FDA Public Workshop on Clinical Trial Endpoints in Prostate Cancer

June 21-22, 2004 – Bethesda, Maryland

Summary

Monday, June 21 – Afternoon session

Hormone-Sensitive Prostate Cancer: Registration Trial Endpoints (Donna J. Griebel, MD)

Dr. Griebel commented that the minutes of FDA advisory committee meetings from the 1980s indicate a lack of consensus on the role and/or significance of PSA as a marker for prostate cancer progression that continues today.

Three classes of hormonal therapies have been approved for the treatment of prostate cancer, Dr. Griebel said: GnRH analogs/agonists (leuprolide, 1985); nonsteroidal anti-androgens (flutamide, 1988); and GnRH antagonists (abarelix, 2003). (The drug diethylstibestrol [DES] was labeled for the treatment of prostate cancer in the 1950s, before passage of the Food, Drug, and Cosmetics Act of 1962, which required that drugs demonstrate effectiveness prior to marketing.) The disease states for which therapies have been approved are locally confined B2 and C disease before and during radiotherapy and advanced carcinoma of the prostate (D2 disease).

Among approvals for the treatment of D2 disease, survival was the primary endpoint for both leuprolide and flutamide. The primary endpoint in the abarelix approval was avoidance of orchiectomy within 4 and 12 weeks of treatment. In 1996, flutamide and goserelin were approved for the treatment of locally confined B2 and C disease before and during radiotherapy; in both cases, the approval endpoint was disease-free survival (DFS). Subsequent approvals of anti-androgens for D2 disease were also based on survival.

Testosterone was used as a surrogate endpoint in the 1989 approval of leuprolide in the depot formulation. In the same year, goserelin was approved on the basis of survival and testosterone. Subsequently, sponsors were asked to show that other GnRH analogs had an impact on testosterone and were non-inferior to leuprolide. Reviewers uniformly accepted Huggins' work demonstrating the impact of hormonal agents on prostate cancer and agreed that testosterone suppression, although it might not improve survival, did make patients feel better.

FDA's 1998 guidance document on effectiveness states (Section d): "When the pathophysiology of a disease and the mechanism of action of therapy are very well understood, it may be possible to link specific pharmacologic effects to a strong likelihood of clinical effectiveness. A pharmacologic effect that is accepted as a validated surrogate endpoint can support ordinary approval." The guidance document goes on to caution against the acceptance of a surrogate in all situations: 'The reasons for the absence of an expected correlation between pharmacologic and clinical effects…can include an incompletely understood relationship between the pharmacologic effect and

the clinical benefit and the presence of other pharmacologic effects attributable to a drug in addition to the effect being measured and thought to be beneficial."

Dr. Griebel noted that although DES and ketoconazole have testosterone-suppressing effects similar to those of GnRH analogs, the scope of the mechanism of action of these non-GnRH drugs is unclear, whereas it is known that GnRH analogs act directly through their impact on testosterone production.

In the case of the GnRH antagonist abarelix, the original intent was to use testosterone as an endpoint. However, data analysis revealed a risk of anaphylaxis and questions about the sustainability of testosterone suppression. It was therefore felt that the risk-benefit analysis should be based on the drug's effect on a true clinical benefit endpoint rather than on a surrogate. The clinical benefit endpoint that served as the basis for approval of abarelix was avoidance of orchiectomy within 4 and 12 weeks of the initiation of treatment.

Endpoints for Clinical Trials in Early Prostate Cancer: PSA Progression After Radical Prostatectomy (Peter T. Scardino, MD)

Dr. Scardino presented findings from a study in which he and his colleagues attempted to determine the prognostic significance of a rising PSA after radical prostatectomy (RP). The investigators analyzed data from 4,035 patients who were treated with RP alone between 1983 and 2003 and followed for a median of 6 years. Close to 60% of patients had palpable tumors, 30% had PSA values >10 ng/ml, and 30% had tumors with a Gleason grade $\geq$7.

Rising PSA was considered recurrence if a PSA value >0.1 ng/ml was confirmed by two subsequent rises (not necessarily consecutive), if secondary therapy for cancer was initiated due to a rise in PSA, or if the absolute PSA value was >0.4 ng/ml. Seventy-four percent of patients included in the analysis never met these criteria for recurrence.

The likelihood of recurrence in an individual patient was most accurately predicted by a nomogram that combined the patient's baseline PSA value, clinical stage of disease, and tumor Gleason grade. (A Palm Pilot version of the nomogram software can be downloaded at no charge at http://www.mskcc.org/prostate/nomograms.) For individual patients, nomogram-based predictions were more accurate than those based on stratification by risk group.

When PSA is rising in an RP-treated patient, it is impossible to tell from the biochemical marker whether the recurrence is local, distant, or both. Additionally, in men with such low PSA levels, imaging studies are generally unhelpful. Dr. Scardino and his colleagues analyzed data from patients with rising PSA who were treated with radiation therapy to determine the prevalence of local vs. distant recurrence (Stephenson et al, JAMA 2004;291:1325-1332). They found that 67% of patients had a complete response to radiation therapy (that is, their PSA levels became undetectable) and 30% had a durable response.

Relatively few studies have examined the probability over time of metastases and prostate cancer death in men with a rising PSA. A competing-risk analysis presented by Dr. Scardino and his colleagues at the 2004 American Society of Clinical Oncology annual scientific meeting concluded that at 15 years the probability of death from prostate cancer or from any other cause was roughly equivalent (one-third of patients died of prostate cancer and one-third of other causes within 15 years). The probability that a patient would remain free of distant metastases was 66% at 10 years and 36% at 15 years. In addition to absolute PSA value, PSA doubling time, tumor Gleason grade, and the presence of seminal vesicle invasion were significant predictors of metastatic progression within 8 years for patients with rising PSA.

