FDA Public Workshop on Clinical Trial Endpoints in Prostate Cancer

June 21-22, 2004 – Bethesda, Maryland

Summary

Tuesday, June 22

Hormone-Refractory Prostate Cancer: Endpoints to Measure Safety and Efficacy of Treatment – A Regulatory Perspective (Bhupinder S. Mann, MBBS, FDA)

Dr. Mann began his presentation with a brief overview of advanced hormone-refractory prostate cancer (HRPC), reviewed the endpoints that FDA has recently accepted for drug approvals in this stage of disease, and discussed the difficulties encountered in reliably measuring the safety and efficacy of treatments for this stage of disease with both traditional and innovative endpoints.

Overview of HRPC

Advanced HRPC is characterized by a variable clinical course, diverse clinical manifestations, and difficult treatment decisions. Traditional endpoints have been of limited utility in evaluating treatment efficacy.

In current clinical practice, patient management and changes in treatment are frequently driven by the monitoring of PSA levels. This has resulted in a fairly large population of patients in whom the only sign of hormone-refractory disease is a rising PSA. These patients have no disease-related symptoms and their bone scans are negative. Although most have a relatively long survival, a few develop rapidly progressive disease and subsequently have a shortened survival.

Given that most patients have no symptoms, have relatively well-preserved performance status (PS) and quality of life (QOL), and relatively long natural survival, acceptable treatment needs to have a highly favorable risk-benefit ratio. Treatment-related toxicity and the treatment's impact on patients' PS and QOL should be minimal. Currently there are no approved treatments for these patients.

Another, similar population of patients with HRPC manifests disease progression by rising PSA levels and positive bone scans. Although many of these patients also have a fairly indolent disease course, some have rapid progression to symptomatic disease. Optimal management of these patients remains undefined as well.

Finally, there are patients with HRPC who develop rapidly progressive disease, with disease-related symptoms, impaired PS and QOL, and shortened survival. The clinical benefit of treatment is well established for these patients.

Endpoints accepted for drug approval in HRPC

Four drugs have been approved for the treatment of HRPC: estramustine (1981), mitoxantrone (1996), zoledronic acid (2003), and docetaxel (2004). The mitoxantrone and zoledronic acid approvals were based on non-survival endpoints.

Docetaxel, in combination with prednisone, was approved in May 2004 on the basis of a significant prolongation in overall survival (OS), demonstrated in a well-controlled clinical trial. OS remains one of the most meaningful endpoints in controlled clinical trials in cancer. Prolongation of OS is a direct measure of treatment efficacy. Additionally, a longer OS provides a reassuring measure of safety because a therapy that causes significant toxicity and possible mortality is unlikely to result in a survival benefit.

Mitoxantrone was approved in November 1996 for use in combination with corticosteroids as initial chemotherapy for pain related to advanced HRPC. Efficacy was shown in an open-label phase 3 controlled clinical trial in which the primary endpoint was palliative response. This endpoint was prospectively defined as a 2-point improvement on a 6-point pain intensity scale, accompanied by a stable analgesic score and having a duration of at least 6 weeks. A palliative response was seen in 29% of patients who received mitoxantrone compared with 12% of patients in the control arm. Median duration of palliative response and median time to disease progression were both significantly longer in patients receiving mitoxantrone. More patients in the mitoxantrone group had PSA-level declines of $\geq$75%. However, no statistically significant difference in survival was shown between the two study arms.

Zoledronic acid, a bisphosphonate, was approved in 2003 for the treatment of patients with progressive bone metastases from prostate cancer. The registration trial used a composite endpoint based on skeletal-related events (SREs). This composite endpoint included the following prospectively defined SREs: pathological bone fractures, spinal cord compression, surgery or radiation therapy to bone to treat or prevent a fracture or cord compression or to provide pain relief. A change in antineoplastic therapy due to increased pain was also included the definition of an SRE. Efficacy was shown by a decrease in the proportion of patients with at least one SRE (33% for patients receiving zoledronic acid vs. 44% for those receiving placebo) and an increase in the median time to the first SRE (NR vs. 321 days).

Issues confounding the interpretation of endpoints in HRPC

The interpretation of OS may be confounded by competing causes of mortality, salvage treatments, and crossover. The use of disease-free survival (DFS) has been suggested as a possible alternative endpoint. The pros and cons of DFS endpoints were discussed at length yesterday.

The use of time to progression (TTP) as an endpoint is difficult in advanced prostate cancer because of the lack of measurable disease. Difficulties in the interpretation of bone-scan findings were also discussed yesterday. For symptom-based endpoints,

clinically relevant treatment effects can be hard to define and lack of blinding can make objective interpretation difficult.

It has been suggested that a TTP endpoint based on PSA measurements might serve as a surrogate for OS. However, a uniform definition of TTP based on PSA derivatives is currently lacking. The validation of PSA derivatives as surrogates for clinical benefit is the subject of this workshop.

As previously noted, in both current clinical practice and clinical trials, treatment changes are frequently driven by changes in patients' PSA levels. Thus, many patients can go off study before any other clinical endpoint of disease progression is reached. As a result, no data on other endpoints of interest may be collected. These data are necessary to define clinical benefit and confirm the validity of the surrogate marker.

