Improved Overall Survival Trends of Men with Newly Diagnosed M1 Prostate Cancer: A SWOG Phase III Trial Experience (S8494, S8894 and S9346)


From the Fred Hutchinson Cancer Research Center (CMT) and University of Washington (CSH), Seattle, Washington, University of Michigan (MHAH), Ann Arbor and Wayne State University (UP), Detroit, Michigan, Johns Hopkins Oncology Center (MAE), Baltimore, Maryland, University of California-San Francisco (EJS), San Francisco, California, University of Wisconsin (GW), Madison, Wisconsin, Tom Baker Cancer Center (BJD), Calgary, Alberta, Canada, Eastern Virginia Urologic Cancer Outreach Program (PFS), Norfolk, Virginia, University of Colorado (EDC), Denver, Colorado, Comprehensive Cancer Centers of Nevada (NJV), Las Vegas, Nevada, and University of Texas Health Science Center (IMT), San Antonio, Texas

Abbreviations and Acronyms
AA = black American
ADT = androgen deprivation therapy
M1 = metastatic
PS = performance status
PSA = prostate specific antigen

Purpose: Frequent prostate specific antigen testing for screening and monitoring prostate cancer has led to significant stage migration. We evaluated whether overall survival in hormone naïve patients with metastatic prostate cancer has improved during the era of prostate specific antigen use. We also assessed whether any patient subsets benefited differentially during this period.

Materials and Methods: We compared overall survival in 3 sequential phase III trials of 3,096 men with hormone naïve, metastatic prostate cancer who received similar androgen deprivation therapy, including 2 trials performed before the prostate specific antigen era (S8494 and S8894) and the other done during this era (S9346). Overall survival was adjusted for patient and disease risk factors in the latter 2 trials. Subgroups were evaluated by interactions of risk factors with trial.

Results: Median overall survival was 30 months in S8494, 33 months in S8894 and 49 months in S9346. Adjusting for risk factors, there was a 22% lower risk of death in S9346 than in S8894 (HR 0.78, 95% CI 0.70, 0.87, p <0.001). The improvement in overall survival was greater in black American men (test of interaction p = 0.008). In S8494 and S8894 median survival for black men was 27 months, and 34 and 35 months for nonblack men, respectively. This racial difference disappeared in S9346 with overall survival of 48 and 49 months in black and nonblack men, respectively.

Conclusions: Adjusting for risk factors, overall survival was significantly improved in the post-prostate specific antigen era trial. However, it cannot be concluded that this was attributable only to prostate specific antigen monitoring. Black men now have overall survival comparable to that of white men. Current estimates of survival should be used to design new trials in this population.

Key Words: prostate, prostatic neoplasms, neoplasm metastasis, mortality, prostate-specific antigen

Submitted for publication March 20, 2012.
Study received institutional review board approval.
Supported by Public Health Service Cooperative Agreement Grants CA32102, CA38926, CA14028, CA59882, CA42777, CA53192, CA64441, CA46282, CA27515, CA68567, CA26361, CA35261, CA53431, CA46398, CA68489, CA65756, CA67863, CA58488, CA87918, CA35281, CA63844, CA45660, CA11080, CA5178, CA69880, CA35119, CA35860, CA98046, CA46407, CA46061, CA46377, CA45800, CA35128, CA35282, CA46133, CA58116, CA46128, CA76132, CA58901, CA68668, CA8182, CA12844, CA22433, CA76404, CA46136, CA37581, CA35176, CA65860, CA35176, CA27585, CA16395, CA12213, CA16462,
Androgen deprivation therapy has been the standard treatment for advanced prostate cancer since the observations by Huggins and Hodges in 1941. During the ensuing 7 decades, several questions regarding the use and efficacy of ADT have been evaluated in phase III trials, including the efficacy of medical vs surgical castration, gonadal suppression vs combined androgen deprivation and intermittent vs continuous ADT. The majority of these trials accrued patients before the implementation of PSA into clinical practice, which generally began around 1990. The routine use of PSA to detect and monitor disease activity has led to a significant stage migration of prostate cancer. Data from the CaPSURE™ database, including 2,078 men with low risk prostate cancer who were diagnosed between 1989 and 2001, demonstrated a significant increase in the proportion of patients with low risk tumor characteristics from 29.8% in 1989 to 45.3% in 1999 to 2001. The most recent data from the SEER (Surveillance, Epidemiology and End Results) program of the National Cancer Institute suggest 99% 5-year relative survival among men with prostate cancer diagnosed in 2000, a significant increase from the 78% among men diagnosed in 1986. Some have argued that improvements in relative and cause specific survival rates do not reflect true increases in life expectancy since survival rates among men diagnosed during the PSA era can be inflated by over diagnosis and lead time. An unanswered and important question is whether men with hormone naïve metastatic prostate cancer have experienced similar improved survival during the PSA era.