Pound et al (JAMA 1999;281(17):1591-7) found that the median time to development of metastases in untreated men with rising PSA was 7.5 to 8 years. Key risk factors for the development of metastases were the time elapsed before PSA began to rise, PSA doubling time, and the pathologic characteristics of the primary tumor.

Dr. Scardino and his colleagues also analyzed the probability that patients with rising PSA after RP would receive androgen deprivation therapy (ADT). They found that the median time from PSA progression to initiation of ADT was 4.3 years. Between 20% and 25% of patients remained untreated after 10 to 15 years although their PSA values continued to rise. The median duration of response to ADT was 10.9 years, making survival a problematic endpoint in this population.

When PSA began to rise after ADT, the median time to development of metastases was 9.5 months and median survival was 26.4 months; no patients survived for 5 years. In this disease state, no risk factors predicted duration of survival. Dr. Scardino concluded that rising PSA may be a useful endpoint in PSA castrate disease.

Evidence Supporting a PSA(t) Endpoint as Proof of Clinical Efficacy (Anthony V. D'Amico, MD, PhD)

PSA level alone provides little useful information about the individual patient in any disease state, said Dr. D'Amico. PSA failure alone is similarly uninformative; as Dr. Scardino showed, PSA failure alone is not an accurate indicator of a patient's risk of death from prostate cancer. Dr. D'Amico said he would present data to demonstrate how PSA doubling time can be used within each disease state to prospectively identify those patients who will die of prostate cancer. Sufficient evidence exists from studies involving a total of 15,000 patients to show that PSA doubling time does predict and may be a surrogate for cancer-specific death. The major limitation of these studies is that they are retrospective analyses rather than randomized trials.

PSA time-dependent parameters as a surrogate for cancer-specific death have been studied in three specific disease states: PSA doubling time following surgery or radiation therapy (RT), PSA velocity in hormone-refractory M0 disease, and PSA velocity in hormone-refractory M1 disease. (*Note:* Data on PSA velocity in hormone-refractory M1

disease were presented by Dr. Daniel Petrylak on Tuesday, June 22.) Additional studies provide ancillary evidence of the importance of time-dependent PSA constructs (PSA velocity prior to diagnosis and PSA slope ratio in both hormone-sensitive and hormone-refractory disease) as prognostic factors.

PSA Doubling Time Following Surgery or RT

In this study (an analysis of data from 8,669 patients who were treated with either surgery or radiation), a PSA doubling time of <3 months accurately identified the 1 in 5 patients who sustained PSA failure who died of prostate cancer. In a Cox regression analysis, a PSA doubling time of <3 months was a prognostic factor for both cancer-specific and all-cause mortality. When PSA doubling time was <3 months, treatment did not predict either cancer death or all-cause death.

In a hazard ratio analysis, the addition of confounders and PSA doubling time to treatment resulted in a hazard ratio of 1.0 (95% confidence interval 0.6–1.9). For patients with PSA doubling times of <3 months, curves for prostate-cancer–specific survival and overall survival were almost identical, indicating that prostate cancer was the sole cause of death in this group of patients. In the setting of PSA failure following local therapy, therefore, a PSA doubling time of <3 months following local therapy may be a surrogate for both cancer-specific and all-cause mortality.

PSA Velocity in the Hormone-Refractory M0 Patient

In this study, currently under review, data were analyzed from 900 patients who received either surgery or RT followed by salvage hormonal therapy, which was given at the time of PSA failure and before patients had positive bone scans. 40% of these patients died of prostate cancer while the remainder died of other causes.

A PSA velocity >1.5 ng/ml/yr was a prognostic factor for both cancer-specific and all-cause mortality. In a hazard ratio analysis, the addition of confounders and PSA doubling time to treatment resulted in a hazard ratio of 1.8 (95% confidence interval 0.7–4.8).

Among patients with PSA velocity >1.5 ng/ml/yr, almost all deaths were due to prostate cancer, whereas among patients with slower PSA velocities there were no deaths from prostate cancer. In the setting of hormonal therapy following PSA failure, therefore, PSA velocity >1.5 ng/ml/yr following hormonal therapy and PSA failure may be a surrogate for both cancer-specific and all-cause mortality.

Ancillary Data

PSA velocity prior to diagnosis
In this study, data were analyzed from 1,095 patients with localized disease who were treated with surgery and enrolled in a screening PSA study between 1989 and 2002. The patients' median PSA value was 4.3 ng/ml; the median length of follow-up was 5 years.

Most patient deaths were due to causes other than prostate cancer. Patients whose PSA velocity was >2 ng/ml/yr prior to diagnosis had significantly higher cancer-specific and all-cause mortality in the 10 years following radical prostatectomy compared with patients with PSA velocities <2 ng/ml/yr. These findings suggest that even prior to diagnosis, PSA kinetics have prognostic significance.

PSA slope ratio
PSA level is not a useful prognostic indicator in the patient on hormonal therapy. Analysis of the PSA slope ratio may be an alternative method of using PSA to define prognosis in this patient population. PSA levels do not decline at the same rate in all patients. Comparing the rate of decline with the rate of increase can provide useful data about survival. A slow rise and quick decline may be considered a good PSA response, whereas a rapid rise and slow decline may be considered a poor PSA response.