PSA-based endpoints may be acceptable surrogates for demonstrating anti-tumor activity (e.g., in phase 2 clinical trials). However, the reliable use of PSA-based endpoints as surrogates for clinical benefit in phase 3 comparative controlled clinical trials remains to be defined. A surrogate endpoint that is "reasonably likely to predict a clinical benefit" can be the basis of an accelerated approval when a new drug provides an advantage over available therapy. However, clinical benefit must be confirmed in subsequent trials.

Endpoints and Trial Design for Evaluating Second-Line Chemotherapy in Hormone-Refractory Prostate Cancer (Mario Eisenberger, MD)

Docetaxel is now the first-line standard therapy for androgen-independent prostate cancer (AIPC). Two large randomized trials (TAX-327, SWOG 9916) have recently reported a median survival of 18 to 19 months for patients receiving docetaxel in combination with either prednisone or estramustine. Median progression-free survival (PFS) for patients receiving first-line docetaxel therapy is probably about 6 months. Most patients remain candidates for further therapy; however, no clearly established standard second-line treatment currently exists.

The conventional approach to developing second-line therapies is to conduct phase 2 screening trials and to follow up with phase 3 trials using survival, PFS, or palliation as endpoints. However, there have been many examples in recent years of positive findings from phase 2 trials failing to be substantiated in phase 3 trials. It is thus difficult to recommend accelerated approval based on the results of phase 2 trials.

For conventional phase 3 studies, improved OS remains the gold standard. Selection of an appropriate control arm is an issue that remains to be resolved. It would be reasonable to use symptomatic treatment such as mitoxantrone plus prednisone, prednisone alone, or second-line hormonal therapy in the control arm. When no clearly established control treatment exists, there is a need to be flexible and creative and to consider innovative designs for registration trials—for example, trials that use internal controls (a + b vs. a vs. b), proof-of-principle trials, trials of combination therapies involving more than one experimental agent, and dose-response assessments.

Dr. Eisenberger said that intuitively he had no difficulty accepting that in patients with extensive disease, in which symptoms are significantly affecting QOL, prolongation of TTP is a reasonable endpoint that constitutes patient benefit. He acknowledged, however, that it is difficult to design trials to prove such benefit. Many different endpoints may be used, some of which are of questionable clinical relevance. For example, when a therapy causes significant toxicity, the significance of prolonging TTP by 6 to 8 weeks is debatable.

Dr. Eisenberger presented findings on PFS in the TAX-327 trial, noting that patient management issues frequently obscure the interpretation of this endpoint in clinical trials. The most common problem relates to the use of multiple criteria to define progression. For example, a patient may demonstrate an improvement in pain and a substantial decline in serum PSA levels, suggesting therapeutic benefit, whereas a protocol-driven bone scan may demonstrate evidence of disease progression. The patient is categorized as having disease progression on the basis of the bone-scan findings, which results in a change in the originally assigned treatment arm, and the other criteria are censored. Frequent censoring occurred with all endpoints.

Complex individual patient management decisions also affected the reliability of progression as an endpoint. Observations were protocol-driven in the sense that when evaluations are repeated frequently, changes will be seen. Although disease progression is certainly a clinically meaningful endpoint, it needs to be adequately defined and evaluations of progression need to be appropriately timed. Clinical trials should be designed with longer intervals between radiological and biochemical assessments.

Finally, Dr. Eisenberger said that the association between PSA kinetics, progression, and survival warrants further study to attempt to define criteria for an accelerated approval.

Clarification Questions

Dr. Sher asked whether the experience gained from TAX-327 might enable a future trial to be designed to reduce early discontinuations. Dr. Eisenberger agreed that there were ways this could be done. He suggested that patients whose condition is stable after 4 to 6 months could be randomized to maintenance therapy vs. no further therapy.

Dr. Raghavan asked whether, in an ideal world, homogeneous populations of hormone-refractory symptomatic and asymptomatic patients could be defined. Dr. Eisenberger responded by noting that many patients enrolled in TAX-327 had no symptoms (e.g., more than 50% had no pain on study enrollment) and that different questions could be asked of these patients than of those with symptoms. For example, a study might be designed in a group of patients with M1 disease who have rising PSA levels after hormonal therapy but in whom bone-scan findings have not worsened; in this population, time to bone-scan progression may be a reasonable endpoint. Dr. Eisenberger added, however, that in TAX-327, a survival advantage was seen in all subgroups of patients who received docetaxel.

Dr. Raghavan said he wondered whether neuroendocrine differentiation could be used to enhance the homogeneity of study populations. He suggested that it could be worthwhile to more closely evaluate the subgroup of patients who present with soft-tissue disease at relapse. Dr. Eisenberger responded that the clinical characterization and significance of patients with the neuroendocrine phenotype requires further study; these patients represent an uncommon subset not necessarily representative of the common form of differentiated prostate cancer.

Dr. D'Amico noted that PFS is flawed as an endpoint because at any given point in time bone-scan findings lag behind the current clinical picture while the PSA level is ahead of the current clinical picture. He said the association with survival needs to be evaluated separately for both bone-scan findings and PSA levels.

