EVALUATION AND FOLLOW-UP OF PATIENTS WITH N1-3 M0 OR NxM1 PROSTATE CANCER IN PHASE III TRIALS

REGINALD HALL, PER OLOV HEDLUND, ROLF ACKERMANN, NICHOLAS BRUCHOVSKY, OTILIA DALESIO, FRANS DEBRUYNE, GERALD P. MURPHY, MAHESH K.B. PARMAR, MICHELE PAVONE-MACALUSO, MIRJA RUUTU, AND PHILIP SMITH

ABSTRACT

Objectives. The aim of this discussion is to review the design and conduct of phase III trials in metastatic prostate cancer, to seek ways of improving their study design, accuracy, relevance to clinical practice, acceptability to patients, and ease of participation by clinicians. We also aim to try to set uniform definitions for the evaluation of the different endpoints used in clinical trials on metastasized prostate cancer.

Methods. The work was started by correspondence between the participants in the group for the year before the consensus meeting. Two comprehensive questionnaires were circulated and the answers were distributed to all the members of the group. The statements were finalized during the consensus meeting.

Results. There were some differing opinions concerning the methods of evaluation of endpoints for follow-up, such as time to tumor progression and time to treatment failure. After the consensus conference, there were no major disagreements within the group.

Conclusions. The aim of phase III trials is to influence clinical management. To obtain a credible result they require a sound statistical basis with appropriate power and encompassing patients from small urologic practices as well as large or academic institutions. However, deviation from routine practice may affect the accrual rate, and the trial procedure should therefore be as similar as possible to routine management. Trials inevitably involve extra work and cost. Both should be kept to a minimum to encourage participation and hasten a timely conclusion. It is mandatory to create uniform ways of designing and evaluating clinical trials in prostate cancer.

UROLOGY 49 (Suppl 4A): 39-45, 1997. ©1997 by Elsevier Science Inc. All rights reserved.

The aim of this discussion is to review the design and conduct of Phase III trials in metastatic prostate cancer for the purpose of seeking ways of improving their study design, accuracy, relevance to clinical practice, acceptability to patients, and ease of participation by clinicians. The benefit sought from a new treatment should be defined at the outset of a trial. Benefit may be an increase in survival, time to progression, or quality of life, or a decrease in the toxicity or cost of treatment. The purpose of Phase III trials in metastatic prostate cancer is to compare the efficacy of 2 or more forms of treatment or patient management in patients who have received no previous treatment for their cancer. Comparison of the treatment's impact on survival is the usual endpoint, but other endpoints, such as disease progression, may be appropriate. Measurement of response of evaluable disease parameters is not usually a subject of study for phase III trials. Demonstration and measurement of response to treatment is the role of phase II studies. The planning of phase III trials is discussed elsewhere, but it is important that study design should include, from the outset, the status of the patients, the extent of disease, and
stratification for prognostic factors of proven significance. The planned size of the trial will need to take into consideration the magnitude of the benefit sought, the number and time scale of expected events, and the anticipated time to accrue the necessary number of patients.

A clear understanding of the information required to answer the question posed by a trial is essential when the study is being designed. The parameters measured and recorded in phase III trials are those needed to determine the chosen endpoint. The frequency of their assessment reflects the expected timing of critical events and the expected duration of survival. These are not necessarily the same as needed for routine clinical management.

The aim of phase III trials is to influence clinical management. To obtain a credible result, phase III trials require a sound statistical basis with appropriate power encompassing patients from small urologic practices as well as large or academic institutions. However, deviation from routine practice may affect the accrual rate and should therefore be as similar as possible to routine management. Trials inevitably involve extra work and cost. Both should be kept to a minimum to encourage participation and hasten a timely conclusion.