Hormone-sensitive disease
In this study, data were analyzed from a cohort of 199 radiation-managed patients from a single institution and a validation cohort of 1,255 patients from 44 institutions who were managed with radiation or surgery. All patients had rising PSA and negative bone scans while on hormonal therapy. In both cohorts, cancer-specific and all-cause mortality were significantly higher in patients who had a rapid rise and a slow decline in PSA levels.

Hormone-refractory disease
This study was a meta-analysis of data from four randomized clinical trials conducted by the Cancer and Leukemia Group B. These trials enrolled 911 patients with hormone-refractory metastatic prostate cancer who received a variety of treatment regimens; 153 patients had 3 pre- and 3-post chemotherapy PSA values that could be analyzed.

Survival in the “best” tertile was 12 months compared with 7 months in the “worst” tertile. However, the proportion of patients who achieved a decline in PSA values of 50% or more was not significantly different in these two groups. Some patients in the “best” tertile benefited from therapy although they did not achieve a 50% decline in PSA. This suggests that the slope, or rate of decline, is more important than the absolute decline in PSA.

Conclusion

Six published studies of PSA doubling time consistently show that a shorter doubling time leads to a shorter interval to a positive bone scan and, in some cases, to cancer-specific death.

Clarification Questions

Dr. D’Amico responded to clarification questions as follows:

- Pre-slope vs. post-slope PSA kinetics were analyzed in the hormone-sensitive patient population. On univariable analysis, the pre-slope and post-slope values and the pre-

slope/post-slope ratio were all significant, but only the ratio remained significant on multivariable analysis.

- The decision to base the analysis on a PSA doubling time of 3 months was made after examining the data for doubling times of 6 months and 12 months and concluding that decreasing the doubling time reduced the ability to identify cancer-specific death.

- Patients for whom data were missing were excluded from analysis.

- Patients with poor results in PSA kinetics analysis died of prostate cancer at a fixed median interval regardless of therapy. However, all of the studies analyzed were performed before any therapy had been shown to improve survival in advanced prostate cancer. Docetaxel has now been shown to improve median survival in this setting. The results of analyses such as those reported here cannot be extrapolated to modes of treatment that were not used in the studies analyzed.

- The nature of retrospective analysis means that it is impossible to be certain that all possible confounding factors have been accounted for. Such a level of certainty can only be achieved in a large randomized clinical trial.

Dr. Albertsen commented that Dr. D'Amico's data suggest PSA doubling time is a surrogate for aggressive disease and therefore a criterion for trial entry. If this is the case, PSA doubling time cannot also be used as a surrogate for treatment response. Dr. Scher agreed.

Dr. Roach observed that PSA doubling time is variable throughout the disease course. He further noted that when hormonal therapy is discontinued, patients' testosterone levels recover and their PSA levels rise correspondingly. In this clinical situation, he said, it is unclear how a rise in PSA due to recovery of testosterone levels would be distinguished from a rise in PSA due to aggressive disease.

Dr. Collette said she interpreted Dr. D'Amico's analysis as showing that PSA doubling time captures the entire effect of initial treatment on survival in patients who relapse but excludes patients who do not relapse. She added that, in her opinion, the use of PSA doubling time as an endpoint in a trial of primary treatment would require that all patients who do not relapse be assumed to have a PSA doubling time of infinity.

Biochemical Failure as a Surrogate Endpoint for Survival in Phase III Clinical Trials of Localized Prostate Cancer (Howard M. Sandler, MD)

Dr. Sandler described results from two trials conducted by the Radiation Therapy Oncology Group, RTOG 86-10 and RTOG 92-02.

RTOG 86-10
This study involved 456 patients with locally advanced disease who were randomly assigned to receive either RT alone or RT plus 4 months of androgen suppression therapy

(AST). AST began about 2 months before the initiation of RT. Because this study was conducted before PSA testing became standard, pre-treatment PSA values were available for only 28% of patients.

About 50% of study participants received salvage AST; of those patients, 70% had PSA-only failure. Patients who received salvage AST did better than those who received RT alone. In terms of both disease-specific survival and overall survival, patients did equally well after salvage AST regardless of whether or not they had received 4 months of AST initially. It was concluded, therefore, that the behavior of prostate cancer after PSA failure was independent of the use of neoadjuvant AST.

RTOG 92-02

This study accrued 1,554 analyzable patients between June 1992 and April 1995. Bilobar, palpable, or more advanced disease was a criterion for study entry. PSA values as high as 150 ng/ml were permitted. Biochemical failure (defined as three consecutive PSA increases, with the time of failure backdated to halfway between the first rise and the last non-rising PSA value) was used as a surrogate for prostate-cancer–specific survival.

Patients received either goserelin and flutamide for 2 months before and during standard RT (short-term androgen deprivation [STAD]) or the same regimen followed by goserelin alone for 24 months (long-term androgen deprivation [LTAD]). The median follow-up period was 5.8 years. At the time of this analysis, 448 patients had died, of whom 142 (32%) died of prostate cancer. A total of 542 patients experienced biochemical failure. PSA failure preceded prostate cancer death by many years.

In terms of the Prentice criteria, treatment was prognostic for both the true endpoint (prostate-cancer–specific survival) and the surrogate endpoint (biochemical failure). Mortality from prostate cancer was lower among patients who received LTAD. This group of patients also had a lower rate of biochemical failure.

However, in terms of prostate-cancer–specific survival from the time of biochemical failure, patients who received LTAD had poorer outcomes than those who received STAD. In Prentice criteria terms, the surrogate was *not* prognostic for the true endpoint because survival from the occurrence of the surrogate depended on the initial treatment.