Dr. Eisenberger observed that in another study involving patients with no evidence of metastasis (M0 disease), he had piloted the practice of collecting patients' blood at frequent intervals, freezing it, and performing PSA tests on the samples only after 6 months had elapsed. Compliance was generally good in this population; however, in patients with evidence of metastasis and possibly symptoms, this approach needs to be further investigated.

Dr. Moul asked whether TAX-327 shed any light on the question of whether patients should be stratified by mode of primary treatment (surgery or radiation). Dr. Eisenberger responded that in his experience the mode of primary treatment is not a strong prognostic indicator; natural history is a more likely explanation for differences in outcomes. Dr. Tangen said that a multivariate analysis of SWOG-8894 had shown that prior prostatectomy was very significant for survival, but that this association was most likely a marker for something else.

Dr. Kramer noted that Dr. Albertsen had presented data yesterday showing that the attribution of the cause of death can be very subjective. Other data, which Dr. Albertsen did not present, suggest that knowledge of a patient's prior treatment can color opinion about the cause of death; for example, a patient who dies after having had a radical prostatectomy with curative intent is less likely to be classified as dying from prostate cancer.

Dr. Yao commented that although patients could be stratified on the basis of many factors, the purpose of randomization is to balance the distribution of such factors, which is in fact what happened in TAX-327.

A member of the audience said that when bone-scan progression is evident prior to study entry, patients may be denied the opportunity to participate in the study. This could be addressed by agreeing prospectively either to discount bone-scan progression seen at 3 months or to accept as the patient's baseline bone-scan finding the worse of the entry-level or 3-month bone scans.

Dr. Eisenberger responded that it might be possible to account for patients' and physicians' likely choices in the study design. For example, if a patient's PSA level and pain scores had both declined, evidence of bone-scan progression might be more carefully scrutinized.

Disease Progression Endpoints in Hormone-Refractory Prostate Cancer: Experience With the Endothelin Receptor Antagonist Atrasentan (Perry Nisen, MD, PhD, Abbott Laboratories)

Atrasentan is an oral, selective, endothelin A receptor antagonist. Because endothelin receptor inhibition monotherapy is cytostatic, not cytotoxic, tumor shrinkage is not expected. Abbott's pivotal study of atrasentan (intended to demonstrate the efficacy and safety of this agent in men with asymptomatic metastatic HRPC) was designed in 2000 as a monotherapy trial with a placebo-control arm, said Dr. Nisen. No approved therapy for HRPC was available at that time. The use of a placebo-control arm necessitated an open-label extension of the study, obscuring the ability to detect a survival benefit.

A further constraint on the study design was that the sponsor did not expect either patients or physicians to wait long for endpoints to be met if PSA was rising. Given these constraints, Dr. Nisen said, TTP was the only acceptable endpoint. Disease progression was defined as having occurred when a patient achieved *either* radiographic progression (demonstrated by a bone scan showing 2 or more new skeletal lesions or by extra-skeletal progression on a CT or MRI scan using modified RECIST criteria) *or* clinical progression (pain due to prostate cancer requiring escalated therapy, a skeletal-related event, or an event due to metastatic prostate cancer requiring intervention, e.g., a malignant pleural effusion).

Bone scans were scheduled for every 3 months. An independent radiologist determined whether or not progression had occurred after reviewing both bone-scan findings and clinical data. Secondary and tertiary endpoints were PSA level, alkaline phosphatase level, QOL, and survival.

Dr. Nisen summarized the lessons learned from this study as follows:

- Two or more new bone-scan lesions appeared very quickly (by 3 months in most men), especially in the presence of extensive baseline bone disease.
- The secondary and tertiary endpoints were useful to support the individual elements of the composite endpoint.
- The existence of independent radiologists and oncology endpoint reviewers and of a clear decision-making hierarchy was critical.
- A high level of diligence and planning is required to conduct studies that involve composite endpoints, extensive radiological assessment, independent radiologists, and independent endpoint reviewers
- Urologists are major recruiters of asymptomatic HRPC patients.
- Global studies are feasible, although frequent isotopic bone scans present concerns in some countries.

Dr. Nisen concluded by offering the following responses to the questions before the panel:

- A TTP endpoint is feasible using clinical and radiographic criteria.
- Progression can be defined in a way that pre-empts PSA-driven dropout but remains clinically meaningful. Patients can be retained on the study assessment protocol after they are off the study drug.
- Symptom endpoints are not useful as the primary endpoint in a monotherapy study conducted in patients prior to chemotherapy.
- Composite endpoints are useful, but secondary and tertiary endpoints are important for substantiation.

Clarification Questions

Dr. Kramer asked whether Abbott had considered weighting the individual components of the composite endpoint to increase confidence in its overall clinical relevance. Dr. Nisen responded that this had not been considered and would be statistically challenge to do. Dr. Kramer added that reaching agreement on what weight to assign to different components would be a further challenge. The advantage of such an approach would be that patient data would not be completely censored as a result of a PSA rise. Dr. Nisen said that an alternative approach would be to simply count events, as had been proposed by a speaker yesterday.