The treatment of metastatic prostate cancer is palliative for the large majority of patients, and as a consequence, time spent without symptoms from the disease or side effects of treatment is important. Asymptomatic patients with distant metastases do not require hormonal treatment for symptomatic progression for 9 to 12 months. For most patients with symptomatic metastases, progression after hormone treatment tends to occur after 12 to 18 months, but for those with “minimal disease” it may be 3 years or longer. In those with lymph node, but no distant, metastases, progression may not occur for several years, and median survival is considerably longer than 5 years. The inevitable delay that this causes in obtaining the result of a trial has led to a search for alternative, shorter, surrogate endpoints.

Tumor progression, treatment failure, and rising serum prostate-specific antigen (PSA) level have all assumed increasing importance in phase III trials. The time from entry into study to the occurrence of 1 or more of these events has been used as a basis for analysis. A rise in PSA causes no physical symptoms but may cause anxiety. Similarly, the development of new hot spots on the bone scan or other radiologic evidence of disease progression is of relatively little concern to the patient unless associated with pain or other symptoms. Time to treatment failure is probably less important to the patient than symptomatic disease progression.

Each of these endpoints provides data that increases our understanding of prostate cancer and helps to refine patterns of care. They are therefore useful endpoints for study in phase III trials. However, the overriding desire of the patient is increased survival and a reduced likelihood of dying from prostate cancer.

ENDPOINTS

Overall survival is the most robust endpoint for phase III study and has the greatest impact for the patient. It is the most accurate, objective, and easiest to measure. Quality of life, freedom from symptoms or unwanted side effects of treatment, and absence of overt disease progression are all important to patients with prostate cancer, but the eventual aim of treating prostate cancer is to prevent death from this disease and its treatment. The cause of death (disease-specific survival) is therefore relevant.

Duration of survival is not just determined by the first treatment, but is made up of several periods of disease control achieved by different treatments. At present, second- and third-line treatment (after hormone relapse) contribute relatively little to overall survival, but may become more important in the future when effective, nonhormonal treatments become available. Thus the duration of benefit from individual treatments may be important to measure. For this, 2 different time periods can be used: time to tumor progression (TTP) or time to treatment failure (TTTF).

For asymptomatic patients, delayed treatment is still an option. Time before commencement of treatment needs to be measured in trials that incorporate this management concept.

SURVIVAL

OVERALL SURVIVAL

Overall survival is very relevant because most men with metastatic prostate cancer will die within the time scale of the study; the majority of prostate cancer. Mean age at diagnosis has been more than 70 years, but appears to be falling and comorbidity has less influence on outcome.

The pattern of care provided by health services varies between countries. Long-term follow-up by the institution that initiated treatment is not always possible and the precise date of death of patients may not be known. However, many countries now have national registration of mortality and participants in phase III trials should be required to obtain confirmation of the date of death from such registries if necessary.

An overall survival benefit should only be considered to be significant clinically if accompanied by an equivalent improvement in disease-specific survival.
DISEASE-SPECIFIC SURVIVAL (DSS)
(CAUSE-SPECIFIC SURVIVAL)
In principle this is the most desirable endpoint but its value is undermined by the difficulty experienced in its definition.
For many men with prostate cancer, it is evident that this disease has been the cause of their death, but for others the exact cause of death is uncertain. Death certificates are not always reliable. Patients with clinical signs of progressive cancer immediately prior to death should usually be counted as cause-specific deaths. Agonal events, such as pulmonary embolism, major hemorrhage, pneumonia, or sepsis, occurring in a patient who is severely ill in the final stages of progressing prostate cancer, should not be identified as the cause of death. These patients should be considered to have died of prostate cancer. However, it should be remembered that even when a patient has progressive cancer, death may be due to a sudden and entirely unrelated cause, such as cardiovascular disease. Elevated PSA alone, in the absence of symptoms or signs of disease progression, is not sufficient to indicate prostate cancer as the cause of death.
It has been suggested that the terms disease-specific survival or cause-specific survival should include only those deaths that are directly attributable to prostate cancer, and not include deaths due to treatment. This definition is rejected. Death due to treatment toxicity, and therefore a consequence of the disease, should also be considered as death due to the cancer for the purpose of calculating disease-specific survival.
The cause of death of a patient in a clinical trial has to be decided by the clinical coordinator, after appropriate inquiry to answer the question “was the patient's death directly, or indirectly, a result of prostate cancer or its treatment?” In addition, it is recommended that phase III trial organization should include an independent “Cause of Death Committee” if DSS is a major endpoint.