An additional observation made from 1990 to 2006 was a 39% decrease in prostate cancer mortality. This decrease in mortality has been largely attributable to widespread PSA screening and yet 2 large prospective trials showed a 20% decrease in prostate cancer mortality or no mortality decrease. It is possible that the observed reduction in prostate cancer mortality is related to improved outcomes in men with locally advanced or metastatic disease, which are disease stages responsible for the majority of prostate cancer deaths.

In the last 3 decades SWOG has performed a series of phase III trials. S8494 (enrolled 1985 to 1987) compared survival in patients randomized to leuprolide or leuprolide with flutamide. S8894 (enrolled 1989 to 1994) evaluated the impact of adding flutamide to bilateral orchiectomy. S9346 (accrued 1995 to 2009) evaluated whether intermittent ADT is noninferior to continuous ADT. These sequential trials with survival end points provide an opportunity to address the question of whether survival has changed since the advent of widespread PSA screening and followup testing. Patient populations and major eligibility criteria were comparable. Patients were enrolled by cooperative groups with members from major cancer centers and community centers around the United States. All patients received protocol specified treatment and were followed for disease progression and survival in a comparable way so that differences in prognosis would be less likely due to access to health care.

We hypothesized that, during an era of increased PSA testing, there has been an improvement in survival for patients with newly diagnosed M1 prostate cancer.

MATERIALS AND METHODS

SWOG has coordinated 3 phase III trials in newly diagnosed stage M1 prostate cancer cases since the mid 1980s. S8494 and S8894 represent the pre-PSA era. However, S8494 cannot be used for multivariate risk analysis because PSA and Gleason score were not collected. S9346 represents the post-PSA era. Although a small fraction of men enrolled on S8894 may have been exposed to PSA screening or disease monitoring, the majority were not.

S8494 (Enrolled 1985 to 1986), PDQ® Registry 10455554

S8494 (intergroup study 0036) enrolled patients with metastatic adenocarcinoma of the prostate who were hormone naïve and had a SWOG PS of 0 to 3, and adequate renal and hematological function. All patients received leuprolide and then were randomized to receive flutamide (250 mg 3 times daily) or matching placebo. Patients were treated until progression and those randomized to placebo could cross over to flutamide at progression. From March 18, 1985 to April 2, 1986 a total of 617 patients were randomized.

S8894 (Enrolled 1989 to 1994), PDQ Registry 8656627

Like S8494, S8894 (intergroup study INT-0105) enrolled hormone naïve patients with metastasis. Eligibility was almost identical. All patients underwent bilateral orchiectomy and were subsequently randomized to receive flut-
amide (250 mg 3 times daily) or a matching placebo. Crossover at progression was allowed. The primary end point of the study was death from any cause. Patients underwent disease assessment at study entry and every 6 months for the first 2 years. After that, radiological studies were done if PSA increased more than 25 ng/ml during any 3-month period. Patients were treated until disease progression. All patients were followed until death. Gleason grading was performed through centralized pathological review of the biopsy specimens obtained before study entry. When not centrally available, the local Gleason score was used.

From December 15, 1989 to September 15, 1994, a total of 1,387 patients were randomized. Two-thirds of the patients came from a total of 99 SWOG institutions in 30 states and the other third came from ECOG (Eastern Cooperative Oncology Group) institutions. A total of 1,286 patients were clinically eligible and evaluated for this analysis.