In response to a question by Dr. Kantoff, Dr. Nisen acknowledged that many of the issues discussed yesterday concerning the time lag before disease progression is evident on a bone scan were manifested in the atrasentan trial. Dr. D'Amico asked whether the time to a doubling of the number of bone-scan lesions for each individual patient had been analyzed. Dr. Andrew Allen, Abbott's global project leader for the atrasentan program, responded that a retrospective analysis had shown a direct relationship between the number of lesions at baseline and the rapidity of further bone metastasis.

In response to a question by Dr. Keegan, Dr. Allen said that the discontinuation rate in the pivotal atrasentan trial had been about 10% in both study arms. However, as a result of extensive efforts by the sponsor to retain patients within the protocol even after discontinuation of the study medication, very few patients were entirely lost to follow-up.

Dr. Keegan asked whether the drug's side-effect profile could have undermined the blinding of the study. Dr. Allen responded that atrasentan is a vasodilator; the most frequent adverse event associated with the drug is headache with rhinitis and peripheral edema, symptoms that are consistent with a respiratory infection. He added that overall the blinding of the study had been well maintained.

In response to a question by Dr. Williams, Dr. Nisen said the primary motivation for the design of the atrasentan study had been to do everything possible to defend against PSA bias, given the trial's complexity and the diversity of participants.

Is PSA Response a Surrogate for Overall Survival? – Analysis of TAX 327 (Martin Roessner, MS, Aventis Pharmaceuticals

TAX-327 was a randomized, controlled study involving 1,006 patients with progressive AIPC in 24 countries, said Mr. Roessner. Patients were randomly assigned to one of three treatment arms: 10 3-weekly cycles of docetaxel, 5 weekly cycles of docetaxel, or 10 3-weekly cycles of mitoxantrone. In addition, all patients received continuous prednisone therapy. Patients were stratified by performance status and pain level. Prior chemotherapy, except estramustine, was not permitted. Median survival in the docetaxel 3-weekly arm was 18.9 months compared with 17.4 months in the docetaxel weekly arm and 16.5 months in the mitoxantrone arm.

PSA response was defined as a decline of $\geq$50% from a baseline PSA value of $\geq$20 ng/ml, confirmed by a second value obtained $\geq$3 weeks later and occurring before progression. Tumor response was defined according to the World Health Organization criteria. PSA response rates were significantly higher in the two docetaxel treatment arms (45.4% for 3-weekly docetaxel, 47.9% for weekly docetaxel) than in the mitoxantrone arm (31.7%). The tumor response rate did not reach statistical significance; however, only 40% of patients had measurable disease and were thus evaluable for tumor response.

Questions addressed in this analysis were whether PSA change correlates with OS, whether PSA change is a surrogate for OS, and whether PSA change can be used as a primary endpoint in clinical trials.

All of the PSA-derived surrogate endpoints that have been proposed have limitations and are more appropriate in certain clinical settings than in others, said Mr. Roessner. One of the main requirements for the use of a surrogate as a primary endpoint in a clinical trial is that all patients must be evaluable for that endpoint. PSA doubling time and time to PSA progression both exclude the initial treatment effect. PSA velocity and PSA response exclude the later progression phase. PSA doubling time, PSA velocity, and time to PSA progression all exclude patients for whom data are incomplete. All current data sets, including the TAX-327 data set, have limitations due to missing data.

PSA response seemed to be the most appropriate surrogate to use in the TAX 327 analysis because it is not affected by missing data later during follow-up; it is censored when tumor, pain, or PSA progression occurs; it accounts for the immediate effect of chemotherapy from baseline; it does not ignore the initial effect (time to nadir); and it includes in the analysis all patients with baseline PSA $\geq$20 ng/ml.

Eighty-five percent of patients had a baseline PSA value of $\geq$20 ng/ml. The median time to nadir ranged from 124 days in the mitoxantrone arm to 164 days in the 3-weekly docetaxel arm. Twenty-five percent of patients had a time to nadir of >7 months, which is significant in the context of a median OS of 16 to 18 months.

In this analysis, PSA response met the first three Prentice criteria for validation of a surrogate endpoint. Treatment had a significant effect on both OS and PSA change (criteria 1 and 2). There was a significant association between PSA change and overall survival (criterion 3). In terms of the final Prentice criterion, PSA response was associated with the treatment effect on overall survival, explaining 49% of the treatment effect. PSA velocity explained 66% of the treatment effect. In an independent analysis conducted by Dr. Sridhara of FDA using the concept of PSA half-life, the proportion of the explained treatment effect on survival ranged from 12% to 21%.

Conclusions

PSA response is a sensitive endpoint for studies of patients treated with docetaxel, but it may not be specific enough to be used as a surrogate for OS. Thus, PSA response may be useful in a phase 1/2 setting to demonstrate drug activity, but it cannot replace OS as a primary endpoint in the phase 3 metastatic setting to confirm ultimate clinical benefit.