DISEASE-FREE SURVIVAL (DFS)
Complete pathological remission of metastatic prostate cancer is rare. Thus “disease-free survival” is not relevant as an endpoint for phase III study in this disease.

RELATIVE SURVIVAL
The endpoint of “relative survival” cannot be applied to phase III trials because survival should only be compared within the randomization of the trial.

LETHAL ADVERSE EVENTS
Lethal adverse events following treatment should be uncommon in phase III trials, but all treatment-related deaths should be identified and reported as a subgroup in survival analyses. They should also be included in DSS calculations.
Since adverse events may be difficult to identify, all events that cannot be related to an expected progression of the malignant disease should be reported and registered, together with evidence to justify the conclusion. The detailed reporting of adverse events as currently required for registration studies of new drugs is burdensome. For non-serious adverse events, it is not necessary for most phase III trials. For serious adverse events, a simpler format of recording these events is required for future phase III trials.

DISEASE PROGRESSION
Two types of disease progression, “clinical” and “biochemical,” are recognized for the purposes of phase III trials.12,13

CLINICAL PROGRESSION
The progression of prostate cancer takes many forms, is often difficult to document until far advanced, and may not be evident until late in the patient’s disease pathway. In a trial, objective evidence of disease progression is usually considered desirable.13
The following events are considered to be of particular importance for patients with progressive prostate cancer, and, in a phase III trial, are to be regarded as evidence of clinical progression: pain, anemia, symptoms or signs of upper tract obstruction, urinary symptoms, malignant lymphadenopathy, lymphoedema, symptoms or signs of spinal cord compression, pathological fracture, or decline in performance status due to prostate cancer.
It is recognized that anemia may be treatment-related rather than signifying progressive disease; prolonged androgen deprivation may result in a drop in hemoglobin of 2 to 3 g/L. In judging disease progression, it is important to ensure that the above factors are related to the malignancy rather than other intercurrent disease or treatment toxicity.
Imaging at regular, predetermined intervals has been used in many previous trials. This experience has revealed that the interpretation of asymptomatic new hot spots on bone scans, with or without x-rays, can be difficult,14 and asymptomatic lung or lymph node metastases seldom warrant treatment. Bone scan is mandatory at entry to phase III trials and is commonly used to evaluate progression. Chest x-ray, computed tomography (CT) or magnetic resonance (MR) scans are reserved for the investigation of symptoms. They will have cost implications if required specifically for a clinical trial. Although they provide useful evidence to support the diagnosis of clinical progression, it is recommended that the planned use
of these investigations at regular intervals is not essential for phase III trials. This is not to say that disease progression is unimportant and should not be recognized in trial design. Because the diagnosis of tumor progression depends on subjective factors and clinical acumen as well as objective measurement, the definition of "clinical progression" cannot be precise. The following definition of clinical progression is recommended: new symptoms or signs, or worsening of those already existing, clearly attributable to prostate cancer.

**Biochemical Progression**

Several biochemical serum parameters have been shown to have prognostic significance when measured at the time of diagnosis or entry into a trial and may also indicate disease progression before this is evident clinically. Those of prognostic significance need to be measured at entry into the trial, but only require repetition during follow-up if they also act as indicators of progression.

It is agreed that prostate-specific antigen (PSA) is the most reliable measure of disease activity currently available and should be measured at regular intervals. A rise in PSA has been shown to predate radiologic and symptomatic disease progression by many months and is generally accepted to be the most reliable first indication of treatment failure in the majority of patients. It is agreed, therefore, that "biochemical progression" based on rising PSA is an endpoint of considerable interest. Furthermore, an association between PSA "response" and survival is becoming apparent, although long-term data are as yet insufficient to support the use of serum PSA as a surrogate endpoint of proven value.