S9346 (Accrued 1995 to 2009), ClinicalTrials.gov identifier NCT00002651

S9346 is a phase III intergroup trial (INT 0162) with participants from CALGB (Cancer and Leukemia Group B), ECOG, EORTC (European Organisation for Research and Treatment of Cancer) and NCIC-CTG (National Cancer Institute of Canada Clinical Trials Group). Its primary objective is to determine whether survival is noninferior with intermittent ADT compared to continuous ADT in patients with newly diagnosed, hormone sensitive metastatic prostate cancer. Key eligibility requirements include stage M1 prostate cancer, a minimum pretreatment PSA of 5 ng/ml and PS 0 to 2.

Step 1 of treatment consists of a 7-month induction course with goserelin and bicalutamide. Patients could have started ADT within 6 months before registration if otherwise eligible. Patients in whom PSA decreased to 4 ng/ml or below at months 6 and 7 were randomly assigned to intermittent or continuous ADT. Patients in whom PSA did not decrease to 4 ng/ml or less at the end of induction were removed from protocol but followed for survival. Standard of care for these patients was continuation of hormonal therapy until progression. PSA was measured at months 1, 4, 6 and 7 of the induction period, and then monthly after randomization.

Final results of S9346 were expected in the summer of 2012. The SWOG Data and Safety Monitoring Committee has given permission for these results to be reported by pooled arms.

All eligible men from S8494 and S8894, and all eligible men from North America from S9346 were considered for this analysis. Only patients with complete risk factor information were included in analysis. EORTC (Europe) patients were excluded from S9346 for this analysis because they did not report race.

Survival was defined for S8494 and S8894 from date of randomization to date of death from any cause or censored at the last contact date. For S9346 the survival interval started at the treatment start date for those who started before enrollment and at the enrollment date for those who had not yet started. A Cox regression model was used to assess the risk of death for S9346 compared to S8894, adjusting for other known risk factors. Extensive disease was defined exactly the same for all 3 trials as diffuse bone disease or visceral organ involvement. The Kaplan-Meier method was used to estimate survival curves. All analyses were done using SAS®, version 9.1. A main effect or interaction test was considered statistically significant at 2-sided p ≤ 0.05.

All patients from these 3 trials provided written institutional review board approved informed consent before study entry.

RESULTS

Table 1 shows the distribution of risk factors for each study. Of S9346, S8894 and S8494 patients 18%, 38% and 16%, respectively, were excluded from these analyses due to missing risk factors, primarily missing PSA and Gleason scores for the 2 more recent trials, leaving 509 from S8494, 791 from S8894 and 1,796 from S9346. There were more extensive disease, a higher prevalence of bone pain,

| Table 1. Patient and disease factors for eligible patients with complete data from each trial |
|-----------------|-----------------|-----------------|
|                | S8494           | S8894           | S9346           |
| No. pts        | 509             | 791             | 1,796           |
| No. extensive vs minimal disease (%) | 439 (86) | 642 (81) | 1,209 (67) |
| No. bone pain (%) | 267 (52) | 364 (46) | 652 (36) |
| No. race (%):  |
| AA             | 99 (19)         | 187 (23.6)      | 317 (17.7)      |
| White          | 409 (80)        | 590 (74.6)      | 1,440 (79.0)    |
| Other          | 1 (0.2)         | 14 (7.8)        | 68 (3.7)        |
| No. PS 2–3 vs 0–1 (%) | 79 (15.5) | 28 (3.5) | 115 (6.3) |
| No. Gleason score (%): Not available | 249 (32) | 313 (17) | |
| 6 or Less      | 240 (30)        | 614 (34)        | |
| 7              | 302 (38)        | 896 (49)        | |
| 8–10           | 68 (62, 74)     | 70 (64, 75)     | 68 (61, 75)     |
| Median age (25%, 75%) | 68 (47, 516) | 51 (17, 195) |
| Median ng/ml PSA (25%, 75%) | Not available | 287 (56%) | 1,341 (74%) |

* Body mass index 25 kg/m² or greater.
higher PSA values, more AA men and fewer obese men in the earlier study, S8894, compared to S9346 on multivariate analysis.