More complete data sets are needed to assess the validity of PSA changes as a surrogate for OS. In addition, there is a need for agreement on which PSA-derived endpoints to use and how they are defined. In the metastatic setting, a PSA-derived endpoint should address both the initial response phase and the later progression phase (e.g., time to nadir combined with PSA velocity or doubling time). The choice of endpoint should depend on the specific requirements of the clinical setting.

Agreement is needed on which criteria should be used to validate surrogate endpoints, how the proportion of the treatment effect explained by the surrogate should be calculated, and what proportion of the treatment effect a surrogate must explain to be considered validated. Any valid surrogate endpoint must be sensitive to treatment effects. Finally, Mr. Roessner said, prospective studies must be conducted that include appropriate measurements for the validation of surrogate endpoints.

Three-Month Change in PSA as a Surrogate Endpoint for Mortality in Advanced Hormone-Refractory Prostate Cancer: Data from Southwest Oncology Group Study S99-16 (Daniel P. Petrylak, MD)

SWOG 99-16 was a randomized trial comparing combination chemotherapy with docetaxel and estramustine to the standard regimen of mitoxantrone and prednisone in patients with advanced HRPC, Dr. Petrylak said. He enumerated several differences in terms of therapeutic dosages and duration of treatment between SWOG 99-16 and TAX-327.

Patients were considered to have progressive disease if they had either a bi-dimensionally measurable lesion assessed within 28 days of study registration or evaluable but not measurable disease (e.g., bone scan) assessed within 42 days of registration. The PSA entry criteria were rising serum PSA, with at least 2 consecutive increasing measurements over baseline obtained at least 7 days apart. The minimum PSA value

considered to represent progressive disease was 5 ng/ml. The primary study endpoint was OS. Secondary endpoints were PFS, objective response rate, and PSA decline rate.

The trial began enrollment in October 1999 and closed in January 2003. A total of 770 patients were entered, of whom 96 (distributed equally in both study arms) were subsequently deemed ineligible. The final survival analysis was completed in September 2004. Inclusion of the ineligible patients in the analysis made no difference to the OS rate. PSA values declined by >50% in 50% of patients in the docetaxel/estramustine group compared with 27% of those in the mitoxantrone/prednisone group.

For this retrospective analysis, PSA decline was defined in two ways: any PSA value within the first 3 months that dipped <50% of the baseline value, or the rate of change of log PSA by month determined by linear regression of the log of all PSA values during the first 3 months (PSA velocity).

Patients were eligible for this analysis if their prostatectomy status was known and the following data were available: a baseline PSA measurement within 28 days prior to study registration and at least two post-baseline PSA measurements. A total of 529 patients met these criteria.

Median survival was 18 months for patients in the docetaxel/estramustine group compared with 16 months for those in the mitoxantrone/prednisone group. Thus, treatment had a significant effect on survival. Median survival was 21 months for patients whose PSA values declined by $\geq$50% compared with 14 months for those whose PSA values did not decline by $\geq$50%. The proportion of the treatment effect explained by the surrogate was 47%.

A positive association between the surrogate and survival was also seen within each treatment group. In both the docetaxel/estramustine and mitoxantrone/prednisone groups, median survival was 21 months for patients whose PSA values declined by $\geq$50%. Median survival for those whose PSA values did not decline by $\geq$50% was 15 months in the docetaxel/estramustine group and 14 months in the mitoxantrone/prednisone group. The analyses of PSA velocity also showed a positive association between the surrogate and survival.

Conclusions

Preliminary analysis shows at least one viable surrogate for survival for chemotherapeutic trials of advanced HRPC. Data collection is continuing and subsequent analyses will have better power. Other proposed surrogate measures will be subjected to the same statistical criteria for surrogacy.

Clarification Questions

Dr. Raghavan asked whether the survival benefit identified for docetaxel in both TAX 327 and SWOG 99-16 (a median of two months) is clinically important and whether it

shifts the paradigm in prostate cancer treatment. Dr. Petrylak responded by noting that it was unfortunate that both TAX 327 and SWOG 99-16 had been conducted after the approval of docetaxel and publication of the phase 2 data that formed the basis for that approval. He added that significant crossover in SWOG 99-16 (35% of patients in the mitoxantrone/prednisone group went on to receive a taxane-based regimen) may have narrowed the survival difference, although the extent of the dilution cannot be quantified. Nevertheless, he said he believed it was significant that a survival difference was seen.

Mr. Roessner commented that the hazard ratio may represent the overall trial results more completely than the median survival. In TAX 327, patients treated with 3-weekly docetaxel had a 24% reduction in the risk of death from prostate cancer. In response to a question, Mr. Roessner added that the survival difference between the 3-weekly docetaxel arm and the weekly docetaxel arm did not reach statistical significance.

Dr. Eisenberger said that docetaxel appeared to provide palliative benefits equal to or better than those for which the mitoxantrone/prednisone regimen was approved. It is relatively rare in oncology to achieve a survival improvement in an end-stage disease, he added. The findings of these two trials put to rest concerns that advanced prostate cancer would not respond to chemotherapy and offer reason for optimism that further improvements in survival can be achieved in the future.

