than those who started salvage ADT without metastasis (median OS after salvage ADT: 4.9 vs 2.8 years).[23] In comparison, for those without metastasis at the time of the initiation of salvage ADT, the difference in median survival as a function of PSA was much smaller: 5.3 years for those with a PSA level < 20 ng/mL at the time of salvage ADT and 4.3 years for those for whom salvage ADT was started with a PSA level of 20 ng/mL or higher. Nevertheless, differences in unmeasured confounders between these groups (eg, PSA kinetics at the time of salvage ADT) may account for the observed differences in outcomes and may not necessarily support an improved outcome with early salvage ADT for those with only biochemical recurrence of prostate cancer. In addition to metastasis, which could have influenced A2 failures and led to worse clinical outcomes, early BF (with a short interval to BF) has also been identified as carrying a significantly increased risk of death from prostate cancer[24,25] or death overall.[26] and it may have been captured in the A2 definition, particularly for those men who had continually rising PSA after RT or recurrence even during ADT.

In conclusion, patients deemed to have A2 BF may have biologically more aggressive disease with worse OS and DM rates, and this warrants additional study. As a result, those with A1 failure by virtue of rising PSA alone may carry less clinical risk. Further investigation of this phenomenon in other clinical trials and with the current definition of BF is warranted.

Acknowledgments

REFERENCES

1. Consensus statement: guidelines for PSA following radiation therapy. American Society for Therapeutic Radiology and Oncology Consensus Panel. Int J Radiat Oncol Biol Phys. 1997; 37:1035–1041. [PubMed: 9169810]

2. Roach M III, 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]

3. Roach M III, 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]

4. 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]

5. Roach M III, Marquez C, Yuo HS, 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. Int J Radiat Oncol Biol Phys. 1994; 28:33–37. [PubMed: 7505775]

6. Kaplan E, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958; 53:457–481.

7. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother Rep. 1966; 50:163–170. [PubMed: 5910392]

8. Kalbfleisch, J.; Prentice, R. The Statistical Analysis of Failure Time Data. John Wiley & Sons; New York, NY: 1980.

9. Gray R. A class of K-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat. 1988; 16:1141–1154.

10. Cox D. Regression models and life tables. J R Stat Soc Ser B. 1972; 34:187–229.

11. Fine J, Gray R. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 1999; 94:496–509.

12. Kleinbaum, DG.; Klein, M. Survival Analysis: A Self-Learning Text. 3rd ed.. Springer-Verlag; New York, NY: 2012.

13. Almeras C, Zerbib M, Eschwege F, Debre B. Nadir PSA and kinetics of PSA decline between the 3rd and 6th month after external beam radiotherapy for T1 T2 Nx M0 localized prostate cancer: value of the prediction of the risk of biological progression [in French]. Prog Urol. 2002; 12:219–225. [PubMed: 12108335]

14. Cury FL, Hunt D, Roach M III, et al. Prostate-specific antigen response after short-term hormone therapy plus external-beam radio-therapy and outcome in patients treated on Radiation Therapy Oncology Group study 9413. Cancer. 2013; 119:1999–2004. [PubMed: 23504930]

15. Loblaw DA, Virgo KS, Nam R, et al. Initial hormonal management of androgen-sensitive metastatic, recurrent, or progressive prostate cancer: 2006 update of an American Society of Clinical Oncology practice guideline. J Clin Oncol. 2007; 25:1596–1605. [PubMed: 17404365]

16. D'Ambrosio DJ, Ruth K, Horwitz EM, Uzzo RG, Pollack A, Buyyounouski MK. How can men destined for biochemical failure after androgen deprivation and radiotherapy be identified earlier? Int J Radiat Oncol Biol Phys. 2008; 70:1487–1491. [PubMed: 18164854]

17. Johnson SB, Jackson WC, Murgic J, Feng FY, Hamstra DA. Time to nadir PSA: of popes and PSA —the immortality bias. Am J Clin Oncol. Sep 21.2013 [Epub ahead of print].