In summary, this study demonstrated a survival advantage for LTAD; PSA failure was associated with an increased risk of prostate cancer death; and the risk of death from the time of biochemical failure was increased in the LTAD arm. Presumably, some of the patients in the LTAD arm who experienced biochemical failure had hormone-refractory disease at the time of failure and were therefore less responsive to salvage androgen ablation. Surrogate endpoints based on biochemical failure may not be appropriate for phase 3 localized prostate cancer trials if the post–surrogate-failure clinical course is influenced by protocol treatment, as was the case in RTOG 92-02.

An analysis of PSA doubling time among patients enrolled in RTOG 92-02 is expected to be completed later this year.

Conclusions

Caution must be exercised in the use of PSA failure as an endpoint in phase 3 trials when study arms employ differing durations of AST. The results of RTOG 86-10 suggest that PSA failure may be an acceptable endpoint when the duration of AST is only 4 months, whereas the findings of RTOG 92-02 suggest that 28 months of AST may be too long for PSA failure to be an acceptable endpoint.

However, in studies that do not involve the use of differing durations of hormonal therapy, time to disease progression as measured by PSA failure may be an adequate endpoint because PSA failure is the impetus for the use of salvage AST, avoidance of which would be a clinically significant benefit to patients.

Clarification Questions

Dr. Sher said he would be cautious about using PSA failure as an endpoint in any type of study because competing risks mean that very few patients with PSA failure actually die of prostate cancer. Dr. Sandler responded that he based his conclusion on the assumption that local therapy for prostate cancer is clinically beneficial and that the failure of local therapy is an adverse development.

Dr. Kramer said it seemed to be a circular argument to suggest that because hormonal therapy is frequently administered when PSA rises, PSA failure is therefore an adequate surrogate. Dr. Sandler noted that in practice it is extremely common for patients to receive AST following biochemical failure. Data presented by Dr. Albertsen suggested that 5 years after biochemical failure, more than 50% of patients without metastatic disease are receiving AST. Dr. Moul commented that in the Department of Defense health care system only about one-third of patients have started AST within 5 years of biochemical failure, suggesting that economic as well as medical considerations may influence the decision to give AST. Dr. Roach noted that large randomized trials have shown that immediate adjuvant hormonal therapy improves survival. Dr. Raghavan observed that the data are equivocal on whether salvage hormone therapy following radiation failure offers a survival advantage.

Dr. Williams noted that patients in the control arm received AST when their PSA levels rose, creating a crossover effect. Dr. Ellenberg said the use of a surrogate may be an unreliable way to detect a modest difference between two treatments, both of which are known to be effective. Dr. Sandler responded that radiation oncologists have adopted the practice of giving higher doses of radiation therapy because studies have shown that higher doses result in a reduced rate of biochemical failure.

Clinical Trial Experience in Early Prostate Cancer: The Casodex Early Prostate Cancer Clinical Trial Program (Kevin Carroll, MSc, AstraZeneca Pharmaceuticals)

Mr. Carroll began by stating that in his opinion progression-free survival (PFS) and overall survival (OS) can no longer sensibly serve as primary trial endpoints in early prostate cancer. Apart from the disease's long natural history, the widespread clinical practice of treating patients with hormones when a rise in PSA occurs prevents the observation of objective progression outcomes and complicates the use of survival as an endpoint. In the future, he said, only trials with endpoints that reflect clinical practice will be feasible. Such endpoints might include the initiation of further prostate cancer therapy and the use of PSA doubling time as a surrogate endpoint. Absent the ability to use such nontraditional endpoints, it will become extremely difficult to encourage further research and drug development in this disease setting, he said.

Mr. Carroll presented data from the Casodex Early Prostate Cancer (EPC) clinical trial program to illustrate this thesis. He noted that when this program was designed in the early 1990s, both changes in PSA and the initiation of further treatment for prostate cancer were excluded as events. It was envisaged that patients would receive additional therapy only following objective evidence of disease progression.

The objective of the EPC program was to determine the benefit of adding 150 mg of bicalutamide (Casodex) to standard care for patients with early-stage prostate cancer. The program, which is ongoing, consists of three double-blinded, placebo-controlled trails involving more than 8,000 patients in 23 countries, powered to detect an improvement in PFS of at least 15%. Progression was determined by means of bone scan, biopsy, or other imaging technique.

In all three trials, the use of bicalutamide was associated with a significant reduction in the risk of a PSA-doubling (PSA-D) event. In two of the three trials, bicalutamide was also associated with a significant reduction in the risk of an objective PFS event; in the North American trial, however, no difference was shown. This result may be explained by the fact that about 80% of patients who received additional therapy in that trial did so in the absence of an objective progression event. The number of patients who received early hormonal therapy exceeded the number who reached the study endpoint of objective progression. When the data for all three trials are re-analyzed including early hormonal therapy as an event, the reduction in the risk of a PFS event becomes significant in all three trials.

Mr. Carroll reiterated that the key issue in validation of a surrogate endpoint is whether the effect of treatment on an earlier (surrogate) endpoint predicts the effect of treatment on a longer-term clinical endpoint. Buyse and Molenbergs (2000) provide a contemporary method of estimating the effect of treatment on a later endpoint given the effect of treatment on an earlier endpoint. Application of this approach to the EPC data suggests that a hormonal therapy that induces a large effect on PSA-D–free survival is likely to also induce a significant effect on objective PFS, suggesting that PSA-D–free survival could be used as the primary endpoint in early prostate cancer trials.