Dr. Pazdur said FDA regards any reasonable survival advantage as clinically significant. He added that, in contrast to a “hard” endpoint such as survival, the clinical relevance of a “softer” endpoint such as TTP is easier to question. An intervention that results in a survival benefit may also have other benefits such as an improved response rate or a delay in tumor progression.

Dr. Scher said he believes the TAX 327 and SWOG 99-16 results do represent a paradigm shift in that they demonstrate that it is possible to accrue sufficient patients to prostate cancer trials to demonstrate a survival benefit. The important question now, he added, is what strategies to employ in future studies to build on these findings.

Dr. Scher also commented that PSA change may predict for outcomes other than survival. He asked whether the study investigators had collected sufficient data when patients went off-study to be able to identify potentially clinically significant events. Dr. Petrylak responded that in SWOG 99-16 all PSA values had been collected as well as data concerning patients’ other therapies. Although these data are difficult to interpret, they could be used to generate hypotheses about the efficacy of second-line therapies.

Dr. Sandler observed that the analysis of SWOG 99-16 had used a nonstandard definition of PSA velocity, which is usually defined as the change in PSA over the change in time. He suggested that, to avoid confusion, the investigators use a term other than PSA velocity for the metric they used. Dr. Coffey, from the audience, noted that PSA can be measured in many different ways and that each derivative provides different information. He added that in both TAX 327 and SWOG 99-16, PSA predicted responders from nonresponders.

Dr. Roach commented that a review of the TAX 327 data could lead to the wrong conclusions about the value of chemotherapy. The PSA data from the weekly-docetaxel study arm suggest that this regimen prolongs survival when, in fact, it resulted in a statistically insignificant survival benefit. Thus, the same rule of thumb may lead to a correct conclusion in one study but to an incorrect conclusion in another.

Dr. D'Amico disputed the point made by Dr. Roach. He added that in TAX 327 PSA velocity was a better predictor of survival than a $\geq$50% decline in PSA value. Finally, he suggested an approach that would avoid the possibility of a PSA endpoint leading to the wrong conclusion: Before conducting a phase 3 study, compare the PSA profiles of docetaxel-treated patients with phase 2 PSA profile data for patients treated with the investigational drug.

Discussion

Questions

- How should response and progression-free survival be defined?
    - Role of bone scan
    - Role of PSA and/or its time-dependent derivatives
- How should dropouts due to PSA be dealt with?
- Are symptom endpoints useful?
    - Is time to symptomatic progression a practical endpoint?
- Are composite endpoints useful (e.g., skeletal-related events)?

Discussion leaders: Dr. George, Dr. Kantoff, Dr. Raghavan

Dr. Pazdur asked whether a PSA-derived endpoint could be considered reasonably likely to predict clinical benefit (the standard for granting accelerated approval [AA]). He noted that some drugs approved on the basis of a modest response rate in solid tumors have subsequently demonstrated a survival advantage. He added that as a regulatory agency FDA must continually strive to balance the desire for more information against the risk of doing harm by delaying drug development.

Dr. Kramer said he was satisfied that, for the classes of drug discussed today, a PSA-derived endpoint could come close to fulfilling the Prentice criteria for a valid surrogate endpoint. He suggested that a prospective, randomized study could be designed using a PSA-derived endpoint as the potential basis for accelerated approval; the study should have sufficient power to ultimately show a survival difference.

In response to a question by Dr. Pazdur, Dr. Kramer said that by class of drug he meant that the drugs being compared should have the same mechanism of action. It would be insufficient simply for both drugs to be cytotoxic because there are many classes of cytotoxic drugs.

Dr. Scher noted that a convincing time-to-event endpoint had led to the approval of bisphosphonate therapy, a treatment that has no effect on PSA values. In a variety of studies, efforts to pre-specify criteria for progression had frequently led to the discontinuation of a potentially beneficial drug. The question then became how to balance changes in PSA kinetics with the practical realities of treating patients. He said studies need to be designed in ways that retain patients for a sufficient period of time to enable a time-to-event endpoint to be reached. One approach might be to allow patients to remain on-study if their PSA is stable and they are not developing symptoms.

Dr. Roach prefaced his comments by saying that he believes PSA to be a very useful tool in patient management and that he considers both PSA doubling time and PSA velocity to be important metrics. Nevertheless, he believed more information is needed on how to use PSA correctly as a study endpoint. He noted that if a drug was given on two different schedules or if another drug was added to therapy, it was not clear that PSA change would accurately predict the drug's impact on survival. Dr. Kramer observed that it was because such a scenario was possible that he had proposed the use of a study design that required an established clinical benefit endpoint to ultimately be met.

Dr. D'Amico said that any PSA-derived endpoint must include a time-dependent element (e.g., doubling time, velocity, slope). Initially at least, patients should be followed to a clinically relevant longer-term endpoint to substantiate the existence of clinical benefit. Dr. Eisenberger agreed that a time-dependent PSA-derived endpoint could be used as the basis for AA as long as survival remained the trial's long-term endpoint.

Dr. George commented that the statistical methodology for validating a surrogate has not been agreed on and would vary according to the mechanism of action of the treatment under study. He added that he would be willing to accept a time-dependent PSA parameter as an endpoint for AA, although he expressed concern that the regulations allowing AA make it difficult to withdraw a drug that is ultimately shown to offer no survival benefit.