18. D'Amico AV, Chen MH, de Castro M, et al. Surrogate endpoints for prostate cancer-specific mortality after radiotherapy and androgen suppression therapy in men with localized or locally advanced prostate cancer: an analysis of two randomised trials. Lancet Oncol. 2012; 13:189–195. [PubMed: 22112793]

19. Pickles T. Prostate-specific antigen (PSA) bounce and other fluctuations: which biochemical relapse definition is least prone to PSA false calls? An analysis of 2030 men treated for prostate cancer with external beam or brachytherapy with or without adjuvant androgen deprivation therapy. Int J Radiat Oncol Biol Phys. 2006; 64:1355–1359. [PubMed: 16406391]

20. Shilkrut M, McLaughlin PW, Merrick GS, Vainshtein JM, Feng FY, Hamstra DA. Interval to biochemical failure predicts clinical outcomes in patients with high-risk prostate cancer treated by combined-modality radiation therapy. Int J Radiat Oncol Biol Phys. 2013; 86:721–728. [PubMed: 23664325]

21. Maffezzini M, Bossi A, Collette L. Implications of prostate-specific antigen doubling time as indicator of failure after surgery or radiation therapy for prostate cancer. Eur Urol. 2007; 51:605–613. [PubMed: 17113217]

22. Valicenti RK, DeSilvio M, Hanks GE, et al. 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]

23. Shipley WU, Desilvio M, Pilepich MV, et al. Early initiation of salvage hormone therapy influences survival in patients who failed initial radiation for locally advanced prostate cancer: a secondary analysis of RTOG protocol 86-10. Int J Radiat Oncol Biol Phys. 2006; 64:1162–1167. [PubMed: 16427211]

24. Buyyounouski MK, Pickles T, Kestin LL, Allison R, Williams SG. Validating the interval to biochemical failure for the identification of potentially lethal prostate cancer. J Clin Oncol. 2012; 30:1857–1863. [PubMed: 22508816]

25. Denham JW, Steigler A, Wilcox C, et al. Time to biochemical failure and prostate-specific antigen doubling time as surrogates for prostate cancer-specific mortality: evidence from the TROG 96.01 randomised controlled trial. Lancet Oncol. 2008; 9:1058–1068. [PubMed: 18929505]

26. Kapadia NS, Olson K, Sandler HM, Feng FY, Hamstra DA. Interval to biochemical failure as a biomarker for cause-specific and overall survival after dose-escalated external beam radiation therapy for prostate cancer. Cancer. 2012; 118:2059–2068. [PubMed: 22009287]
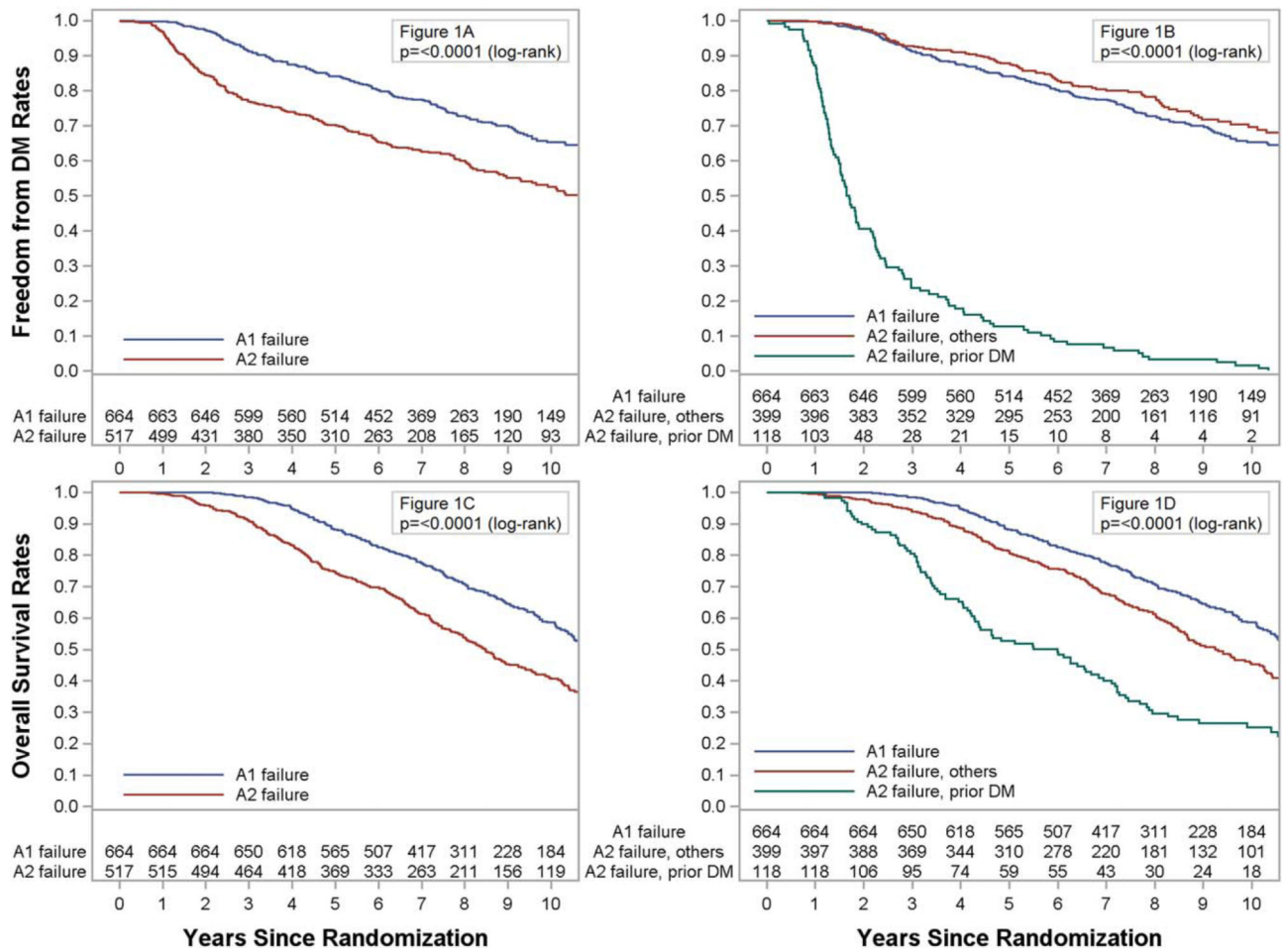
Figure 1.
(A) Freedom from distant metastasis (DM) as a function of A1 biochemical failure versus A2 biochemical failure. (B) Freedom from DM as a function of A1 or A2 biochemical failure or initial DM. (C) Overall survival as a function of A1 biochemical failure versus A2 biochemical failure. (D) Overall survival as a function of A1 or A2 biochemical failure or initial DM.