Figure 1 shows survival curves stratified by clinical trial. Median survival was 30 months in S8494, 33 months in S8894 and 49 months in S9346. The HR of 0.70 indicated a 30% decreased risk of death in the more recent S9346 compared to that in the older S8894 trial (95% CI 0.64, 0.77, p < 0.001). If all data from all eligible men from each trial were used regardless of the completeness of risk factor information, the survival HR would remain 0.70, suggesting that those with complete risk factor data are representative of the 2 trials.

Table 2 shows the corresponding multivariate adjusted analysis of all-cause mortality. Factors associated with an increased risk of death included extensive (vs minimal) disease, bone pain, older age, AA vs other race, body mass index less than 25 kg/m², performance status 2 or 3 and Gleason score 8 to 10 (vs 2 to 6). PSA at study entry and Gleason score 7 (vs 6 or less) were not significantly associated with survival (p < 0.05). After adjusting for these factors, S9346 showed a lower risk of death than S8894 (HR 0.78, 95% CI 0.70, 0.87, p < 0.001). Adjustment for risk factors explained some of the difference in survival between the 2 trials but not completely. There was still a 22% lower risk of death on S9346 compared to S8894.

The interaction of each risk factor with clinical trial was assessed. The interaction test evaluated whether the prognostic significance of a risk factor with survival was different for the 2 trials. Only the interaction with AA patients was statistically significant (p = 0.008). AA patients had worse survival than nonAA patients in the S8894 trial after risk factor adjustment but the disparity did not persist in S9346. Figure 2 shows the unadjusted survival curves stratified by clinical trial and AA status. Survival patterns were similar for the 2 pre-PSA trials. Each racial group showed improvement with time compared to S9346 but the AA group survival improvement was greater than that of the nonAA group (median increase 21 vs 14 to 15 months). For S9346 AA and nonAA patients had survival curves that overlapped, indicating that racial group was not a significant predictor of survival in patients with M1 prostate cancer in the post-PSA era trial. We included nonwhite, nonAA patients in the nonAA category, including 2 in S8494, 14 in S8894 and 68 in S9346. When excluded from analysis, the results remained unchanged.

Table 3 shows a comparison of patient and disease characteristics for AA patients in the 3 trials. There was a higher prevalence of extensive disease and bone pain in the earlier trials, and PSA values were also substantially higher in S8894 than in S9346. Age was comparable but there was more obesity in the more recent trial, S9346.

**DISCUSSION**

The wide use of PSA to detect and monitor disease activity has led to significant stage migration of...
prostate cancer. Better local control and more detailed imaging may also have contributed to improved outcomes in the PSA era. We evaluated the potential change in survival in the setting of metastatic prostate cancer using 3 sequential, phase III, cooperative group trials with comparable eligibility criteria. Our analysis indicates overall improvement in risk adjusted survival for nonAA and AA men. Of note is the resolution of the disparity in survival between the races in S9346. There are several possible explanations for the favorable survival observed in the latter trial. After adjusting for known risk factors, such as PS, disease extent, PSA and bone pain, the risk of death on S9346 was lower than on S8894. This may be due to the model not containing precise enough risk factors to adequately model survival, such as detailed enumeration and extent of bone lesions. There are no validated risk models in the setting of hormone sensitive disease. We did not have data on hemoglobin, lactate dehydrogenase and alkaline phosphatase for these trials, which are possibly important measures since they were identified as prognostic factors in the Halabi nomogram.13 Additionally, assessment of PSA velocity in the period before the metastatic diagnosis or more detailed disease extent information might also explain some of the discrepancy between the trials.14,15 Unfortunately, this level of detail is not available for these trials. We used the best disease predictors that we could identify but the multivariate model is likely incomplete.

Each trial used contemporary hormonal therapy. Different antiandrogens (flutamide vs bicalutamide) may have had some role in the different trial prognoses. Although half of S8894 patients were randomized to placebo after bilateral orchiectomy, the survival differences between trials cannot be explained by a lack of antiandrogen use since at long-term followup there was no statistically significant survival difference between the 2 arms of S8894 (median survival 2.8 years for flutamide and 2.5 years for placebo, log rank p = 0.16). We also assessed the flutamide by race interaction (p = 0.59).