Dr. Pazdur said the regulations do provide for the withdrawal from the market of a drug that is ultimately shown not to provide clinical benefit. However, there may be many reasons (e.g., crossover, changes in medical practice) why a drug does not ultimately show a survival benefit.

Dr. Scher asked whether and to what extent a PSA change would need to be sustained to be considered reasonably likely to predict clinical benefit. Dr. Pazdur said FDA usually considers both duration and rate of response in its review of a drug. Dr. Keegan commented that the definition of PSA response had differed in each of the studies presented and that only TAX 327 had required that a response be sustained (by stipulating that a response be confirmed 3 weeks later).

Dr. Raghavan noted that FDA has access to numerous fairly comparable sets of randomized data, which presents an opportunity to objectively assess the relationship between PSA change and health outcomes in a large number of patients. Dr. Sridhara,

FDA statistician, enumerated several concerns with the data that had been presented to the panel. All of the findings had been derived from retrospective analysis, she said. Because none of the studies have complete PSA data, conclusions are based on small numbers of patients. Dr. D'Amico's analysis looked at PSA increase whereas the other analyses looked at PSA decline. In Dr. D'Amico's data set PSA was measured at 3-month intervals; however, in metastatic HRPC, PSA is likely to be measured more frequently. Dr. Raghavan responded that additional data exist that have not yet been submitted to FDA and that should be included in the analysis he proposed. For example, in SWOG 99-16, prospective data are still being collected.

Dr. Sridhara suggested that a sensitivity analysis of all the randomized studies conducted to date could offer guidance for a future prospective study as to which PSA parameters seem to be the most promising potential surrogate endpoints. She further noted that the findings for PSA slope and velocity would vary depending upon the period of time that was included in the analysis.

Dr. Ellenberg said she supported conducting an analysis of prospectively collected data from existing randomized studies. She noted that even in a prospective study missing data are frequently a problem. Dr. Moul commented that the data contributed to Dr. D'Amico's analysis from the Department of Defense database had been prospectively collected.

Mr. Kazmierczak, the panel's patient advocate, observed that a leap of faith is needed at some point in order to get new drugs approved. He added that, as a potential candidate for a prostate cancer clinical trial, he would want to be involved in a trial that was based on good science and he was not yet convinced that PSA parameters alone were a sufficient predictor of clinical benefit.

A member of the audience said that most new trials in advanced prostate cancer will likely involve not comparisons of docetaxel to another agent in the taxane class but either comparisons of docetaxel to an agent of a different class or of docetaxel alone to docetaxel plus an agent of a different class. Thus, stipulating that a PSA-derived surrogate endpoint may be used only in trials of drugs within the same class will severely limit the ability to use such an endpoint.

Dr. Kramer responded that his point was that using a surrogate endpoint to compare drugs of different classes introduces additional variability. He said care must be exercised in the choice of a surrogate because risk is involved and it may or may not be possible to withdraw a drug from the market once it has been approved.

Dr. Pazdur commented that the intent of the AA program is to enable FDA to approve drugs that offer an improvement over existing therapy for patients with refractory disease. Drugs have received AA on the basis of response rates of 15% to 20%. Although such response rates appear modest, they most likely result from novel mechanisms of action, given that the patients being treated have in most cases received several previous courses of therapy that have failed to halt the progression of their disease.

Dr. Robert Kane of FDA, from the audience, commented that the underlying problem facing the panel is not regulatory roadblocks to drug approval but the absence of drugs that are more than marginally effective. He added that he was impressed by the analyses that found PSA could explain 50% of the survival benefit of docetaxel and suggested that this might be the best that could be expected of a surrogate endpoint, especially given the impact of competing mortality risks. In response to a question by Dr. Pazdur, he said that for drugs approved by FDA on the basis of response rate, the response rate explained less than 50% of the treatment effect.

Dr. Roach said it was uncontroversial to say that PSA can be useful for anticipating how effective a drug might be, but "the devil is in the details" and consensus is lacking on exactly how, under what circumstances, and in what patient populations to apply PSA parameters.

Dr. Moul asked whether it would be feasible to retrospectively collect pre-treatment PSA values for patients enrolled in TAX 327 and SWOG 99-16 and use that data to produce a pre-treatment/post-treatment PSA slope ratio. Dr. Petrylak responded that in SWOG 99-16 three pre-treatment PSA values were collected for all patients. Dr. Eisenberger said pre-treatment PSA data are also available for patients enrolled in TAX 327. Dr. Nisen said that Abbott is conducting a prospective, randomized, placebo-controlled trial in high-risk, hormone-naïve patients in which the primary endpoint is change in PSA doubling time. Data from this trial will be available in 2005.

Dr. Pazdur asked how the pre-treatment/post-treatment PSA slope ratio could be used in a trial. Dr. D'Amico proposed that in a randomized trial, PSA values should be collected at specified intervals before treatment (e.g., 6 months, 3 months, and baseline) and after randomization. The rapidity of action of the drug being tested (e.g., whether it has an immediate or a delayed effect) should be taken into account in determining the appropriate intervals for post-randomization PSA testing. Subjecting the ratio (R) of the pre-treatment and post-treatment slopes to a statistical test would determine whether the value of R was significantly different between the two arms of the trial.