TABLE 1

Pretreatment Characteristics for RTOG 9202 and RTOG 9413 According to the ASTRO Definition of Biochemical Failure

	RTOG 9202			RTOG 9413		
	A1 (n = 352)	A2 (n = 311)	Total (n = 663)	A1 (n = 312)	A2 (n = 206)	Total (n = 518)
Age, y						
Median	69	69	69	69	69	69
Range	48-84	43-88	43-88	50-82	44-85	44-85
Prostate-specific antigen, ng/mL						
Median	24.8	24.9	24.8	26.50	25.90	26.20
Range	0.20-250.00	0.123-228.40	0.123-250.00	2.00-92.50	2.85-98.20	2.00-98.20
Clinical stage, n (%)						
T2	161 (46)	124 (40)	285 (43)	79 (25)	32 (15)	111 (21)
T3	177 (50)	167 (54)	344 (52)	35 (11)	16 (8)	51 (10)
T4	14 (4)	20 (6)	34 (5)	198 (63)	158 (77)	356 (69)
Gleason score, n (%)						
<7	121 (34)	90 (29)	211 (32)	68 (22)	43 (21)	111 (21)
7-10	206 (59)	195 (63)	401 (60)	244 (78)	163 (79)	407 (49)
Missing	25 (7)	26 (8)	51 (8)			
Nodal status, n (%)						
Negative	109 (31)	90 (29)	199 (30)	NA		
Positive	22 (6)	23 (7)	45 (7)			
Not done	221 (63)	198 (64)	419 (63)			

Abbreviations: ASTRO, American Society for Therapeutic Radiology and Oncology; NA, not applicable; RTOG, Radiation Therapy Oncology Group.