A concern about the use of PFS as an endpoint is that the precise time of the initiation of tumor progression often cannot be determined. This lack of precision may lead to overestimation of PFS and to bias in the comparison of treatments. One simple approach to resolving this problem might be to compare treatments on the basis of an overall event count over the follow-up period.

This approach would permit treatments to be compared free from the potential biases associated with the timing of progression. An overall event count approach also offers the opportunity to simplify clinical trial design. Further, it can be shown that an overall event count analysis is more powerful if the effect of treatment is delayed. Of course, both treatment arms must be followed equally and there must be no serious imbalances in the numbers of dropouts or the amount of missing data.

In the EPC program, the regular log-rank analysis of PFS time and the event-count analysis (based on the number of events occurring in the first 9 months of follow-up) yield virtually identical results.

Conclusion

Data from the EPC program show that a large treatment effect on PSA-D– free survival is predictive of a positive treatment effect on objective PFS, suggesting that PSA-D–free survival can be employed as a surrogate endpoint in early prostate cancer trials.

Clarification Questions

Dr. Roach commented that if the definition of progression differs for treated and untreated patients, there can be little certainty about the significance of progression. He added that in Canada bicalutamide had been withdrawn from the market as monotherapy for early prostate cancer and suggested that the use of PSA doubling time as a trial endpoint might result in the failure to identify negative effects of a drug.

Dr. Raghavan asked what evidence Mr. Carroll had that the surrogate endpoints he proposed constituted patient benefit. Mr. Carroll responded that the broader issue is that the EPC program, which was designed with objective progression endpoints, was superseded by changes in clinical practice. It is very difficult to design clinical trials when even a very small change in PSA level triggers immediate administration of the most effective treatment available.

Dr. Ellenberg said the data presented by Mr. Carroll could be interpreted to mean that adjuvant therapy is ineffective and unnecessary. Dr. D'Amico responded that several studies have shown a benefit from adjuvant therapy; another interpretation of Mr. Carroll's data could be that salvage therapy is either unnecessary (for the 4 out of 5 patients who will ultimately die of a cause other than prostate cancer) or ineffective (for patients with short PSA doubling times who do die of prostate cancer).

Dr. Mann said it remained important to power trials to identify a survival difference to ensure that treatment did not worsen survival. Dr. Hirschfeld added that it may be necessary to rethink definitions of disease progression but that it is premature to conclude that PFS is not a valid clinical trial endpoint.

Dr. Pazdur said FDA must evaluate whether drugs offer a benefit within the context of current medical practice. He added that in diseases such as breast and colorectal cancer, for which there are many effective therapies, drugs usually have a long history of development in a refractory setting before they are tested in the adjuvant setting. In general there is a greater willingness to accept the risks of a therapy in the treatment of advanced disease when the prognosis is poor and no effective treatment is available.

PSA as Surrogate for Survival in Hormone-Naïve Metastatic Prostate Cancer
(Laurence Collette, MSc, EORTC)

Ms. Collette reported the findings of an EORTC study that used the meta-analytic approach to surrogate endpoint validation (Buyse at al, 2000) to assess various PSA endpoints as potential surrogates for overall survival in metastatic, hormone-naïve (M1) prostate cancer. This approach differs from that taken in the Prentice criteria in that it necessitates replicate estimates of treatment effects on both the potential surrogate and the true endpoint and, thus, large patient numbers, but does not require statistically significant treatment differences on either the true or the surrogate endpoint.

The study utilized data from more than 2,000 patients in the EPC program, excluding those with locally advanced disease. All patients included in the analysis had M1 disease. Ninety percent had a performance status of 0 or 1, although more than 50% presented with more than five bone metastases. The median PSA value at study entry was about 200 ng/ml. The data were subdivided by the patients' country of residence into 21 units of analysis. At a median follow-up of 3.25 years, median survival was 2.3 years; 71.3% of patient deaths were due to prostate cancer.

Four PSA endpoints were assessed: PSA normalization (defined as a decline from baseline to 4 ng/ml or below at two subsequent observations more than 4 weeks apart); time to PSA progression (defined as [1] an increase of >20% above the nadir and exceeding 4 ng/ml and [2] an increase above the moving average nadir of >50% and exceeding 10 ng/ml, which was either the last observed value or was sustained for more than 4 weeks); and longitudinal PSA profile (the complete series of PSA measurements in each patient). The statistical method used permitted direct estimation of the association between endpoints both within a patient (the biomarker/prognostic factor association) and between treatment effects within trials (the surrogate endpoint association).

These analyses confirmed that PSA is a strong prognostic factor for survival in patients with metastatic, hormone-naïve disease. However, none of the tested PSA endpoints was validated as a surrogate endpoint for testing hormonal treatments in M1 disease.

PSA progression defined as a 50% rise to at least 10 ng/ml appeared to be the most promising PSA endpoint, although predictions of effects on overall survival from this endpoint remained too imprecise for it to be useful in new trials.

Clarification Questions

In response to a question, Ms. Collette said that the precision of the data did not allow for the assessment of PSA decline to undetectable levels. Dr. D'Amico commented that none of the PSA endpoints tested in this analysis was time-dependent and asked why endpoints such as PSA velocity were not assessed. Ms. Collette responded that because the date of PSA assessment had to be approximated by the visit date, measures of PSA velocity would not be reliable enough in this database to warrant analysis.

Dr. Tangen suggested that, given there was a non-significant treatment effect in the overall clinical trial, the somewhat artificial subdivision of the data into 21 units of analysis (rather than having data from a number of independent trials, some of which might produced a significant treatment effect) might have made it more difficult to validate the surrogate endpoint.