Panel members discussed other approaches to incorporating PSA parameters as trial endpoints. Dr. D'Amico suggested that if it was considered important to capture the duration of a PSA response, a two-component variable could be constructed that consisted of the magnitude of the PSA velocity and the duration of the response. Dr. Williams noted that it would be important to design the analysis to include all patients (i.e., those whose PSA values rose or remained stable as well as those whose PSA values declined).

Mr. Roessner commented that the pre-treatment PSA slope is a prognostic variable that should be taken into account in calculating the effect of treatment. He added that 75% to 80% of patients in TAX 327 experienced an initial PSA decline that was followed by a gradual return to the baseline value. He suggested that the time it took for PSA to return

to the patient's baseline value might be another endpoint that could be examined for a possible relationship with survival.

Dr. Kramer expressed concern that consensus was lacking on which of a large number of PSA parameters would make the best surrogate endpoint. Dr. Pazdur said that a sponsor intending to use a PSA-derived endpoint would be required to prospectively define that endpoint. The recommended design of a trial intended to be the basis for AA is a randomized study that incorporates an interim analysis based on the prospectively defined surrogate endpoint and a final analysis based on an established clinical benefit endpoint.

Dr. Raghavan observed that it may be unrealistic to rely on a single predictive marker for a solid tumor. He suggested that tissues and serum should be prospectively collected for later analysis for pre-specified markers. Otherwise, heterogeneity may be overlooked and data misinterpreted.

A patient advocate in the audience commented that the reality for patients is that treatment of prostate cancer is "one big clinical trial." She endorsed Dr. Coffey's earlier comment that the greatest harm to patients with prostate cancer arises from the failure to introduce new therapies for the disease. She urged panel members and FDA to refrain from taking an overly dogmatic approach.

Dr. Pazdur responded that the purpose of the AA program is to speed the introduction of new, potentially beneficial drugs in life-threatening diseases. He noted, however, that the quality and amount of evidence required for AA is not different from that required for regular approval. The applicant must show substantial evidence of the measured effect from well-controlled clinical trials; borderline evidence is not acceptable. The difference is that the evidence may focus on a surrogate endpoint that is only *reasonably likely to predict benefit* rather than on an accepted clinical benefit endpoint.

The panel did not directly respond to the questions posed for discussion at this session. The consensus was that these questions had been addressed in the course of the general discussion. In response to a question from Dr. Roessner, Dr. Williams summarized the views expressed by panel members with regard to the role of bone-scan findings as follows: Evidence of bone-scan progression must be confirmed by other findings such as, at a minimum, a rising PSA level. Additionally, further thought needs to be given to how to define disease progression on the basis of the magnitude and extent of changes in bone lesions.

In conclusion, Dr. Pazdur said that FDA will ultimately present a summary of the discussions at this workshop to the Oncology Drugs Advisory Committee, the agency's statutory advisory body on issues related to oncology drugs. He reminded members of the audience that they were welcome to submit comments to him by electronic mail after the meeting. He thanked panel members and members of the audience for their participation in the workshop.

Addendum

Dr. Mario Eisenberger submitted the following additional comments after the workshop:

The relatively common practice of initiating androgen deprivation in patients who present with rising serum PSA levels as the only evidence of disease (M0, hormone naïve) has resulted in a growing patient population in whom the only sign of active disease is a rising PSA (androgen-resistant, castrate, nonmetastatic disease). These patients have no disease-related symptoms and their bone scans are negative. Although most have a relatively long survival, their natural history remains poorly defined. Sequential utilization of secondary endocrine manipulations is the most common therapeutic practice. Clinical trials in this patient population are clearly needed. At the present time, the only large study addressing treatment questions in this population is ECOG 1899, which compares a commonly used second-line hormonal manipulation (ketoconazole + hydrocortisone) with docetaxel-based chemotherapy. The endpoint of this study is time to development of clinically evident bone metastasis.

In patients with androgen-independent castrate (M0) disease, prolongation of time to clinically evident distant metastasis is most likely a clinically significant endpoint as long as treatment-related toxicity does not impact negatively on quality of life.

Another population of patients with hormone-refractory prostate cancer manifest disease progression by rising PSA levels and positive bone scans. The outcome of patients with hormone refractory disease and clinically evident metastasis (M1 disease) remains poor. Current data suggest that time from PSA progression to bone scan progression is relatively short (6 months or less). These patients may be candidates for chemotherapy since the benefits of second-line hormonal therapies are relatively modest and of short duration.

In TAX 327, the pivotal study that served as the basis for the approval of docetaxel, in combination with prednisone, for the treatment of hormone-refractory prostate cancer. subgroup analysis indicated that docetaxel + prednisone offered a significant survival benefit in all subgroups according to age, presence or absence of pain, and Karnofsky performance status of 70 or worse/80 or better. Furthermore, secondary objectives demonstrated better pain control and improved QOL for the docetaxel arms. Treatment was associated with acceptable toxicity.