TABLE 2

PSA Kinetics in the Group With A2 Biochemical Failure

PSA Variable	Pattern 1: Rising PSA After RT Without Prior DM (n = 54 [10.4%])	Pattern 2: Falling PSA After RT With Nadir Without Prior DM (n = 284 [54.9%])	Pattern 3: Other Without Prior DM (n = 61 [11.8%])	Pattern 4: DM Occurring Before, at the Same Time, or Within 1 Month of Initiation of ADT (n = 118 [22.8%])
PSA at time of salvage HT, ng/mL	n = 54	n = 284	n = 26	n = 117
Median (range)	10.9 (1.4-179.8)	2.2 (0.0-172)	1.9 (0.1-297)	11.2 (0.0-780)
Mean (standard error)	–	5.5 (12.3)	17.1 (58.1)	47.1 (120.3)
PSA doubling time, mo	n = 36	n = 176	n = 34	n = 97
Median (range)	7.78 (1.27-110.7)	22.6 (3.02-2718)	18.7 (1.37-NA)[a]	11.4 (1.21-270.3)
Mean (standard error)	15.8 (21.3)	98.9 (292.8)	NA[a]	24.8 (46.5)
PSA nadir, ng/mL				n = 70
Median (range)	NA	0.20 (0.00-8.7)	NA	0.20 (0.00-26.8)
Mean (standard error)		0.43 (0.90)		1.21 (3.96)
Time to PSA nadir, mo				n = 70
Median (range)	NA	14.0 (0.99-106.5)	NA	11.8 (3.65-59.3)
Mean (standard error)		17.4 (13.9)		16.5 (13.3)
PSA change/mo[b]				n = 70
Median (range)	NA	0.063 (0-8.1)	NA	0.27 (0-33.4)
Mean (standard error)		0.24 (0.66)		1.86 (5.24)
5-year failure/survival rate, % (95% CI)[c]				
Overall survival	51.0 (36.9-63.5)	90.3 (86.2-93.3)	63.9 (50.1-74.8)	52.8 (43.3-61.4)
DM	37.3 (24.4-50.1)	7.9 (5.1-11.4)	6.7 (2.1-15.1)	87.3 (80.6-92.5)
Local failure	24.1 (13.7-36.2)	21.0 (16.4-25.9)	21.8 (12.3-33.0)	20.4 (13.6-28.1)

Abbreviations: ADT, androgen-deprivation therapy; CI, confidence interval; DM, distant metastasis; HT, hormone therapy; NA, not applicable; PSA, prostate-specific antigen; RT, radiation therapy.

[a] The results are too big because of the random patterns of PSA.

[b] PSA change = [(PSA at failure – PSA at nadir)/(Date of PSA failure – Date of PSA nadir] × 30.5.

[c] The cumulative incidence method was used to estimate distant and local failure rates, and Kaplan-Meier estimation was used for overall survival rates.

TABLE 3

Survival and Failure Rates at 5 Years According to the ASTRO Definition of Biochemical Failure

Endpoint	Biochemical Failure	Patients, n	Failure, n	5-Year Failure/Survival Rate, % (95% CI)[a]	HR (95% CI)[b]	P
Combined data (n = 1181)						
Overall survival	A1	664	267	88.2 (85.7-90.7)	RL	
	A2	517	299	74.6 (70.8-78.4)	1.68 (1.48-1.99)	<.0001
Distant metastasis	A1	664	193	15.7 (12.9-18.4)	RL	
	A2	517	212	29.0 (25.1-33.0)	1.60 (1.32-1.95)	<.0001
Local failure	A1	664	176	19.6 (16.5-22.6)	RL	
	A2	517	135	21.3 (17.7-24.8)	1.01 (0.81-1.27)	.92
RTOG 9202 (n = 663)						
Overall survival	A1	352	178	88.1 (84.7-91.5)	RL	
	A2	311	201	75.1 (70.3-79.9)	1.59 (1.30-1.94)	<.0001
Distant metastasis	A1	352	127	17.3 (13.4-21.3)	RL	
	A2	311	139	29.3 (24.2-34.4)	1.42 (1.12-1.80)	.004
Local failure	A1	352	102	22.2 (17.8-26.5)	RL	
	A2	311	90	25.1 (20.3-30.0)	1.03 (0.78-1.37)	.086
RTOG 9413 (n = 518)						
Overall survival	A1	312	89	88.3 (84.6-91.9)	RL	
	A2	206	98	73.8 (67.7-80.0)	1.97 (1.48-2.62)	<.0001
Distant metastasis	A1	312	66	13.7 (9.9-17.6)	RL	
	A2	206	73	28.6 (22.4-34.9)	1.94 (1.39-2.70)	<.0001
Local failure	A1	312	74	16.6 (12.5-20.8)	RL	
	A2	206	45	15.6 (10.5-20.7)	0.93 (0.64-1.35)	.70

Abbreviations: ASTRO, American Society for Therapeutic Radiology and Oncology; CI, confidence interval; HR, hazard ratio; RL, reference level; RTOG, Radiation Therapy Oncology Group.