A question was asked concerning whether a correlation coefficient of 1 between the surrogate and the true endpoint at the individual patient level (i.e., a perfect prognostic factor) would have equated to a valid surrogate. Responding, Dr. Kramer referred to the hypothetical example presented in the morning session by Dr. Baker, in which the surrogate endpoint indicated that Treatment B was superior to Treatment A, whereas the true endpoint (death from prostate cancer) indicated the opposite. Thus, even a perfect correlate may not be a valid surrogate.

Correlation of PSA Endpoints with Overall Survival in Men with Hormone-Naive D2 Prostate Cancer: Southwest Oncology Group Study S8894 (Cathy Tangen, DrPH)

Dr. Tangen presented data from SWOG S8894, an intergroup trial of patients with hormone-naïve metastatic prostate cancer that was conducted between 1989 and 1994. All patients were treated with bilateral orchiectomy and were then randomly assigned, in a double-blind fashion, to receive either flutamide or a placebo. Eligible patients were not required to have measurable disease. Prior surgery or radiation therapy was permitted; prior hormonal therapy, chemotherapy, or biological response modifier therapy was not. Patients with a performance status of 3 were eligible if their performance status was due to bone pain.

The study was not designed for the collection of PSA data; these data were extracted retrospectively from flow sheets. Assay methods were not required to be reported; data were excluded if the assay method could not be determined or if an upper limit of normal of 4 ng/ml could not be verified. Additionally, the schedule of follow-up PSA assessments (prestudy, week 1, month 1, and every 3 months thereafter until progression) was not rigid.

Out of a total of 1,302 randomized, eligible patients, there were 742 who were alive at 9 months and for whom two or more on-study PSA values within the first 6 months and a specified, adequate PSA assay method were available. Seventy-eight percent of patients had extensive disease, 52% had bone pain, 16% had a performance status of 2 or 3, and 13% had a prior prostatectomy. The mean PSA value was 586 ng/ml. Sixteen percent of patients were African American.

Four potential PSA surrogates for survival were assessed: PSA normalization (a PSA level $\leq$4.0 ng/ml or less); PSA response (a PSA decline from baseline of $\geq$50%); PSA progression (an increase of $\geq$50% over the nadir and of at least 5 ng/ml); and PSA slope (slope of the $\log_e$ of PSA from baseline to cutoff).

PSA progression occurred in 392 patients. The median time to PSA progression was 11 months (range 1 to 36 months). PSA normalization occurred in 551 patients. The median time to normalization was 2 months (range 1 to 15 months).

The four potential PSA endpoints were evaluated in all patients at 3, 6, and 9 months post-randomization and the endpoint that maximized the survival difference between the two groups was selected. Nine months was chosen for both PSA normalization and PSA progression. PSA response was not pursued because almost all patients (98%) experienced it.

At 9 months, 72% of patients had PSA normalization and 21% had PSA progression. Median survival was 44 months for patients who achieved PSA normalization, compared with 23 months for those who did not. Median survival was 17 months for patients who had PSA progression, compared with 45 months for those who did not. Among those with PSA progression, half had objective progression (determined by bone scan) within 3 months.

In summary, the PSA endpoints assessed were reasonably correlated with survival. The strongest correlation was observed for PSA progression. These might be interesting endpoints to use in phase 2 screening studies; however, because of the absence of a significant treatment effect, the investigators could not evaluate whether such PSA endpoints would be valid in a phase 3 study. Finally, for many patients, PSA progression precedes objective progression by only a few months.

Discussion

Questions

- How should disease-free survival (DFS) be defined using existing modalities?
- Is the DFS definition different for adjuvant therapy post-surgery than for adjuvant therapy post-radiation therapy?
- Is DFS a surrogate for survival in any setting?
- Is DFS a sufficient endpoint for drug approval in any setting?

Discussion Leaders: Dr. Eisenberger, Dr. Kramer, Dr. Sandler

How should DFS be defined using existing modalities?

Dr. Eisenberger said that most of the large trials to date have been conducted by the Radiation Therapy Oncology Group (RTOG). The two trials that had a no-treatment control arm were both designed in the pre-PSA era and provided little baseline PSA data. RTOG trials have shown differences in favor of hormonal therapy following radiation with regard to biochemical DFS, distant DFS, and local relapse, but no trials conducted in the United States have shown a survival difference. By contrast, in the EORTC trial, combination hormonal therapy for 3 years resulted in a survival difference.

Dr. Sandler noted that long-term follow-up of one of the two RTOG trials, RTOG-8531 (designed in 1985), has now shown a survival benefit for adjuvant hormonal therapy. These results were reported at the American Society of Clinical Oncology and American Society for Therapeutic Radiation Oncology scientific meetings in 2003.

The Messing surgical trial, which began in 1987, included only patients who were node positive at the time of radical prostatectomy (RP), continued Dr. Eisenberger. Further, this trial achieved only about one-third of its projected accrual. Although this was a positive trial, because of its design and limited size additional confirmatory trials are required.

Ideally, the definition of DFS should include an objective measure of disease progression, Dr. Eisenberger said. However, this goal is unattainable given that the standard of care is to begin hormonal therapy as soon as PSA begins to rise and before there is objective evidence of disease progression. Currently, in the United States, 70% of men with a short PSA doubling time receive hormonal therapy before they develop bone metastases. Given this reality of clinical practice, it is necessary for clinical trials to strictly define criteria for changing therapy.