S9346 was designed as a noninferiority trial. Since the Data and Safety Monitoring Committee has let this trial continue to the final planned analysis, it is unlikely that either arm is substantially better than the other. Another potential contributor to the observed differences in survival is the overall improved health care for nonprostate cancer and prostate cancer, including the introduction of docetaxel as a standard therapy for castration resistant prostate cancer. However, S9346 patients who were randomized before and after 2004, when docetaxel received Food and Drug Administration approval, showed no difference in median survival (46.3 months before and 45.1 after 2004).

Most gratifying is the improvement in advanced prostate cancer outcomes overall and particularly for AA men. Our previous analysis of S8894 showed that these men experienced worse outcomes from the disease even in the context of carefully overseen therapy in a clinical trial.16 When we evaluated ZIP Code summary information regarding income and education between the 2 trials (table 3), there was no evidence of a shift in the socioeconomic status with time. Also, the 5-year survival of AA and white patients in the SWOG trials are comparable to those reported by SEER for men with newly diagnosed metastatic prostate cancer,17 suggesting this is a fairly representative sample. A strength of this analysis is that all patients received the same protocol specified treatment, and comparable followup and disease monitoring.

Therefore, we hypothesize that the survival improvement for all patients, particularly AA men, is based on a significant shift to less extensive disease. A similar finding of improved disease-free survival was reported among AA men compared to white men diagnosed with clinically localized prostate cancer and treated with radical prostatectomy (1991 to 1995 vs 1996 to 1999).18 We would further hypothesize that

### Table 3. Risk factors among AA men stratified by clinical trial

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>S8484</th>
<th>S8894</th>
<th>S9346</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. AA men</td>
<td>99</td>
<td>187</td>
<td>317</td>
</tr>
<tr>
<td>No. extensive disease (%)</td>
<td>90 (91)</td>
<td>165 (88)</td>
<td>219 (70)</td>
</tr>
<tr>
<td>No. bone pain (%)</td>
<td>65 (66)</td>
<td>101 (54)</td>
<td>110 (35)</td>
</tr>
<tr>
<td>No. PS 2–3 (%)</td>
<td>26 (26)</td>
<td>13 (7)</td>
<td>33 (10)</td>
</tr>
<tr>
<td>No. Gleason grade (%):</td>
<td>Not available</td>
<td>Not available</td>
<td>Not available</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>60 (32)</td>
<td>111 (35)</td>
</tr>
<tr>
<td>9–10</td>
<td>49 (49)</td>
<td>81 (43)</td>
<td>202 (64)</td>
</tr>
<tr>
<td>Median age (25%, 75%)</td>
<td>67 (62, 72)</td>
<td>68 (63, 75)</td>
<td>66 (60, 73)</td>
</tr>
<tr>
<td>Median ng/ml PSA (25%, 75%)</td>
<td>Not available</td>
<td>288 (98, 1000)</td>
<td>117 (37, 526)</td>
</tr>
<tr>
<td>% Individuals with college education</td>
<td>15 (11, 23)</td>
<td>17 (12, 26)</td>
<td></td>
</tr>
<tr>
<td>Income (× $1,000)</td>
<td>19.0 (14.2, 25.0)</td>
<td>21.7 (15.3, 30.1)</td>
<td></td>
</tr>
</tbody>
</table>
this improvement is based on greater awareness of prostate cancer and improved health seeking behavior among AA men. The disappearance of this disparity is an important advance in the care of men with prostate cancer. However, AA men had a twofold to threefold greater incidence of newly diagnosed metastatic prostate cancer compared to white men at ages 40 to 69 years from 1995 to 2007, which contributes to a similarly increased mortality rate.\(^{17,19}\) A greater effort is needed to eliminate disparities in prostate cancer.

**CONCLUSIONS**

Based on SWOG trial data, men with newly diagnosed M1 prostate cancer who are hormone naive have significantly better survival in the PSA era after accounting for risk factors. AA men have greater absolute gains in survival compared to nonAA men. The current prognosis in AA men is similar to that in nonAA groups. Results suggest that these gains may be at least partially attributable to PSA monitoring.

**REFERENCES**