[a] The cumulative incidence method was used to estimate distant and local failure rates, and Kaplan-Meier estimation was used for overall survival rates.

[b] An HR quantifies how much more (less) risk that patients at some level have in comparison with those at the RL. A CI that includes 1 indicates no difference between these 2 subgroups.

TABLE 4

Multivariate Models for Overall Survival, Distant Metastasis, Local Failure, and Salvage ADT Use for Patients Treated in RTOG 9202 or RTOG 9413

Endpoint	Covariate	Comparison	HR (95% CI)	P
Overall survival[a]	BF type	A1	RL	
		A2	1.66 (1.40-1.97)	<.0001
	Age, y	<70	RL	
		70	1.43 (1.20-1.69)	<.0001
	Gleason score	2-6	RL	
		7-10	1.54 (1.26-1.89)	<.0001
Distant metastasis[a]	BF type	A1	RL	
		A2	1.60 (1.31-1.95)	<.0001
	Gleason score	2-6	RL	
		7-10	1.53 (1.18-1.97)	.0012
	Stage	T2	RL	
		>T2	1.70 (1.31-2.20)	<.0001
	Study	9202	RL	
		9413	0.68 (0.54-0.85)	.0007
Local failure[a]	BF type	A1	RL	
		A2	0.97 (0.77-1.22)	.78
	Stage	T2	RL	
		>T2	1.54 (1.18-2.02)	.0016
	Study	9202	RL	
		9413	0.69 (0.54-0.88)	.0024
Time to salvage ADT after BF (A1 failure only)	Age, y	<70	RL	
		70	0.91 (0.75-1.10)	.34
	Gleason score	2-6	RL	
		7-10	1.12 (0.90. 1.39)	.32
	Stage	T2	RL	
		>T2	1.07 (0.87-1.31)	.54
	Study	9202	RL	
		9413	1.40 (1.15-1.70)	<.001

Abbreviations: ADT, androgen-deprivation therapy; BF, biochemical failure; CI, confidence interval; HR, hazard ratio; RL, reference level; RTOG, Radiation Therapy Oncology Group.

[a] All the outcomes of the combined data in the multivariate proportional hazards models were adjusted by BF type (A1 [RL]), age (<70 [RL] vs 70 years), combined Gleason score (2-6 [RL] vs 7-10), PSA (30 [RL] vs >30 ng/mL), clinical classification (T2 [RL] vs >T2), and study (RTOG 9202 [RL] vs RTOG 9413).

TABLE 5

Multivariate Proportional Hazards Models for RTOG 9202 and RTOG 9413: Time to Biochemical Failure for Those Meeting the A1 Definition of Biochemical Failure

Covariate	Comparison	Adjusted HR (95% CI)[a]	P
PSA nadir Study	Continuous	0.93 (0.81-1.1)	.34
	9202	RL	
	9413	1.54 (1.32-1.79)	<.0001
Time to PSA nadir Study	Continuous	0.99 (0.99-0.995)	<.0001
	9202	RL	
	9413	1.43 (1.24-1.66)	<.0001
Rate of PSA rise[b] Study	Continuous	1.19 (1.12-1.27)	<.0001
	9202	RL	
	9413	1.55 (1.33-1.80)	<.0001

Abbreviations: CI, confidence interval; HR, hazard ratio; PSA, prostate-specific antigen; RL, reference level; RTOG, Radiation Therapy Oncology Group.

[a] An HR quantifies how much more (less) risk that patients at some level have in comparison with those at the RL. A CI that includes 1 indicates no difference between these 2 subgroups. All the outcomes of the combined data in the multivariate proportional hazards models were adjusted by age (<70 [RL] vs 70 years), combined Gleason score (2-6 [RL] vs 7-10), PSA (30 [RL] vs >30 ng/mL), clinical classification (T2 [RL] vs >T2), and study (RTOG 9202 [RL] vs RTOG 9413).

[b] Rate of PSA rise = (PSA at failure – PSA at nadir)/(Date of PSA failure – Date of PSA nadir).