Dr. Eisenberger suggested that DFS might be defined in terms of the number of patients who develop distant metastatic or androgen-independent disease. A measure of PSA kinetics defined by the time of initiation of hormonal therapy would recognize the reality of current clinical practice. However, none of these proposed endpoints has been thoroughly evaluated in clinical trials.

Dr. Pazdur noted that FDA has accepted DFS as an endpoint in breast and colon cancer; however, the agency has specific concerns about the use of DFS as an endpoint in prostate cancer for reasons that have been referred earlier in today's discussion.

Is the DFS definition different for adjuvant therapy post-surgery than for adjuvant therapy post-radiation therapy?

Dr. Kramer prefaced his remarks by commenting that while it may be argued that the perfect is the enemy of the good, the corollary argument is that the good is the enemy of

"good enough." Every constituency has a different perspective on what is "good enough." The Prentice criteria set a high standard for validation of a surrogate endpoint. Dr. Baker's hypothetical example during the morning session demonstrated how a trial result for the true endpoint of death from prostate cancer may be the opposite of the outcome for the surrogate endpoint of elevated PSA. The results of the Women's Health Initiative study of combination hormone replacement therapy in menopausal women show that it is possible to assume for many years, based on widely accepted surrogate endpoints, that a treatment is beneficial when in fact it is harmful.

Addressing the specific question posed, Dr. Kramer said there is no single, commonly accepted definition of DFS in the literature. Because surgery and radiation therapy have very different mechanisms of action, there may be no reason why DFS should be defined in the same way for both modalities. Furthermore, to achieve the goal of boosting accrual to prostate cancer clinical trials by increasing the involvement of community physicians, definitions that are developed in research settings must be tested for validity in the community oncology setting. Dr. Kramer added that definitions that depend on the physician's choice of therapy should be avoided because they are subject to too many continually changing variables and subjective drifts in practice patterns.

Dr. Kramer also expressed the opinion that there should be a better balance between the effort expended on the validation of surrogate markers of benefit and that expended on the identification of surrogate markers of harm. Conceivably, a trial could be stopped early on the basis of a surrogate marker of benefit before an important harm caused by the treatment (which could alter the risk-benefit balance) became apparent.

Dr. D'Amico said that a consistent theme of the presentations heard today was that PSA failure, although it appears to be prognostic, does not meet the higher standard required for a valid surrogate endpoint. He suggested that a DFS composite endpoint might include an objective marker of progression, such as a positive bone scan, in combination with a time-dependent PSA parameter. For example, for hormone-naïve patients with negative bone scans, a positive bone scan and/or a PSA doubling time of <3 months might be an appropriate composite endpoint. For patients who already have positive bone scans, the endpoint might be either bone-scan progression or a PSA velocity >1.5 ng/ml/yr.

Dr. Eisenberger commented that in patients with a PSA doubling time of <10 months, the median time to bone-scan progression is nearly 4 years. However, Dr. D'Amico noted that for patients whose PSA doubling time is <3 months, the time to a positive bone scan is 18 months.

Dr. Scher suggested that focusing drug development efforts on patients whose disease has recurred would produce a higher yield. Another possibility would be to use a discontinuation-of-hormonal-therapy study design. Mr. Kazmierczak, the patient-advocate member of the panel, suggested that different endpoints might be appropriate at different disease stages.

Dr. Albertsen noted that DFS works best as a marker of treatment efficacy in patients with high-risk disease. In patients with early disease and/or disease that progresses very slowly, DFS is a less useful endpoint. However, the time point at which DFS is no longer a valid endpoint has not been defined.

Dr. D'Amico said that the use of a time-dependent PSA construct as an endpoint would require the recruitment of a study population at extremely high risk of death from prostate cancer. Dr. Albertsen responded that it is not clear whether such patients could be identified at a sufficiently early stage in the disease process for treatment to change the outcome.

Dr. Scardino noted that long-term follow-up of patients who received androgen deprivation therapy regardless of PSA doubling time had shown that once biochemical failure occurs, patients have a uniformly poor prognosis. He suggested that consideration be given to the use of a PSA-derived endpoint in a clinical trial in this patient population.

Is DFS a surrogate for survival in any setting?

Dr. Sandler said he had nothing to add to the discussion that had already taken place on this question.

Is DFS a sufficient endpoint for drug approval in any setting?

Dr. Kramer said the short answer to this question is no. He added that the definition of "sufficient" is influenced by whether one is most concerned with being able to offer new therapies to as many people as possible or with protecting as many people as possible from possible harm.

The data presented by Dr. D'Amico, although retrospective, add strength and rigor to the available evidence for the use of PSA-based endpoints to predict a treatment effect, Dr. Kramer added. The next step should be to conduct a prospective study to determine the most appropriate time point for measuring PSA doubling time or velocity. Whatever the time point turns out to be, he cautioned, clinicians will be faced with the difficult issue of having to ask patients to wait to undergo another PSA test in a certain number of months. Even 3 months is a long time to ask a patient to wait, he said.

Dr. D'Amico commented that each patient's repeat PSA tests must be done with same assay to assure that the results of a follow-up test can be reliably compared with those of an earlier test. He added that 3-month PSA doubling time can be estimated using values obtained at shorter time intervals (e.g., 2 months, 2 weeks). Dr. Kramer, however, cautioned that such estimation techniques introduce additional inherent variability.

Dr. Ellenberg said that, in addition to ensuring that surrogates do not cause harm, it is important that they be able to clearly demonstrate that a treatment is providing benefit to patients. Dr. Donald Coffey, from the audience, contended that the greatest harm to patients with prostate cancer arises from the failure to introduce new therapies for the

disease. Dr. D'Amico commented that at advanced stages of disease the likelihood of doing harm to the patient is lower than it might be at earlier stages.