It should be recognized that PSA expression is hormone-dependent, and a considerable decrease may not necessarily represent significant disease regression. Furthermore, some antitumor effects from nonhormonal treatment may not lead to a decrease in PSA levels. It should also be remembered that PSA almost certainly will be replaced by a better biochemical marker in the future.

In patients with advanced prostate cancer, a rise in serum creatinine usually indicates ureteric obstruction; elevated alkaline phosphatase suggests either progressive bone or liver metastases; and a fall in hemoglobin probably, but not necessarily, indicates bone marrow involvement. Each of these biochemical or hematologic measurements contributes to the diagnosis of clinical progression and should be measured during follow-up, but as individual parameters, they do not amount to "biochemical progression." A fall in PSA to normal or near-normal following protocol treatment is a significant prognostic factor for the duration of survival. It may therefore be used to stratify patients as a determinant of intermittent treatment or randomization to a second treatment. Fall in PSA cannot be used to stratify for analysis of outcome. In the future, prostate-specific membrane antigen (PSMA), neuroendocrine markers, radio-labeled antibodies, and other methodologies may provide additional or alternative indicators of progression. For the foreseeable future it is recommended that biochemical progression should be defined as a sustained and continuing rise in serum PSA. The percentage change in PSA or the threshold PSA level will need to be defined for individual trials. In some patients a fall in PSA is not observed. For these a "no change" category may need to be defined.

**Time to Progression (TTP)**

Progression, clinical, biochemical, or both, is a useful endpoint in phase III trial design. The time at which this event occurs therefore requires definition. As evidence of progression, it is the time at which PSA rises (after falling to a nadir, or remaining unchanged after treatment) that is relevant. It is proposed that time to tumor progression be defined as the time at which clinical or biochemical progression is first observed as defined above.

**Time to Treatment Failure (TTTF)**

This endpoint is similar to TTP but also includes, the "failure" of treatment due to toxicity as well as treatment failure manifest by cancer progression. For the comparison of treatments that can be stopped, it has greater practical relevance than TTP.

TTTF differs from TTP in that prostate cancer may remain unchanged, or even improved for a considerable time when treatment is withdrawn. As a consequence, TTTF will usually be shorter than TTP when applied to the comparison of 2 different treatments.

TTTF is defined as the time from entry into the trial to clinical or biochemical progression or withdrawal of treatment because of treatment toxicity, or death, whichever is the shortest.

**Symptom-Free Interval**

The management of some patients by careful monitoring and delayed hormone treatment or by intermittent hormone therapy creates another clinical scenario that requires consideration in phase III trial design. This concept differs from immediate, continuous treatment because it includes periods of time when the patient receives no treatment and is asymptomatic. These periods comprise freedom from the morbidity of prostate cancer and may be single (before delayed therapy) or multiple (within intermittent therapy). They
are important because they address the problem of disease morbidity and quality of life.

If the symptom-free interval is to be used as an endpoint, it will be important to avoid bias from the tendency to treat the patient sooner if an "effective" treatment is still available compared with the inevitable reluctance to investigate symptoms of possible progression if further treatment choice is limited. Whether treatment is delayed initially or given intermittently, it is the total time free of symptoms as a proportion of overall duration of survival that is relevant. The symptom-free interval is defined as the total period of time, within the trial protocol, during which the patient has no symptoms attributable to his prostate cancer.

FOLLOW-UP

Investigations

The investigations to be performed and recorded during the follow-up of patients in phase III trials are those required to document the following: compliance with treatment, toxicity, and any adverse events related to treatment, symptoms of prostate cancer, progression of prostate cancer (clinical or biochemical), and death or survival.

Requirements for phase III trial follow-up may not be the same as those for clinical practice, particularly as the latter appears to differ according to institution or country. There is therefore a need to agree on a minimum data set for follow-up that should be required for all phase III trials. These comprise quality of life (including urinary symptoms), pain due to prostate cancer, performance status (change due to prostate cancer), clinical examination, serum PSA, serum creatinine, hemoglobin, alkaline phosphatase, bone scan or other imaging techniques as indicated for assessment of progression, and survival. Other blood tests or imaging will be employed as indicated clinically, but are not to be included in the minimum data set.

Frequency of Follow-Up

To facilitate patient recruitment and clinician participation, and to minimize unnecessary work and cost, follow-up visits in phase III trials should coincide with clinical management whenever possible. Currently, these tend to be at 3-monthly intervals for the first year, progressing to 6-monthly and possibly annually. However, this pattern of follow-up visits may not be the most appropriate for the likely time scale of important events. If TTTF or TTP are study endpoints, frequent evaluation is necessary during the first year to ensure treatment compliance and detection of toxicity. Thereafter, the interval between follow-up visits should be designed according to the expected duration of survival in a particular study and the anticipated timing of critical events, such as tumor progression. For trials evaluating intermittent therapy more frequent visits will probably be necessary. The data to be collected on these additional visits should be limited to that upon which treatment decisions are based, namely, PSA (or another marker) and quality of life questionnaire. Whatever pattern of follow-up is employed, it is essential that the frequency of visits is the same in both or all arms of the trial.

It has been proposed that, for trial purposes, data sheets need only be completed at annual intervals to record data concerning critical events retrospectively, but leaving the frequency of follow-up to individual preference. The advantage of this trial design is its simplicity and savings in data management costs. Such trials minimize the amount of work required specifically for the trial and, in their simplest form, could amount to no more than the recording of the date of death from national cancer registries. The main disadvantage of this trial design is that the recorded time to critical events depends upon the frequency of the follow-up visits and excessive variation in this will undermine the quality or strength of the endpoint in question. Reliance solely on annual follow-up is recommended only for trials that address the single endpoint of overall survival, with other endpoints being relatively unimportant. This will rarely be the case for patients with metastatic disease.

Retrospective Sequential Method

Although trial protocols encourage the definition of clear endpoints and the date of occurrence of certain events, experience teaches that such precision is often lacking in everyday practice. Typically, the clinician suspects that disease progression is occurring, but the definite diagnosis of disease progression can only be made after several visits, over a considerable period of time, or when the results of a number of investigations become available. The definition of TTP or TTTF has been a cause of considerable difficulty in many previous trials. Inevitably, the diagnosis of treatment failure or disease progression is made retrospectively and usually in a sequential way. It is therefore recommended that the retrospective sequential method be used in future phase III trial designs, whereby the first signs of deterioration in any parameter are noted and the finding is followed consecutively (at the predetermined follow-up intervals) for confirmation of progressive disease or for cancellation, if tumor relapse is not confirmed. The first date of suspicion of progression (the date of first rise in PSA, the onset of pain, or the first fall in hemoglobin) is eventually counted as the time point for TTP or TTTF.
**QUALITY OF LIFE**

Quality of life (QOL) evaluation is complex and both questionnaires and their analysis may appear to take a disproportionate amount of time. For this reason, they may not be considered to be necessary for all phase III trials. For phase III trials in which TTP or survival are main endpoints, QOL measurement may be necessary for a selected proportion of the trial population only. Nonetheless, if QOL is to be studied, this decision should be made when designing the trial and arrangements made for QOL assessment in such a manner as to ensure that it is recorded at the same time as all other critical parameters. Quality of life evaluation is a relatively new field of research and its contribution to clinical decision making and disease management has yet to be established. Since in metastatic prostate cancer treatment offers palliation rather than cure, the quality of the remaining life is of considerable importance.