Dr. Eisenberger said that overall survival should continue to be the endpoint for clinical trials in patients with lethal phenotypes of prostate cancer. However, DFS may be a reasonable endpoint for trials in patients who may not have a lethal phenotype but may still be at high risk for developing metastatic disease.

General Discussion and Audience Questions

A patient advocate in the audience observed that patients often see no advantage to enrolling in a clinical trial and that trials must be designed in ways that patients perceive as beneficial. She urged panel members to arrive at a consensus. Dr. Raghavan reiterated Dr. Pazdur's introductory comment at the morning session that the purpose of this meeting was to have a wide-ranging discussion about the positive and negative aspects of various endpoints for trials of drugs to treat prostate cancer. He said the tone of the discussion reflected, in part, panel members' recognition of their responsibility to avoid putting FDA in the position of approving treatments that offer no benefit to patients.

Dr. Pazdur added that it was not the purpose of this meeting to reach consensus. Issues highlighted by the panel will subsequently be discussed at a meeting of the Oncology Drugs Advisory Committee (ODAC), the FDA's statutory advisory body on issues related to oncology drugs. By statute, FDA can take advice related to oncologic drugs only from ODAC.

Dr. Robert Temple, FDA, from the audience, drew a distinction between the use of PSA-derived parameters as trial entry criteria, which he said was uncontroversial, and the use of such parameters as trial endpoints. He suggested that it could be beneficial to further explore how well different degrees of PSA response correlate with the ultimate outcome measure.

Dr. D'Amico responded that it is not known whether there is a PSA level that offers a survival advantage. However, there is evidence that a PSA doubling time of <3 mo or a PSA velocity of <1.5 ng/ml/yr reliably correlates with death from prostate cancer and that a rapid rise in PSA following any intervention indicates a poor prognosis. Dr. Temple said the crucial question is whether there is convincing evidence that intervening to reduce PSA level improves survival.

Dr. Roach said that designing trials with a focus on patients with the shortest PSA doubling times could result in the selection of patients whose disease is too severe to respond to any therapy and exclude patients who might benefit from therapy. Dr. Temple responded that in a variety of disease settings new interventions are tested first in the most seriously ill populations and later in less severely ill groups of patients.

Dr. Scher suggested that a trial of cytotoxic therapy might be conducted in patients refractory to hormone therapy, with a survival endpoint and eligibility criteria based on PSA kinetics.

Dr. Coffey, from the audience, said that PSA kinetics are the best mechanism for distinguishing aggressive from indolent prostate cancer. Most patients have indolent disease and are currently receiving unnecessary treatment. A surrogate that would identify patients who do not need to be treated would be very beneficial, he said. He commented that since 1942 only two new therapies have become available for prostate cancer (mitoxantrone and docetaxel), both of which offer marginal benefit.

Dr. Kramer agreed that identifying the patients who will do well without treatment is a key issue facing the prostate cancer research community. However, he said he was not persuaded that the current knowledge base enables this distinction to be reliably made. He added that because the relationship between PSA kinetics, tumor burden, and survival is not clearly understood, the identification of surrogate endpoints derived from PSA kinetics is not a simple matter.

A member of the audience stated that in his opinion the oncology drug development paradigm of testing agents first in metastatic, refractory disease and later in the adjuvant setting has been an abject failure and is the reason so few advances have occurred during the past 30 to 40 years in the treatment of metastatic solid tumors. He said a new paradigm is urgently needed and questioned whether more extensive surgery or radiation therapy would improve prostate cancer survival rates.

Dr. Sandler, noting that his own practice focuses on the treatment of early disease, said that biochemical failure following local treatment occurs in 25% or more of patients with prostate cancer and is usually the impetus for a change in treatment. Patients would benefit if improvements in surgery could reduce the rate of biochemical failure to 10%. Further, more intense radiation therapy has been shown to improve patient outcomes.

Dr. Raghavan said he believes the existing oncology drug development paradigm has in fact worked very well. He noted that only 1 in 5 drugs that enter testing is ultimately shown to be active. The number of confounding variables in the adjuvant setting complicates the ability to demonstrate treatment efficacy. Bone marrow transplantation with high-dose chemotherapy in breast cancer advanced rapidly and without proper validation from the metastatic to the adjuvant setting, he pointed out, and has since been shown to be ineffective.

Dr. Hirschfeld observed that initiating the testing of new agents at earlier stages of disease could pose ethical problems because patients could be exposed to potentially toxic new therapies when they might have other therapeutic options.

Dr. D'Amico referred to a study conducted by the United Kingdom Medical Research Council, presented at the American Society of Clinical Oncology annual scientific meeting earlier this month (June 2004), which had tested bisphosphonates as adjuvant

therapy in high-risk patients with prostate cancer. Earlier studies in the metastatic, hormone-refractory setting had shown that bisphosphonates delayed the onset of the first skeletal-related event. In the adjuvant trial, however, no such delay was observed. Thus, one cannot necessarily extrapolate results obtained in studies of advanced disease to earlier disease settings.

A patient advocate in the audience asked whether FDA was interested in endpoints associated with prostate cancer prevention. Dr. Pazdur said the agency does intend to convene a panel of prevention experts at a later date to address cancer prevention issues.

A member of the audience drew a distinction between the treatment of individual patients and the design of a clinical study protocol, which is a systematic effort to determine whether a treatment effect exists and for which endpoints must be pre-specified.

The meeting adjourned for the day.