**STATISTICAL CONSIDERATIONS**

The objectives of a phase III trial should be achievable and applicable to clinical practice. Analysis of the results of phase III trials is a specialist statistical issue, but one that requires clinical input. The statistical principles for the analysis of phase III trials have been well defined and there is a considerable body of international expertise to advise on the design of future trials.\(^4,5,23-26\)

Avoidance of bias by the use of proper randomization methods and analysis on the basis of “intention to treat” is fundamental. Trial design based on an awareness of relevant prognostic factors and appropriate stratification is also essential. Evaluation of outcomes is a statistical exercise, but “clinical significance” is at least as important as “statistical significance” if the results of phase III trials are to influence future clinical practice. Cancers differ in their biological behavior, their response to treatment, and the duration of their overall “patient pathway.” Wide variation is often seen in prostate cancer, which means that all phase III trials will not be identical. Rather, the choice of endpoint, the pattern of follow-up, and the treatment will depend on the category and extent of prostate cancer to be studied.

For the immediate future, modifications of currently available treatments may improve survival, but differences between treatments will probably be fairly small. A meta-analysis of trials of maximum androgen blockade suggested nonsignificant differences of only a few months\(^27\) and it is clear that some colleagues consider such small differences to be of little practical importance. To assess mortality differences of only moderate size, studies of thousands of patients are necessary to avoid the play of chance.\(^26\)

To ensure that trials have clinical significance as well as statistical validity, participants should decide at the outset the magnitude of benefit (increase in survival, TTTF, time free of symptoms) that would lead to the routine use of the new treatment. The statistical design and the number of patients will be determined by these clinical considerations. If the desired benefit is larger than the benefit that can be expected on the basis of epidemiologic or pharmacologic data, the study should not be started.

The lack of curative treatment has led to a greater emphasis on quality of life, freedom from symptoms, and freedom from treatment side effects, based on the assumption that survival will be the same for the treatments compared, that is, the “equivalence” of survival. For example, if a new treatment is found to be less toxic, the trial needs to show reliably that the survival obtained is not compromised compared with the standard treatment. In practice, proof of absolute equivalence is not feasible. Thus, if phase III trials in future are to be of this nature, there is a need to define a limit to the difference in survival that is acceptable as “equivalent.” For example, assuming a median survival of 36 months for M1 patients, would a new treatment be considered beneficial if symptoms and side effects are reduced but median survival is also reduced? If so, by how much? Three, 6, 9, or 12 months? Indeed, is any reduction in survival acceptable? Conversely, what increase in survival is considered clinically significant for a new treatment? Does the answer depend upon treatment toxicity or the cost of the treatment in question? Answers to these questions will undoubtedly be contentious, but if controversy is to be avoided when the results of individual trials or meta-analyses are published, consensus guidelines should be agreed to as a basis for future trials.

For patients with a medium or good prognosis and expected median survival of 3 years or more, a decrease in survival of up to 6 months, but no more, is suggested as an acceptable basis for the statistical definition of equivalence. A 6-month
"trade-off" is also suggested as the maximum acceptable, even when longer survival is expected. For poor prognosis patients whose median survival is only 18 months, a possible reduction of up to 3 months might be accepted by patients, depending on the advantages to be gained from the new treatment.

Clinicians are uncomfortable with the idea that any new treatment that may reduce survival could be described as "acceptable." Nonetheless, as already recognized, proof of absolute equivalence is not feasible. Thus, for clinical trial purposes, definitions of equivalence will be essential to safeguard patients and ensure that the possibility of trial investigators accepting a large survival discrepancy is excluded by the trial design. The need for such definitions is a reflection of the palliative nature of current treatments for metastatic prostate cancer.

In the future, new treatments will become available that offer cure rather than palliation. These may replace or be combined with hormone treatment. This consensus document seeks to set guidelines that will be relevant to such future situations. Overall survival will be the main endpoint for these future phase III trials.

ACKNOWLEDGMENTS. We are particularly grateful to Dr. P.O. Hedlund for his organization of the discussion upon which this manuscript is based. We also thank Drs. W. Kiebert and R. Sylvester of the European Organization for Research and Treatment of Cancer (EORTC) Data Centre Brussels for their contributions.

REFERENCES


